Recursion Theory Notes, Fall 2011

Lecturer: Lou van den Dries

0.1 Introduction

Recursion theory (or: *theory of computability*) is a branch of mathematical logic studying the notion of computability from a rather theoretical point of view. This includes giving a lot of attention to what is *not* computable, or what is computable relative to any given, not necessarily computable, function. The subject is interesting on philosophical-scientific grounds because of the Church-Turing Thesis and its role in computer science, and because of the intriguing concept of Kolmogorov complexity. This course tries to keep touch with how recursion theory actually impacts mathematics and computer science. This impact is small, but it does exist.

Accordingly, the first chapter of the course covers the basics: primitive recursion, partial recursive functions and the Church-Turing Thesis, arithmetization and the theorems of Kleene, the halting problem and Rice's theorem, recursively enumerable sets, selections and reductions, recursive inseparability, and index systems. (Turing machines are briefly discussed, but the arithmetization is based on a numerical coding of combinators.)

The second chapter is devoted to the remarkable negative solution (but with positive aspects) of Hilbert's 10th Problem. This uses only the most basic notions of recursion theory, plus some elementary number theory that is worth knowing in any case; the coding aspects of natural numbers are combined here ingeniously with the arithmetic of the semiring of natural numbers.

The last chapter is on Kolmogorov complexity, where concepts of recursion theory are used to define and study notions of randomness and information content. This also involves a bit of measure theory.

The first and last chapter are largely based on a set of notes written by Christian Rosendal, and in the second chapter I follow mostly the treatment in Smorynski's excellent book *Logical Number Theory I.*

An obvious omission in the material above is that of *relative* computability. This topic merges with effective descriptive set theory and deserves to be included, but there is only so much one can do in one semester. I have opted for a rather careful treatment of fewer topics.

Notation: $\mathbb{N} = \{0, 1, 2, \ldots\}$ is the set of natural numbers, including 0, and a, b, c, d, e and k, l, m, n (sometimes with subscripts or accents) range over $\mathbb{N}$.

Chapter 1

Basic Recursion Theory

In this chapter, x, y, z, sometimes with subscripts, range over $\mathbb{N}$ as well. (In later chapters, x, y, z can be something else.)

1.1 Primitive Recursion

We define $\mathcal{F}_d$ to be the set of all functions $f : \mathbb{N}^d \to \mathbb{N}$, and put $\mathcal{F} := \bigcup_d \mathcal{F}_d$. We identify $\mathcal{F}_0$ with $\mathbb{N}$ in the obvious way. For $i = 1, \dots, d$, the ith projection function $P^i_d : \mathbb{N}^d \to \mathbb{N}$ is given by $P^i_d(x_1, \dots, x_d) = x_i$. The *successor function* $S : \mathbb{N} \to \mathbb{N}$ is given by $S(x) = x + 1$. To denote functions, we sometimes borrow a notation from λ-calculus: if $t(x_1, \dots, x_d)$ is an expression such that $t(a_1, \dots, a_d) \in \mathbb{N}$ for all $a_1, \dots, a_d$, then $\lambda x_1 \dots x_d.t(x_1, \dots, x_d)$ denotes the corresponding function

$$(a_1, \dots, a_d) \mapsto t(a_1, \dots, a_d) : \mathbb{N}^d \to \mathbb{N}.$$

For example, $S = \lambda x.x + 1$ and $P^i_d = \lambda x_1 \dots x_d.x_i$.

Now we define **substitution**. For $g \in \mathcal{F}_n$ and $f_1, \dots, f_n \in \mathcal{F}_d$, $g(f_1, \dots, f_n)$ is the function $\lambda x_1 \dots x_d.g(f_1(x_1, \dots, x_d), \dots, f_n(x_1, \dots, x_d))$ in $\mathcal{F}_d$.

For any $g \in \mathcal{F}_d$ and $h \in \mathcal{F}_{d+2}$, there is a unique $f \in \mathcal{F}_{d+1}$ such that for all $x_1, \dots, x_d, y$:

$$\begin{aligned} f(x_1, \dots, x_d, 0) &= g(x_1, \dots, x_d), \\ f(x_1, \dots, x_d, y + 1) &= h(x_1, \dots, x_d, y, f(x_1, \dots, x_d, y)). \end{aligned}$$

This function f is said to be obtained by **primitive recursion** from g and h.

1.1.1 Primitive Recursive Functions

Loosely speaking, the functions obtainable by the above procedures are the primitive recursive functions. More precisely,

Definition 1. *The set* PR *of* ***primitive recursive functions*** *is the smallest subset of $\mathcal{F}$ such that, with* $\mathrm{PR}_d := \mathrm{PR} \cap \mathcal{F}_d$*,*

(a) all constant functions in $\mathcal{F}$ belong to PR*;*

(b) $S \in \mathrm{PR}_1$*;*

(c) $P^i_d \in \mathrm{PR}_d$ *for* $i = 1, \ldots, d$*;*

(d) if $g \in \mathrm{PR}_n$*,* $f_1, \ldots, f_n \in \mathrm{PR}_d$*, then* $g(f_1, \ldots, f_n) \in \mathrm{PR}_d$ *(closure under substitution);*

(e) if $g \in \mathrm{PR}_d$ *and* $h \in \mathrm{PR}_{d+2}$ *and* $f \in \mathcal{F}_{d+1}$ *is obtained by primitive recursion from* g, h *then* $f \in \mathrm{PR}_{d+1}$ *(closure under primitive recursion).*

Essentially all functions in $\mathcal{F}$ that arise in ordinary mathematical practice are primitive recursive. For example, the greatest common divisor function gcd is primitive recursive, but proving such facts will be much easier once we have some lemmas available.

In addition to functions, some *sets* will be called primitive recursive. Recall that the *characteristic function* of a set $A \subseteq \mathbb{N}^d$ is the function $\chi_A : \mathbb{N}^d \to \mathbb{N}$ defined by $\chi_A(\vec{x}) = 1$ if $\vec{x} \in A$ and $\chi_A(\vec{x}) = 0$ if $\vec{x} \in \mathbb{N}^d \setminus A$.

Definition 2. *A set $A \subseteq \mathbb{N}^d$ is said to be* ***primitive recursive*** *if its characteristic function χ_A is primitive recursive.*

A set $A \subseteq \mathbb{N}^d$ is often construed as a d-ary relation on $\mathbb{N}$, and so, when $\vec{x} \in A$, we often write $A(\vec{x})$ instead of $\vec{x} \in A$.

1.1.2 Basic Examples

- The addition operation $+ : \mathbb{N}^2 \to \mathbb{N}$ is primitive recursive: let $g = P^1_1 \in \mathcal{F}_1 \cap E$, $h = S(P^3_3) \in \mathcal{F}_3 \cap E$, i.e. $g(x) = x$ and $h(x, y, z) = z + 1$. Then $+$ is obtained by primitive recursion from g, h:

  $$x + 0 = x = g(x), \qquad x + (y + 1) = (x + y) + 1 = h(x, y, x + y).$$

- The multiplication operation $\cdot : \mathbb{N}^2 \to \mathbb{N}$ is primitive recursive. For let $g \in \mathcal{F}_1 \cap E$, $h \in \mathcal{F}_3 \cap E$ be given by

  $$g(x) = 0, \qquad h(x, y, z) = z + x \ (= +(P^3_3, P^1_3)(x, y, z)).$$

  Then $\cdot$ is obtained by primitive recursion from g, h, since $x \cdot 0 = 0 = g(x)$ and $x \cdot (y + 1) = xy + x = h(x, y, xy)$.

- $\lambda xy.x^y$ is primitive recursive: $x^0 = 1$, $x^{y+1} = x^y \cdot x$.

- The monus function $\dot{-} : \mathbb{N}^2 \to \mathbb{N}$, where

  $$x \dot{-} y = \begin{cases} x - y & \text{if } x \geq y, \\ 0 & \text{otherwise.} \end{cases}$$

First, we observe that $\lambda x.x \dot{-} 1$ is primitive recursive, as $0 \dot{-} 1 = 0$ and $(x+1) \dot{-} 1 = x$. Then $\dot{-}$ itself is primitive recursive: $x \dot{-} 0 = x$, $x \dot{-} (y+1) = (x \dot{-} y) \dot{-} 1$.

- $\mathsf{sign} : \mathbb{N} \to \mathbb{N}$, given by

$$\mathsf{sign}(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x = 0, \end{cases}$$

  is primitive recursive, for $\mathsf{sign}(0) = 0, \mathsf{sign}(x+1) = 1$.

- The set $A = \{(x,y) \in \mathbb{N}^2 : x < y\} \subseteq \mathbb{N}^2$ is primitive recursive since $\chi_A(x,y) = \mathsf{sign}(y \dot{-} x)$.

- The diagonal $\Delta = \{(x,y) \in \mathbb{N}^2 : x = y\} \subseteq \mathbb{N}^2$ is primitive recursive:

$$\chi_\Delta(x,y) = 1 \dot{-} ((x \dot{-} y) + (y \dot{-} x)).$$

- For $A, B \subseteq \mathbb{N}^d$ we have $\chi_\emptyset = 0$, $\chi_{A \cap B} = \chi_A \cdot \chi_B$, and $\chi_{-A} = 1 \dot{-} \chi_A$. Therefore, the primitive recursive subsets of $\mathbb{N}^d$ are the elements of a boolean algebra of subsets of $\mathbb{N}^d$.

- **Definition by cases:** Suppose $A \subseteq \mathbb{N}^d$ and $f, g \in \mathcal{F}_d$ are primitive recursive. Then the function $h : \mathbb{N}^d \to \mathbb{N}$ given by

$$h(\vec{x}) = \begin{cases} f(\vec{x}) & \text{for } \vec{x} \in A, \\ g(\vec{x}) & \text{for } \vec{x} \notin A, \end{cases}$$

  is primitive recursive because $h = f \cdot \chi_A + g \cdot \chi_{-A}$.

- **Iterated sums/products:** Let $f \in \mathcal{F}_{d+1}$ and let

$$g = \lambda x_1 \ldots x_d y. \sum_{i=0}^{y} f(x_1, \ldots, x_d, i).$$

  Note that $g(\vec{x}, 0) = f(\vec{x}, 0)$ and $g(\vec{x}, y+1) = g(\vec{x}, y) + f(\vec{x}, y+1)$. Thus if f is primitive recursive, so is g, and vice-versa. Similarly, if f is primitive recursive, so is

$$\lambda x_1 \ldots x_d y. \prod_{i=0}^{y} f(x_1, \ldots, x_d, i).$$

Notation. Let $p(i)$ be a condition on $i \in \mathbb{N}$: for each $i \in \mathbb{N}$ either $p(i)$ holds, or $p(i)$ does not hold. Then $\mu i_{\leq y}\ p(i)$ denotes the least $i \leq y$ such that $p(i)$ holds if there is such an i, and if no such i exists, then $\mu i_{\leq y}\ p(i) := 0$. For example, if $A \subseteq \mathbb{N}$, then $\mu i_{\leq 3}\ i \in A$ is one of the numbers $0, 1, 2, 3$.

- **Bounded search:** Let $A \subseteq \mathbb{N}^{d+1}$, and define $f \in \mathcal{F}_{d+1}$ by

$$f(\vec{x}, y) = \mu i_{\leq y}\ A(\vec{x}, i).$$

  If A is primitive recursive, then so is f. This is because $f(\vec{x}, 0) = 0$ and

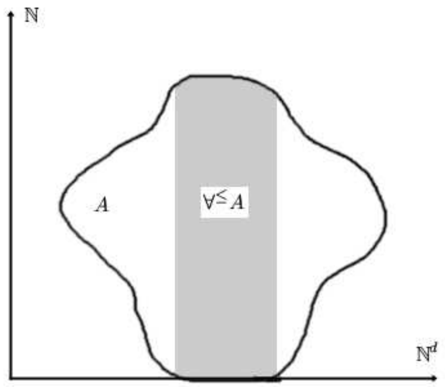
$$f(\vec{x}, y+1) = \begin{cases} f(\vec{x}, y) & \text{if } \sum_{i=0}^{y} \chi_A(\vec{x}, i) \geq 1, \\ y+1 & \text{if } \sum_{i=0}^{y} \chi_A(\vec{x}, i) = 0 \text{ and } A(\vec{x}, y+1), \\ 0 & \text{if } \sum_{i=0}^{y} \chi_A(\vec{x}, i) = 0 \text{ and } A(\vec{x}, y+1). \end{cases}$$

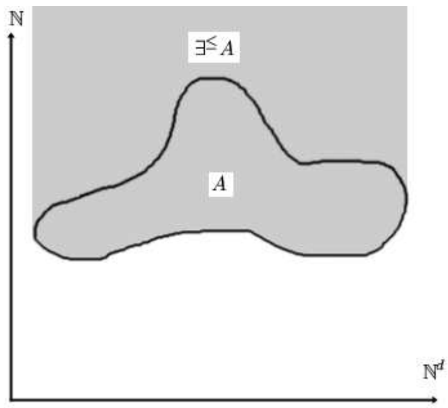
  More generally, if $A \subseteq \mathbb{N}^{d+1}$ and $g \in \mathcal{F}_{d+1}$ are primitive recursive, so is the function $\lambda \vec{x} y . \mu i_{\leq g(\vec{x}, y)}\ A(\vec{x}, i)$ (this is a function in $\mathcal{F}_{d+1}$).

- **Bounded quantification:** Let $A \subseteq \mathbb{N}^{d+1}$ and define:

$$\begin{aligned} \exists^{\leq} A &= \{(\vec{x}, y) \in \mathbb{N}^{d+1} : \exists i \leq y\ A(\vec{x}, i)\} \subseteq \mathbb{N}^{d+1}, \\ \forall^{\leq} A &= \{(\vec{x}, y) \in \mathbb{N}^{d+1} : \forall i \leq y\ A(\vec{x}, i)\} \subseteq \mathbb{N}^{d+1}. \end{aligned}$$

  Since $\chi_{\exists^{\leq} A}(\vec{x}, y) = \mathsf{sign}\big(\sum_{i=0}^{y} \chi_A(\vec{x}, i)\big)$ and $\chi_{\forall^{\leq} A}(\vec{x}, y) = \mathsf{sign}\big(\prod_{i=0}^{y} \chi_A(\vec{x}, i)\big)$, it follows that if A is primitive recursive, then so are $\exists^{\leq} A$ and $\forall^{\leq} A$.





Exercise. Prove that gcd : $\mathbb{N}^2 \to \mathbb{N}$ is primitive recursive. Here, $\gcd(0, 0) = 0$, and $\gcd(x, y)$ is the greatest d such that $d|x$ and $d|y$, for x, y not both zero.

1.1.3 Coding

Enumerate $\mathbb{N}^2$ diagonally as follows:

$$\begin{array}{cccccccc} (0,0) & (1,0) & (0,1) & (2,0) & (1,1) & (0,2) & (3,0) & \cdots \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \\ 0 & 1 & 2 & 3 & 4 & 5 & 6 & \cdots \end{array}$$

This bijection $(x,y) \mapsto |x,y| \ : \ \mathbb{N}^2 \to \mathbb{N}$ is given by:

$$|x,y| = \frac{(x+y)(x+y+1)}{2} + y.$$

This bijection is primitive recursive: use that

$$|x,y| \ = \ \mu i_{\leq f(x,y)} \ 2i = f(x,y)$$

with $f : \mathbb{N}^2 \to \mathbb{N}$ given by $f(x,y) = (x+y)(x+y+1) + 2y$, so f is primitive recursive. Note also that $x \leq |x,y|$ and $y \leq |x,y|$ for all x, y.

Proposition 3. *For $d \geq 1$, we have primitive recursive functions*

$$\langle \ \rangle^d : \mathbb{N}^d \to \mathbb{N} \text{ and } (\ )^d_1, \dots, (\ )^d_d : \mathbb{N} \to \mathbb{N},$$

such that $\langle \ \rangle^d$ is a bijection and $\langle (x)^d_1, \dots, (x)^d_d \rangle = x$ for all x.

Proof. For $d = 1$, take both $\langle \ \rangle^1$ and $(\ )^1_1$ as the identity function on $\mathbb{N}$. For $d = 2$, take $\langle x, y \rangle^2 = |x,y|$ and define $(\ )^2_1$ and $(\ )^2_2$ by bounded search:

$$(x)^2_1 := \mu i_{\leq x} \ (\exists y \leq x \ |i,y| = x), \qquad (x)^2_2 = \mu i_{\leq x} \ (\exists y \leq x \ |y,i| = x).$$

We now proceed by induction. Given $\langle \ \rangle^d$ and $(\ )^d_1, \dots, (\ )^d_d$ with the desired properties, and $d \geq 2$, put

$$\begin{aligned} &\langle x_1, \dots, x_d, y \rangle^{d+1} = |\langle x_1, \dots, x_d \rangle^d, y| \ \text{ and} \\ &(x)^{d+1}_i = ((x)^2_1)^d_i \ \text{ for } i = 1, \dots, d, \quad (x)^{d+1}_{d+1} = (x)^2_2. \end{aligned}$$

Note that the identities just displayed also hold for $d = 1$. □

1.1.4 More General Recursions

Let us consider first *double recursions*: Suppose $g, g' \in \mathcal{F}_d$ and $h, h' \in \mathcal{F}_{d+3}$ are primitive recursive. Let $f, f' \in \mathcal{F}_{d+1}$ be given by:

$$\begin{aligned} f(\vec{x}, 0) &= g(\vec{x}), & f(\vec{x}, y+1) &= h\big(\vec{x}, y, f(\vec{x}, y), f'(\vec{x}, y)\big), \\ f'(\vec{x}, 0) &= g'(\vec{x}), & f'(\vec{x}, y+1) &= h'\big(\vec{x}, y, f(\vec{x}, y), f'(\vec{x}, y)\big). \end{aligned}$$

Then f is primitive recursive. To see this, we use the above coding method and define $t \in \mathcal{F}_{d+1}$ by $t(\vec{x}, 0) = |g(\vec{x}), g'(\vec{x})|$, and

$$t(\vec{x}, y+1) = |h\big(\vec{x}, y, (t(\vec{x}, y))^2_1, (t(\vec{x}, y))^2_2\big), \ h'\big(\vec{x}, y, (t(\vec{x}, y))^2_1, (t(\vec{x}, y))^2_2\big)|.$$

Thus t is primitive recursive. We have $f(\vec{x}, y) = (t(\vec{x}, y))^2_1$ and $f'(\vec{x}, y) = (t(\vec{x}, y))^2_2$, so f and f' are primitive recursive.

It turns out that if a function is defined recursively in terms of several of its previous values (the most general case being that $f(\vec{x}, y+1)$ is computed in terms of $f(\vec{x}, 0), \dots, f(\vec{x}, y)$), then it is still primitive recursive. To deal with this situation, we use a Skolem trick. Given $f \in \mathcal{F}_{d+1}$, define $\widehat{f} \in \mathcal{F}_{d+1}$ by $\widehat{f}(\vec{x}, 0) = f(\vec{x}, 0)$, and $\widehat{f}(\vec{x}, y) = |f(\vec{x}, y), \widehat{f}(\vec{x}, y-1)|$ for $y > 0$. So $\widehat{f}(\vec{x}, y)$ encodes the values of f at $(\vec{x}, 0), (\vec{x}, 1), \ldots, (\vec{x}, y)$.

Let $g \in \mathcal{F}_d$ and $h \in \mathcal{F}_{d+2}$, and define $f \in \mathcal{F}_{d+1}$ by

$$f(\vec{x}, 0) = g(\vec{x}) \text{ and } f(\vec{x}, y+1) = h(\vec{x}, y, \widehat{f}(x, y)).$$

Claim. If g and h are primitive recursive, then so is f.

Proof. Assume g and h are primitive recursive. To obtain that f is primitive recursive, it suffices to show that $\widehat{f}$ is primitive recursive, because

$$f(\vec{x}, y) = \begin{cases} \widehat{f}(\vec{x}, y) & \text{if } y = 0, \\ (\widehat{f}(\vec{x}, y))_1^2 & \text{if } y > 0. \end{cases}$$

But $\widehat{f}(\vec{x}, 0) = g(\vec{x})$, and

$$\begin{aligned} \widehat{f}(\vec{x}, y+1) &= \langle f(\vec{x}, y+1), \widehat{f}(\vec{x}, y) \rangle \\ &= \langle h(\vec{x}, y, \widehat{f}(\vec{x}, y)), \widehat{f}(\vec{x}, y) \rangle, \end{aligned}$$

so $\widehat{f}$ is primitive recursive, as desired. □

Exercise. Show that the Fibonacci sequence (F_n) defined by $F_0 = 0, F_1 = 1$ and $F_n = F_{n-1} + F_{n-2}$ for $n \geq 2$, is primitive recursive.

For later use we also need an encoding of finite sequences of variable length such that the length of the sequence can be decoded from its code. We define

$$\langle m_1, \ldots, m_d \rangle := \langle m_1, \ldots, m_d, d \rangle^{d+1}.$$

Note that then

$$\langle m_1, \ldots, m_d \rangle = \langle n_1, \ldots, n_e \rangle \iff d = e \text{ and } m_1 = n_1, \ldots, m_d = n_d.$$

Define $\mathrm{lh} : \mathbb{N} \to \mathbb{N}$ by $\mathrm{lh}(n) = (n)_2^2$, so that $\mathrm{lh}(\langle n_1 \ldots, n_d \rangle) = d$. It is clear that lh is primitive recursive. Let $B \subseteq \mathbb{N}$ be the set of all $\langle n_1, \ldots, n_d \rangle$ for all d.

Claim. $B = \{b \in \mathbb{N} : (b)_2^2 > 0\} \cup \{0\}$.

To see this, note that if $(b)_2^2 = d > 0$, then $b = |a, d|$ with $a \in \mathbb{N}$, so $a = \langle n_1, \ldots, n_d \rangle^d$ for suitable $n_1, \ldots, n_d$, hence $b = \langle n_1, \ldots, n_d \rangle \in B$. It remains to note that the only element $\langle n_1, \ldots, n_d \rangle$ of B with $d = 0$ is $\langle 0 \rangle^1 = 0$.

It follows from the claim that B is primitive recursive. Next, we introduce a primitive recursive function $(n, j) \mapsto (n)_j : \mathbb{N}^2 \to \mathbb{N}$ such that for $n = \langle n_1, \ldots, n_d \rangle$ and $1 \leq j \leq d$ we have $(n)_j = n_j$, and thus $(n)_j < n$. First we define a primitive recursive $f : \mathbb{N}^2 \to \mathbb{N}$ by

$$f(n, 0) = n, \qquad f(n, i+1) = \big(f(n, i)\big)_1^2.$$

It is easy to check that for $n = \langle n_d, n_{d-1}, \dots, n_1 \rangle$ and $1 \le i \le d$ we have

$$f(n,i) = \langle n_d, \dots, n_i \rangle^{d+1-i}.$$

Now define the primitive recursive function $(n,j) \mapsto (n)_j \ : \mathbb{N}^2 \to \mathbb{N}$ by

$$(n)_j := f(n, \mathrm{lh}(n) + 1 \dot{-} j)_2^2 \text{ for } j \neq 1, \quad (n)_1 := f(n, \mathrm{lh}(n)).$$

This function has the desired properties as is easily verified.

1.2 Partial Recursive Functions

From the way we defined "primitive recursive", it is clear that each primitive recursive function is *computable* in an intuitive sense. A standard diagonalization argument, however, yields a computable function $f : \mathbb{N} \to \mathbb{N}$ that is not primitive recursive. This argument goes as follows: one can effectively produce a list $f_0, f_1, f_2, \dots$ of all functions in PR_1. (Any primitive recursive function $\mathbb{N} \to \mathbb{N}$ may appear infinitely often in this list.) Here, *effective* means that we have an algorithm/program that on any input (m,n) computes $f_m(n)$. Now define

$$f : \mathbb{N} \to \mathbb{N}, \qquad f(n) := f_n(n) + 1.$$

Then f is computable in the intuitive sense, but $f \neq f_n$ for every n, so f cannot be primitive recursive. More generally, any effective method of characterizing the intuitive notion of (total) computable function $\mathbb{N} \to \mathbb{N}$ must fail by a similar diagonalization.

It turns out that we *can* effectively characterize a more general notion of *partial* computable function $\mathbb{N} \rightharpoonup \mathbb{N}$ that *does* correspond to the intuitive notion of what is computable by an algorithm; the key point is that algorithms may fail to terminate on some inputs. Accordingly, we extend our notion of primitive recursive function to that of partial recursive function, essentially by allowing *unbounded search.*

Notation

For sets P, Q, we denote by $f : P \rightharpoonup Q$ a partial function f from P into Q, that is, a function from a set $D(f) \subseteq P$ into Q; then

$$f(p)\downarrow \qquad \text{(in words: } f \text{ converges at } p\text{)}$$

means that $p \in D(f)$, and

$$f(p)\uparrow \qquad \text{(in words: } f \text{ diverges at } p\text{)}$$

means that $p \in P$ but $p \notin D(f)$.

Let $f : \mathbb{N}^{d+1} \rightharpoonup \mathbb{N}$. Then $\lambda\vec{x}.\ (\mu y f(\vec{x}, y) = 0)$ denotes the partial function $g : \mathbb{N}^d \rightharpoonup \mathbb{N}$ such that for $\vec{x} \in \mathbb{N}^d$:

$$g(\vec{x})\downarrow \iff \exists y \big[\forall i \le y\ f(\vec{x}, i)\downarrow \text{ and } \forall i < y\ f(\vec{x}, i) \neq 0 \text{ and } f(\vec{x}, y) = 0\big],$$

and if $g(\vec{x})\downarrow$, then $g(\vec{x})$ is the unique y witnessing the righthandside in the above equivalence. Let $g : \mathbb{N}^n \rightharpoonup \mathbb{N}$ and $f_1, \dots, f_n : \mathbb{N}^d \rightharpoonup \mathbb{N}$. Then

$$g(f_1, \dots, f_n) \ : \ \mathbb{N}^d \rightharpoonup \mathbb{N}$$

is given by:

$$g(f_1, \dots, f_n)(\vec{x})\downarrow \iff f_1(\vec{x})\downarrow, \dots, f_n(\vec{x})\downarrow \text{ and } g(f_1(\vec{x}), \dots, f_n(\vec{x}))\downarrow$$

and if $g(f_1, \dots, f_n)(\vec{x})\downarrow$, then $g(f_1, \dots, f_n)(\vec{x}) = g(f_1(\vec{x}), \dots, f_n(\vec{x}))$. Given $g : \mathbb{N}^d \rightharpoonup \mathbb{N}$, $h : \mathbb{N}^{d+2} \rightharpoonup \mathbb{N}$ there is clearly a unique $f : \mathbb{N}^{d+1} \rightharpoonup \mathbb{N}$ such that for all $\vec{x} \in \mathbb{N}^d$ and $y \in \mathbb{N}$:

- $f(\vec{x}, 0)\downarrow \iff g(\vec{x})\downarrow$; if $f(\vec{x}, 0)\downarrow$, then $f(\vec{x}, 0) = g(\vec{x})$.
- $f(\vec{x}, y+1)\downarrow \iff f(\vec{x}, y)\downarrow, \quad h(\vec{x}, y, f(\vec{x}, y))\downarrow$; if $f(\vec{x}, y+1)\downarrow$, then $f(\vec{x}, y+1) = h(\vec{x}, y, f(\vec{x}, y))$.

This f is said to be obtained by primitive recursion from g, h.

Definition 4. *The set of partial recursive functions is the smallest set of partial functions $\mathbb{N}^d \rightharpoonup \mathbb{N}$ for $d = 0, 1, 2, \dots$ such that:*

(a) the function $O : \mathbb{N}^0 \to \mathbb{N}$ with value 0 is partial recursive, and the successor function $S : \mathbb{N} \to \mathbb{N}$ is partial recursive;

(b) the functions P^i_d, for $1 \leq i \leq d$, are partial recursive;

(c) whenever $g : \mathbb{N}^m \rightharpoonup \mathbb{N}, h_1, \dots, h_m : \mathbb{N}^d \rightharpoonup \mathbb{N}$ are partial recursive, so is $g(h_1, \dots, h_m)$;

(d) whenever $g : \mathbb{N}^d \rightharpoonup \mathbb{N}, h : \mathbb{N}^{d+2} \rightharpoonup \mathbb{N}$ are partial recursive, then so is the function obtained from g, h by primitive recursion;

(e) whenever $f : \mathbb{N}^{d+1} \rightharpoonup \mathbb{N}$ is partial recursive, then so is

$$\lambda\vec{x}.\ (\mu y f(\vec{x}, y) = 0) : \mathbb{N}^d \rightharpoonup \mathbb{N}.$$

A *recursive* function is a partial recursive function $f : \mathbb{N}^d \to \mathbb{N}$, the notation here indicating that $D(f) = \mathbb{N}^d$. A set $A \subseteq \mathbb{N}^d$ is *recursive* if $\chi_A : \mathbb{N}^d \to \mathbb{N}$ is recursive. It is easy to check that primitive recursive functions are recursive.

Lemma 5. *Given $f : \mathbb{N}^d \to \mathbb{N}$, the following are equivalent:*

(a) f is recursive;

(b) $\mathrm{graph}(f) \subseteq \mathbb{N}^{d+1}$ *is recursive.*

Proof. Use that $\chi_{\mathrm{graph}(f)}(\vec{x}, y) = \chi_\Delta(f(\vec{x}), y)$. In the other direction, use

$$f(\vec{x}) = \mu y(\chi_{\mathrm{graph}(f)}(\vec{x}, y) = 1).$$

□

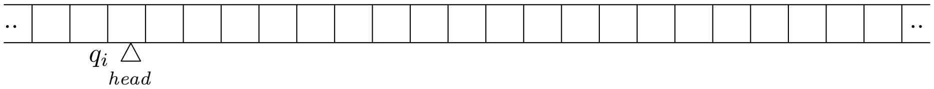
By similar arguments as for primitive recursive sets one obtains:

Theorem 6. *The class of recursive subsets of $\mathbb{N}^d$ is closed under finite unions, intersections and complements. If $A \subseteq \mathbb{N}^{d+1}$ is recursive, so are $\exists^{\leq} A$ and $\forall^{\leq} A$ as subsets of $\mathbb{N}^d$.*

Exercise. Show that if $f : \mathbb{N} \to \mathbb{N}$ is a recursive bijection, then so is its inverse. Show that there is a primitive recursive bijection $\mathbb{N} \to \mathbb{N}$ whose inverse is not primitive recursive. (For the second problem, use that there is a recursive function $\mathbb{N} \to \mathbb{N}$ that is not primitive recursive.)

1.2.1 Turing Machines

Definition 7. *A Turing machine consists of a biinfinite tape of successive boxes,*



and a head that can read, erase, write in the box in case it is blank, and move to the box on the left or the right. Moreover, there are given

- *A finite alphabet $\Sigma = \{s_0, s_1, \ldots, s_n\}$ with $n \geq 1, s_i \neq s_j$ when $i \neq j$, with distinguished symbols $s_0 := \#$ for 'blank', and $s_1 := 0$.*
- *A finite set of internal states $Q = \{q_0, q_1, \ldots, q_m\}$ with $m \geq 1$, $q_i \neq q_j$ for $i \neq j$, with distinguished states q_0 (the initial state), and $q_m := q_f$ (the final state).*
- *A finite set $\{I_1, \ldots, I_p\}$ of instructions, each of one of the following three types:*

  (a) $q_a s_b s_c q_d$: if in state q_a reading s_b, erase s_b, write s_c in its place, and enter state q_d;

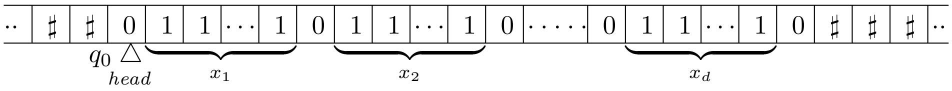
  (b) $q_a s_b R q_d$: if in state q_a reading s_b, move to the box on the right, and enter state q_d;

  (c) $q_a s_b L q_d$: if in state q_a reading s_b, move to the box on the left, and enter state q_d.

This set of instructions should be complete: for each $q_a \neq q_f$, and each symbol s_b there is exactly one instruction of the form $q_a s_b \ldots$, and there is no instruction of the form $q_f \ldots$.

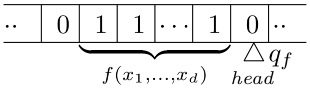
Formally, a Turing machine is just a triple $(\Sigma, Q, \{I_1, \ldots, I_p\})$ as above.

Definition 8. *Let M be a Turing machine with alphabet $\{\#, 0, 1\}$. We say that M computes the partial function $f : \mathbb{N}^d \rightharpoonup \mathbb{N}$ if for any input $(x_1, \ldots, x_d)$ as follows:*



the machine, applying the instructions,

- *either never enters q_f, and then $f(\vec{x})\uparrow$.*
- *or eventually enters state q_f, and then $f(\vec{x})\downarrow$, with its head and tape as follows:*



Fact. *Turing Computable = Partial Recursive.*

Turing gave a compelling analysis of the intuitive concept of computability, in 1936, and this led him to identify it with the precise notion of Turing computability. Turing machines remain important as a model of computation in connection with complexity theory. But in the rest of this course we deal directly with partial recursive functions without using Turing machines.

1.3 Arithmetization and Kleene's Theorems

We shall give numerical codes to the programs that compute partial recursive functions. First, we introduce formal expressions that specify these programs. These expressions will be called combinators and they are words on the alphabet with the following distinct symbols:

(a) the two symbols O and S,

(b) for each pair i, d such that $1 \leq i \leq d$ a symbol P^i_d,

(c) for each pair m, d, a symbol S^m_d (a substitution symbol),

(d) for each d a symbol R_d (a primitive recursion symbol),

(e) for each d a symbol S_d (an unbounded search symbol).

Each combinator has a specific arity (a natural number) and each d-ary combinator f is a word on this alphabet, and has associated to it a partial recursive function $\hat{f} : \mathbb{N}^d \rightharpoonup \mathbb{N}$. The definition is inductive:

(a) The word O of length 1 is a nullary combinator with associated function $\mathbb{N}^0 \to \mathbb{N}$ taking the value 0. The word S of length 1 is a unary combinator with associated function $x \mapsto x + 1 : \mathbb{N} \to \mathbb{N}$.

(b) For $1 \leq i \leq d$, the word P^i_d is a d-ary combinator of length 1 with associated function $(x_1, \ldots, x_d) \mapsto x_i : \mathbb{N}^d \to \mathbb{N}$.

(c) If g is an m-ary combinator and $h_1, \ldots, h_m$ are d-ary combinators, then $\mathrm{S}^m_d\, gh_1 \ldots h_m$ is a d-ary combinator with associated function $\hat{g}(\hat{h}_1, \ldots, \hat{h}_m)$.

(d) If g is a d-ary combinator and h is a $(d+2)$-ary combinator, then $\mathrm{R}_d\, gh$ is a $(d+1)$-ary combinator whose associated function $\mathbb{N}^{d+1} \rightharpoonup \mathbb{N}$ is obtained by primitive recursion from $\hat{g}$ and $\hat{h}$.

(e) If g is a $(d+1)$-ary combinator, then $\mathrm{S}_d\, g$ is a d-ary combinator whose associated function is

$$\lambda x_1 \ldots x_d.\ (\mu y\, (\hat{g}(x_1, \ldots, x_d, y) = 0)\,.$$

Next we associate inductively to each combinator g a number $\#g \in \mathbb{N}$:

(a) $\#\,\mathrm{O} = \langle 1, 0\rangle$, $\#\,\mathrm{S} = \langle 1, 1\rangle$,

(b) $\#\,\mathrm{P}^i_d = \langle 2, i, d\rangle$,

(c) $\#(\mathrm{S}^m_d\, fg_1 \ldots g_m) = \langle 3, \#f, \#g_1, \ldots, \#g_m, d\rangle$,

(d) $\#(\mathrm{R}_d\, gh) = \langle 4, \#g, \#h, d+1\rangle$,

(e) $\#\,\mathrm{S}_d\, g = \langle 5, \#g, d\rangle$.

Let Co be the set of all combinators. Then $\# : \mathrm{Co} \to \mathbb{N}$ is clearly injective. It is easy to check that the primitive recursive function $\alpha : \mathbb{N} \to \mathbb{N}$ given by $\alpha(n) := (n)_{\mathrm{lh}(n)}$ has the property that if $n = \#g$ with $g \in \mathrm{Co}$, then $\alpha(n)$ is the arity of g.

Lemma 9. *The set* $\#\,\mathrm{Co}$ *is primitive recursive.*

Proof. With B as in subsection 1.1.4, consider the following subsets of B:

(a) $B_1 := \{\langle 1, 0\rangle, \langle 1, 1\rangle\}$;

(b) $B_2 := \{n \in B : (n)_1 = 2, \mathrm{lh}(n) = 3, 1 \leq (n)_2 \leq (n)_3\}$;

(c) B_3 is the set of all $n \in B$ such that $(n)_1 = 3$ and such that $m := \alpha((n)_2)$ satisfies $\mathrm{lh}(n) = m + 3, \alpha((n)_3) = \alpha((n)_4) = \cdots = \alpha((n)_{m+2}) = (n)_{\mathrm{lh}(n)}$;

(d) $B_4 := \{n \in B : (n)_1 = 4, \mathrm{lh}(n) = 4, \alpha((n)_3) = \alpha((n)_2) + 2 = (n)_4 + 1\}$;

(e) $B_5 := \{n \in B : (n)_1 = 5, \mathrm{lh}(n) = 3, \alpha((n)_2) = (n)_3 + 1\}$.

Note that these sets B_i are primitive recursive. Let $A := \#\mathrm{Co}$, so $\chi_A(n) = 0$ if $n \notin B$, and also if $n \in B \setminus \bigcup_{i=1}^5 B_i$. It remains to describe χ_A on the B_i. Clearly, $\chi_A(n) = 1$ for $n \in B_1 \cup B_2$. For $n \in B_3$ and $m := \alpha((n)_2)$ we have

$$\chi_A(n) = \prod_{i=2}^{m+2} \chi_A((n)_i),$$

For $n \in B_4$ we have $\chi_A(n) = \chi_A((n)_2)\chi_A((n)_3)$, and for $n \in B_5$ we have $\chi_A(n) = \chi_A((n)_2)$. □

Note that for any partial recursive function $\phi : \mathbb{N}^d \rightharpoonup \mathbb{N}$, there are infinitely many d-ary combinators f such that $\hat{f} = \phi$. This is because for any d-ary combinator f we have $\hat{f} = \hat{g}$ where $g := \mathrm{S}^1_d\,\mathrm{P}^1_1\,f$.

We have encoded programs—more precisely, combinators—by numbers, and next we shall encode *terminating computations* using these programs.

Given a d-ary combinator f and $\vec{x} = (x_1, \dots, x_d)$, we also write $f(\vec{x})$ instead of $\hat{f}(\vec{x})$.

Let f be a combinator and $f(\vec{x}) = y$. The latter means that $\vec{x}$ is in the domain of $\hat{f}$ and $\hat{f}(\vec{x}) = y$. Then we assign to the triple $(f, \vec{x}, y)$ a number $\mathrm{ct}(f, \vec{x}, y) \in \mathbb{N}$, called the *computation tree* of f at input $\vec{x}$. This assignment is defined inductively:

(a) if f is O, S, or P^i_d, then, with $\vec{x} = (x_1, \dots, x_d)$,

$$\mathrm{ct}(f, \vec{x}, y) := \langle \#f, \vec{x}, y \rangle := \langle \#f, x_1, \dots, x_d, y \rangle;$$

(b) if f is $\mathrm{S}^m_d g h_1 \dots h_m$, then, with $\vec{x} = (x_1, \dots, x_d)$, $h_1(\vec{x}) = u_1, \dots, h_m(\vec{x}) = u_m$, $\vec{u} = (u_1, \dots, u_m)$, and $g(\vec{u}) = y = f(\vec{x})$,

$$\mathrm{ct}(f, \vec{x}, y) := \langle \#f, \vec{x}, \mathrm{ct}(h_1, \vec{x}, u_1), \dots, \mathrm{ct}(h_m, \vec{x}, u_m), \mathrm{ct}(g, \vec{u}, y), y \rangle;$$

(c) if f is $\mathrm{R}_d g h$, then, with $\vec{x} = (x_1, \dots, x_d, x_{d+1})$ and $\vec{x}_d := (x_1, \dots, x_d)$,

$$\mathrm{ct}(f, \vec{x}, y) := \begin{cases} \langle \#f, \vec{x}_d, 0, \mathrm{ct}(g, \vec{x}_d, y), y \rangle & \text{if } x_{d+1} = 0 \\ \langle \#f, \vec{x}_d, n+1, \mathrm{ct}(h, \vec{x}_d, n, f(\vec{x}_d, n), y), y \rangle & \text{if } x_{d+1} = n+1; \end{cases}$$

(d) if f is $\mathrm{S}_d g$, then, with $\vec{x} = (x_1, \dots, x_d)$,

$$\mathrm{ct}(f, \vec{x}, y) := \langle \#f, \vec{x}, \mathrm{ct}(g, \vec{x}, 0, g(\vec{x}, 0)), \dots, \mathrm{ct}(g, \vec{x}, y, g(\vec{x}, y)), y \rangle.$$

Lemma 10. *The map*

$$(f, \vec{x}, y) \mapsto \mathrm{ct}(f, \vec{x}, y) : \{f \text{ a combinator, } f(\vec{x}) = y\} \to \mathbb{N}$$

is injective, and its image $T \subseteq \mathbb{N}$ is primitive recursive.

Proof. As before. □

Recall that we introduced the primitive recursive function

$$\alpha : \mathbb{N} \to \mathbb{N}, \qquad \alpha(z) := (z)_{\mathrm{lh}(z)}.$$

Theorem 11 (Kleene). *For each d we have a primitive recursive $T_d \subseteq \mathbb{N}^{d+2}$ such that for every d-ary combinator f and all $\vec{x} \in \mathbb{N}^d$:*

- *if $f(\vec{x})\downarrow$, then $T_d(\#f,\vec{x},z)$ for a unique z and $f(\vec{x})=\alpha(z)$ for this z;*
- *if $f(\vec{x})\uparrow$, then $T_d(\#f,\vec{x},z)$ for no z.*

Proof. Define $T_d \subseteq \mathbb{N}^{d+2}$ as follows:

$$\begin{aligned} T_d(e,\vec{x},z) \iff\ & z \text{ is the computation tree of some } d\text{-ary combinator} \\ & f \text{ with } \#f = e \text{ at input } \vec{x}=(x_1,\dots,x_d) \\ \iff\ & T(z), (z)_1 = e, (z)_2 = x_1, \dots, (z)_{d+1} = x_d, \alpha(e) = d. \end{aligned}$$

Thus, T_d is primitive recursive. If f is a d-ary combinator, and $f(\vec{x})=y$, then clearly $T_d(\#f,\vec{x},z)$ for a unique z, namely $z=\mathrm{ct}(f,\vec{x},y)$, and $y=(z)_{\mathrm{lh}(z)}=\alpha(z)$ for this z. If f is a d-ary combinator and $f(\vec{x})\uparrow$, then there is no z with $T_d(\#f,\vec{x},z)$. □

Corollary 12. *Suppose $\phi : \mathbb{N}^d \rightharpoonup \mathbb{N}$ is partial recursive. Then $\phi = \hat{f}$ for some d-ary combinator f with exactly one occurrence of an unbounded search symbol.*

Proof. Take any d-ary combinator g such that $\phi=\hat{g}$. Then for all $x_1,\dots,x_d$,

$$\phi(x_1,\dots,x_d) \simeq \alpha\big(\mu z(T_d(\#g,x_1,\dots,x_d,z))\big).$$

[*Note: The symbol $\simeq$ indicates that either both sides are defined and equal, or both sides are undefined.*] Thus $\phi=\hat{f}$ for some d-ary combinator f in which S_d occurs exactly once, and no other unbounded search symbol occurs. □

Definition 13. $\phi_e^{(d)} = \lambda x_1 \dots x_d . \alpha\big(\mu z(T_d(e,x_1,\dots,x_d,z))\big)\ :\ \mathbb{N}^d \rightharpoonup \mathbb{N}.$

Corollary 14. *Each $\phi_e^{(d)}$ is partial recursive, and for each partial recursive $\phi : \mathbb{N}^d \rightharpoonup \mathbb{N}$ there is an e such that $\phi = \phi_e^{(d)}$.*

Another consequence is that the recursive functions, as defined earlier, are exactly the computable functions, as defined in MATH 570.

Tracing back the definition of T_d in terms of combinators and computation trees we see that if $e=\#f$ with f a d-ary combinator, then $\phi_e^{(d)}=\hat{f}$, while if $e\neq\#f$ for all d-ary combinators f, then $\phi_e^{(d)}$ is the partial function $\mathbb{N}^d \rightharpoonup \mathbb{N}$ with empty domain. This fact is needed to prove:

Lemma 15. *Let $d\geq 1$ be given. Then there is a primitive recursive function $\rho:\mathbb{N}^2\to\mathbb{N}$ (depending on d) such that for all $x_1,\dots,x_d$,*

$$\phi_e^{(d)}(x_1,\dots,x_d) \simeq \phi_{\rho(e,x_1)}^{(d-1)}(x_2,\dots,x_d).$$

Before giving the proof we note that for all a,d there is a d-ary combinator c_a^d whose associated function is the constant (total) function $\mathbb{N}^d\to\mathbb{N}$ taking the value a. Indeed, one can construct c_a^d in such a way that, for any given d, the map

$$a\mapsto \#(\mathrm{c}_a^d):\mathbb{N}\to\mathbb{N}$$

is primitive recursive. We leave this construction as an exercise to the reader.

Proof. Let f be a d-ary combinator, $d \geq 1$. Then

$$f_a := \mathrm{S}^d_{d-1}\, f\, \mathrm{c}^{d-1}_a\, \mathrm{P}^1_{d-1} \cdots \mathrm{P}^{d-1}_{d-1}$$

is a $(d-1)$-ary combinator such that for all $x_2, \ldots, x_d$,

$$\hat{f}_a(x_2, \ldots, x_d) \simeq \hat{f}(a, x_2, \ldots, x_d).$$

Note that

$$\#(f_a) = \langle 3, \#f, \#(c^{d-1}_a), \langle 2, 1, d-1\rangle, \ldots, \langle 2, d-1, d-1\rangle, d-1\rangle.$$

Thus the primitive recursive function $\rho : \mathbb{N}^2 \to \mathbb{N}$ defined by

$$\rho(e, a) := \langle 3, e, \#(c^{d-1}_a), \langle 2, 1, d-1\rangle, \ldots, \langle 2, d-1, d-1\rangle, d-1\rangle$$

has the property that if f is any d-ary combinator with $\#f = e$, then f_a is a $(d-1)$-ary combinator with $\#(f_a) = \rho(e, a)$, while if there is no d-ary combinator f with $\#f = e$, then there is no $(d-1)$-ary combinator g with $\#g = \rho(e, a)$. It is clear from the remark preceding the lemma that this function ρ has the desired properties. □

If we have to indicate the dependence on d we let ρ_d be a function ρ as in the lemma. The next consequence has the strange name of "s-m-n theorem".

Corollary 16. *Given m, n, there is a primitive recursive $s^m_n : \mathbb{N}^{m+1} \to \mathbb{N}$ such that for all $e, x_1, \ldots, x_m, y_1, \ldots, y_n$:*

$$\phi^{(m+n)}_e(x_1, \ldots, x_m, y_1, \ldots, y_n) \simeq \phi^{(n)}_{s^m_n(e, x_1, \ldots, x_m)}(y_1, \ldots, y_n).$$

Proof. Take $s^0_n := \mathrm{id}_{\mathbb{N}}$, and put

$$s^{m+1}_n(e, x_1, \ldots, x_{m+1}) := \rho_{n+1}(s^m_n(e, x_1, \ldots, x_m), x_{m+1}).$$

□

The next result corresponds to the existence of a *universal computer*:

Corollary 17. *There is a partial recursive $\psi : \mathbb{N}^2 \rightharpoonup \mathbb{N}$ such that for all d, e and $\vec{x} \in \mathbb{N}^d$ we have*

$$\phi^{(d)}_e(\vec{x}) \simeq \psi(e, \langle\vec{x}\rangle).$$

Proof. Exercise. □

The next three results are due to Kleene, and are collectively referred to as the *Recursion Theorem.* They look a bit strange, but are very useful.

Theorem 18. *Let $f : \mathbb{N}^{d+1} \rightharpoonup \mathbb{N}$ be partial recursive. Then there exists an e_0 such that for all $\vec{x} \in \mathbb{N}^d$,*

$$f(e_0, \vec{x}) \simeq \phi^{(d)}_{e_0}(\vec{x}).$$

Proof. Define partial recursive $g : \mathbb{N}^{d+1} \rightharpoonup \mathbb{N}$ by $g(e, \vec{x}) \simeq f(s^1_d(e,e), \vec{x})$. Take a such that $g = \phi^{(d+1)}_a$. Then for all $\vec{x} \in \mathbb{N}^d$,

$$f(s^1_d(a,a), \vec{x}) \simeq g(a, \vec{x}) \simeq \phi^{(d+1)}_a(a, \vec{x}) \simeq \phi^{(d)}_{s^1_d(a,a)}(\vec{x}).$$

Thus $e_0 = s^1_d(a,a)$ works. □

From now on, $\phi_e := \phi^{(1)}_e$, so $\phi_0, \phi_1, \phi_2, \dots$ is an enumeration of the set of partial recursive functions $\mathbb{N} \rightharpoonup \mathbb{N}$.

Note also that the function $(e,x) \mapsto \phi_e(x) : \mathbb{N}^2 \rightharpoonup \mathbb{N}$ is partial recursive.

Theorem 19. *Let $h : \mathbb{N} \to \mathbb{N}$ be recursive. Then there exists an e_0 such that $\phi_{e_0} = \phi_{h(e_0)}$.*

Proof. Define a partial recursive $f : \mathbb{N}^2 \rightharpoonup \mathbb{N}$ by $f(e,x) \simeq \phi_{h(e)}(x)$. By the previous theorem we get $e_0 \in \mathbb{N}$ such that for all x,

$$\phi_{e_0}(x) \simeq f(e_0, x) \simeq \phi_{h(e_0)}(x),$$

so $\phi_{e_0} = \phi_{h(e_0)}$. □

Theorem 20. *There is a primitive recursive $\beta : \mathbb{N} \to \mathbb{N}$ such that for each total ϕ_e,*

$$\phi_{\beta(e)} = \phi_{\phi_e(\beta(e))}.$$

This is a uniform version of the previous theorem: given recursive $h : \mathbb{N} \to \mathbb{N}$ with index e, that is, $h = \phi_e$, the number $e_0 = \beta(e)$ satisfies $\phi_{e_0} = \phi_{h(e_0)}$.

Proof. We claim that $\beta : \mathbb{N} \to \mathbb{N}$ defined by $\beta(k) = s^1_1(s^1_2(b,k), s^1_2(b,k))$ does the job, where b is an index of the partial recursive function $g : \mathbb{N}^3 \rightharpoonup \mathbb{N}$ defined by $g(e,x,y) \simeq \phi_{\phi_e(s^1_1(x,x))}(y)$. Verification of claim:

$$\begin{aligned}
\phi_{\beta(e)}(y) &\simeq \phi_{s^1_1(s^1_2(b,e), s^1_2(b,e))}(y) \\
&\simeq \phi^{(2)}_{s^1_2(b,e)}(s^1_2(b,e), y) \\
&\simeq \phi^{(3)}_b(e, s^1_2(b,e), y) \\
&\simeq g(e, s^1_2(b,e), y) \\
&\simeq \phi_{\phi_e(s^1_1(s^1_2(b,e), s^1_2(b,e)))}(y) \\
&\simeq \phi_{\phi_e(\beta(e))}(y).
\end{aligned}$$

□

1.4 The Halting Problem and Rice's Theorem

Instead of saying that $A \subseteq \mathbb{N}^d$ is recursive, we also say that A is *decidable*. Assuming the Church-Turing Thesis, for A to be decidable means to have an effective procedure for deciding membership in A, that is, an algorithm that decides for any input $\vec{x} \in \mathbb{N}^d$ whether $\vec{x} \in A$.

Theorem 21 (Turing)**.** *The halting problem,*

$$K_0 := \{(e, n) : \phi_e(n) \downarrow\}$$

is undecidable.

Proof. Suppose otherwise. Then the set $\{e : \phi_e(e) \downarrow\}$ is decidable. Define $f : \mathbb{N} \rightharpoonup \mathbb{N}$ by

$$f(e) = \begin{cases} 0 & \text{if } \phi_e(e) \uparrow, \\ \uparrow & \text{if } \phi_e(e) \downarrow . \end{cases}$$

Note that f is partial recursive since $f(e) = \mu y(y = 0 \ \& \ \phi_e(e) \uparrow)$. So we have an e_0 such that $f = \phi_{e_0}$. But $f(e_0) \uparrow \iff \phi_{e_0}(e_0) \downarrow$, that is,

$$\phi_{e_0}(e_0) \uparrow \iff \phi_{e_0}(e_0) \downarrow,$$

a contradiction. □

This proof also shows that the set

$$K := \{e : \phi_e(e) \downarrow\}$$

is undecidable. Most natural undecidability results in mathematics can be reduced to the halting problem, although the reduction is not always obvious.

Let $\mathcal{A}$ be a set of partial recursive functions $\mathbb{N} \rightharpoonup \mathbb{N}$. In many cases it is natural to ask whether there is an effective procedure for deciding, for any unary combinator f, whether the associated partial function $\hat{f}$ belongs to $\mathcal{A}$. For example, is it decidable whether a unary combinator defines

(a) a partial function with nonempty domain?

(b) a partial function with finite domain?

(c) a partial function with finite image?

(d) a total function?

and so on. Assuming the Church-Turing Thesis, the next theorem of Rice says that all such questions have a negative answer. To understand this reading of Rice's theorem, note that one can effectively compute from any unary combinator f its index $e = \#f$, and in the other direction, we can decide for any given e whether e is the index of a unary combinator, and if so, find such a combinator. So these decision problems about unary combinators translate to equivalent decision problems about natural numbers.

Theorem 22 (Rice)**.** *Let $\mathcal{A}$ be a nonempty proper subset of the set of all partial recursive functions $\mathbb{N} \rightharpoonup \mathbb{N}$. Then $\{e : \phi_e \in \mathcal{A}\}$ is undecidable.*

Proof. Suppose $\{e : \phi_e \in \mathcal{A}\}$ is decidable. Take a, b such that $\phi_a \in \mathcal{A}$ and $\phi_b \notin \mathcal{A}$. Define partial recursive $f : \mathbb{N}^2 \rightharpoonup \mathbb{N}$ by

$$\begin{aligned} f(e,n) &\simeq \phi_a(n) \text{ if } \phi_e \notin \mathcal{A}, \\ f(e,n) &\simeq \phi_b(n) \text{ if } \phi_e \in \mathcal{A}. \end{aligned}$$

By the recursion theorem we have an e such that $\phi_e = f(e, .)$. If $\phi_e \notin \mathcal{A}$, then $\phi_e = f(e, .) = \phi_a \in \mathcal{A}$. If $\phi_e \in \mathcal{A}$, then $\phi_e = f(e, .) = \phi_b \notin \mathcal{A}$. Thus we have a contradiction. □

Of course, if $\mathcal{A}$ is empty or $\mathcal{A}$ is the entire set of partial recursive functions $\mathbb{N} \rightharpoonup \mathbb{N}$, then $\{e : \phi_e \in \mathcal{A}\}$ is empty or equal to $\mathbb{N}$, so the restrictions on $\mathcal{A}$ in Rice's theorem cannot be dropped.

Corollary 23. *The set $\{(a, b) : \phi_a = \phi_b\}$ is undecidable.*

Proof. If $\{(a, b) : \phi_a = \phi_b\}$ were decidable, so would be the set $\{a : \phi_a = \phi_0\}$. By Rice's Theorem, the latter set is undecidable, and thus the former must be undecidable. □

Here is another typical application of the recursion theorem:

There is an e such that $D(\phi_e) = \{e\}$.

To get such an e, let $\psi : \mathbb{N}^2 \rightharpoonup \mathbb{N}$ be the partial recursive function given by

$$\psi(x,y) \simeq \begin{cases} 0 & \text{if } x = y, \\ \uparrow & \text{if } x \neq y. \end{cases}$$

The recursion theorem gives an e such that $\psi(e, \cdot) = \phi_e$. Then $D(\phi_e) = \{e\}$.

1.5 Recursively Enumerable Sets

Definition 24. *A set $A \subseteq \mathbb{N}$ is called* recursively enumerable*, abbreviated* r.e. *(or* computably enumerable*, abbreviated* c.e.*) if $A = \emptyset$ or $A = f(\mathbb{N})$ for some recursive $f : \mathbb{N} \to \mathbb{N}$.*

Proposition 25. *Let $A \subseteq \mathbb{N}$. The following are equivalent:*

(a) A is recursively enumerable;

(b) $A = \mathrm{Im}(f)$ for some partial recursive $f : \mathbb{N} \rightharpoonup \mathbb{N}$;

(c) There is a primitive recursive $S \subseteq \mathbb{N}^2$ such that for all x,

$$x \in A \iff \exists y\ S(x,y);$$

(d) $A = D(f)$ *for some partial recursive function* $f : \mathbb{N} \rightharpoonup \mathbb{N}$*;*

(e) $A = \emptyset$*, or* $A = f(\mathbb{N})$ *for some primitive recursive function* $f : \mathbb{N} \to \mathbb{N}$*.*

Proof. The case that $A = \emptyset$ is trivial, so we assume that A is non-empty. The implication $(a) \implies (b)$ is clear. For $(b) \implies (c)$, let $A = \mathrm{Im}(\phi_e)$. Then

$$\begin{aligned} y \in A &\iff \exists x \exists z \big(T_1(e, x, z) \ \&\ y = \alpha(z) \big) \\ &\iff \exists n \big(T_1(e, (n)_1, (n)_2) \ \&\ y = \alpha((n)_2 \big). \end{aligned}$$

Now use that the set

$$\{(y, n) : T_1(e, (n)_1, (n)_2) \ \&\ y = \alpha((n)_2)\}$$

is primitive recursive. For $(c) \implies (d)$, let S be as in (c). Define a partial recursive function $f : \mathbb{N} \rightharpoonup \mathbb{N}$ by $f(x) \simeq \mu y\ S(x, y)$. Then $A = D(f)$.

For $(d) \implies (e)$, let $A = D(\phi_e)$. Then for all x,

$$x \in A \iff \exists z \big(T_1(e, x, z) \big).$$

Pick $a \in A$ and define $f : \mathbb{N} \to \mathbb{N}$ by

$$f(n) = \begin{cases} (n)_1 & \text{if } T_1\big(e, (n)_1, (n)_2\big), \\ a & \text{otherwise.} \end{cases}$$

Then f is primitive recursive with $f(\mathbb{N}) = A$. It is clear that $(e) \implies (a)$. □

If $A \subseteq \mathbb{N}$ is recursive, then A is recursively enumerable: use (d) above with $f : \mathbb{N} \rightharpoonup \mathbb{N}$ defined by

$$f(x) = \begin{cases} 1 & \text{if } x \in A, \\ \uparrow & \text{otherwise.} \end{cases}$$

For arbitrary d we call a set $A \subseteq \mathbb{N}^d$ *recursively enumerable* if there is a partial recursive function $f : \mathbb{N}^d \rightharpoonup \mathbb{N}$ such that $A = D(f)$.

Lemma 26. *Let* $A \subseteq \mathbb{N}^d$*. Then* A *is recursively enumerable iff* $A = \pi(S)$ *for some primitive recursive set* $S \subseteq \mathbb{N}^{d+1}$*, where* $\pi : \mathbb{N}^{d+1} \to \mathbb{N}^d$ *is given by* $\pi(x_1, \dots, x_d, y) = (x_1, \dots, x_d)$*.*

We leave the proof as an exercise. The reader should check that the lemma still holds when “primitive recursive” is replaced by “recursive”.

Lemma 27. *If* $A, B \subseteq \mathbb{N}^d$ *are recursively enumerable, so are* $A \cup B$ *and* $A \cap B$*. If* $C \subseteq \mathbb{N}^{m+n}$ *is recursively enumerable, then so is* $\pi(C) \subseteq \mathbb{N}^m$*, where*

$$\pi : \mathbb{N}^{m+n} \to \mathbb{N}^m, \quad \pi(x_1, \dots, x_m, y_1, \dots, y_n) = (x_1, \dots, x_m).$$

If $A \subseteq \mathbb{N}^d$ *and* $B \subseteq \mathbb{N}^e$ *are recursively enumerable, so is* $A \times B \subseteq \mathbb{N}^{d+e}$*.*

Proof. Suppose $A', B' \subseteq \mathbb{N}^{d+1}$ are primitive recursive such that

$$A = \{\vec{x} \in \mathbb{N}^d : \exists y \ A'(\vec{x}, y)\}, \quad B = \{\vec{x} \in \mathbb{N}^d : \exists y \ B'(\vec{x}, y)\}.$$

Then $A \cup B = \{\vec{x} \in \mathbb{N}^d : \exists y \ \big(A'(\vec{x}, y) \text{ or } B'(\vec{x}, y)\big)\}$. Since $A' \cup B'$ is primitive recursive, $A \cup B$ is recursively enumerable. Also,

$$A \cap B = \{\vec{x} \in \mathbb{N}^d : \exists y \ \big(A'(\vec{x}, (y)_1) \ \& \ B'(\vec{x}, (y)_2)\big)\},$$

hence $A \cap B$ is recursively enumerable.

Let $C' \subseteq \mathbb{N}^{m+n+1}$ be primitive recursive and

$$C = \{(\vec{x}, \vec{y}) \in \mathbb{N}^{m+n} : \exists z \ C'(\vec{x}, \vec{y}, z)\}.$$

Then

$$\pi(C) = \{\vec{x} \in \mathbb{N}^m : \exists y \ C'(\vec{x}, (y)_1, \dots, (y)_{n+1})\},$$

hence C is recursively enumerable.

The asssertion on cartesian products of r.e. sets is left as an exercise. □

Exercise. Show that if $A \subseteq \mathbb{N}^{d+1}$ is recursively enumerable, then so is $\forall^{\leq} A$.

Some notation: For $i = 1, \dots, n$, let $f_i : \mathbb{N}^m \rightharpoonup \mathbb{N}$ be such that $D(f_i) = D(f_j)$ for $i, j \in \{1, \dots, n\}$. This yields the partial map $f = (f_1, \dots, f_n) : \mathbb{N}^m \rightharpoonup \mathbb{N}^n$. For $A \subseteq \mathbb{N}^m$ and $B \subseteq \mathbb{N}^n$ we set

$$\begin{aligned} f(A) &:= f(A \cap D(f)) = \{f(\vec{x}) : \vec{x} \in A \cap D(f)\}, \\ f^{-1}(B) &:= \{\vec{x} \in D(f) : f(\vec{x}) \in B\}. \end{aligned}$$

We call f *partial recursive* if each f_i is partial recursive.

Lemma 28. *For $f : \mathbb{N}^m \rightharpoonup \mathbb{N}^n$ as above, f is partial recursive iff* $\operatorname{graph}(f) \subseteq \mathbb{N}^{m+n}$ *is recursively enumerable.*

Proof. We assume $n = 1$ since the general case follows easily from this case. Suppose $f = \phi_e^{(m)}$. Then for all $(\vec{x}, y) \in \mathbb{N}^{m+1}$,

$$(\vec{x}, y) \in \operatorname{graph}(f) \iff \exists z \big(T_m(e, \vec{x}, z) \ \& \ \alpha(z) = y\big).$$

For the converse, suppose $\operatorname{graph}(f) \subseteq \mathbb{N}^{m+1}$ is recursively enumerable. Take recursive $S \subseteq \mathbb{N}^{m+2}$ such that for all $(\vec{x}, y) \in \mathbb{N}^{m+1}$,

$$(\vec{x}, y) \in \operatorname{graph}(f) \iff \exists z \ S(\vec{x}, y, z).$$

Then $f(\vec{x}) \simeq \big(\mu n \ S(\vec{x}, (n)_1, (n)_2)\big)_1$. □

Corollary 29. *Let $f : \mathbb{N}^m \rightharpoonup \mathbb{N}^n$ be partial recursive, and let $A \subseteq \mathbb{N}^m$ and $B \subseteq \mathbb{N}^n$ be recursively enumerable. Then $f(A) \subseteq \mathbb{N}^n$ and $f^{-1}(B) \subseteq \mathbb{N}^m$ are recursively enumerable.*

Proof. Use that for all $\vec{y} \in \mathbb{N}^n$,

$$\vec{y} \in f(A) \iff \exists \vec{x}\big(A(\vec{x}) \ \&\ (\vec{x}, y) \in \mathrm{graph}(f)\big),$$

and that projecting a recursively enumerable set to $\mathbb{N}^n$ yields a recursively enumerable set. Also, for all $\vec{x} \in \mathbb{N}^m$,

$$\vec{x} \in f^{-1}(B) \iff \exists \vec{y} \in \mathbb{N}^n \big(B(\vec{y}) \ \&\ (\vec{x}, \vec{y}) \in \mathrm{graph}(f)\big),$$

showing that $f^{-1}(B)$ is recursively enumerable. □

It will be convenient to approximate ϕ_e by functions $\phi_{e,s}$ with finite domains. Here $s \in \mathbb{N}$ and $\phi_{e,s} : \mathbb{N} \rightharpoonup \mathbb{N}$ is given by

$$\phi_{e,s}(x) \simeq \begin{cases} \phi_e(x) & \text{if } \exists z \le s\big(T_1(e,x,z)\big), \\ \uparrow & \text{otherwise.} \end{cases}$$

Note that if $T_1(e,x,z)$, then $x \le z$, so $D(\phi_{e,s}) \subseteq \{0, \dots, s\}$.

We also use the following traditional notations:

$$W_e := D(\phi_e), \quad W_{e,s} := D(\phi_{e,s}).$$

Note that $(W_e)_{e \in \mathbb{N}}$ is an enumeration of the set of recursively enumerable subsets of $\mathbb{N}$, and that $W_e = K_0(e)$.

Corollary 30. *With these notations, we have:*

(a) The halting set

$$K_0 := \{(e,n) \in \mathbb{N}^2 : \phi_e(n) \downarrow\} = \{(e,n) : \ \exists \boldsymbol{z}\ T_1(e,n,z)\}$$

is recursively enumerable;

(b) $K := \{e \in \mathbb{N} : \ \phi_e(e) \downarrow\} = \{e \in \mathbb{N} : \ \exists z\ T_1(e,e,z)\}$ *is recursively enumerable;*

(c) $W_{e,s} \subseteq W_{e,s+1}$*, and* $\bigcup_{s \in \mathbb{N}} W_{e,s} = W_e$*;*

(d) The set $\{(e,n,s) \in \mathbb{N}^3 : n \in W_{e,s}\} = \{(e,n,s) \in \mathbb{N}^3 : \ \exists z \le s(T_1(e,n,z))\}$ *is primitive recursive.*

Theorem 31. *Let* $B \subseteq \mathbb{N}$ *be recursively enumerable. Then there is a primitive recursive function* $\beta : \mathbb{N} \to \mathbb{N}$ *such that for all* n*,* $B(n) \iff K(\beta(n))$*.*

Proof. Let $B = D(\phi_b)$. Define a partial recursive function $f : \mathbb{N}^2 \rightharpoonup \mathbb{N}$ by $f(x,y) \simeq \phi_b(x)$. Take c such that $f = \phi_c^{(2)}$. Then for all n,

$$\begin{aligned} B(n) &\iff \phi_b(n) \downarrow \iff f(n, s_1^1(c,n)) \downarrow \\ &\iff \phi_c^{(2)}(n, s_1^1(c,n)) \downarrow \iff \phi_{s_1^1(c,n)}(s_1^1(c,n)) \downarrow \iff K(s_1^1(c,n)). \end{aligned}$$

Thus we can take $\beta(n) := s_1^1(c,n)$ for this c. □

The significance of this result is that deciding membership in any recursively enumerable subset of $\mathbb{N}$ reduces computably to deciding membership in K.

1.6 Selections & Reductions

Theorem 32 (Selection Theorem)**.** *Let $A \subseteq \mathbb{N}^{d+1}$ be recursively enumerable. Then there is a partial recursive function $f : \mathbb{N}^d \rightharpoonup \mathbb{N}$ such that for all $\vec{x} \in \mathbb{N}^d$,*

(a) $f(\vec{x}) \downarrow \iff \exists y\, A(\vec{x}, y)$,

(b) $f(\vec{x}) \downarrow \implies A(\vec{x}, f(\vec{x}))$.

Proof. Take a primitive recursive $S \subseteq \mathbb{N}^{d+2}$ such that for all $(\vec{x}, y) \in \mathbb{N}^{d+1}$, $A(\vec{x}, y) \iff \exists z\, S(\vec{x}, y, z)$. Define $f : \mathbb{N}^d \rightharpoonup \mathbb{N}$ by

$$f(\vec{x}) \simeq \big(\mu n(S(\vec{x}, (n)_1, (n)_2)\big)_1.$$

Then f has the desired properties. □

An f as in the theorem above is called a partial recursive selector for A.

Theorem 33 (Reduction Theorem)**.** *Let $A, B \subseteq \mathbb{N}$ be recursively enumerable. Then there are recursively enumerable sets $C, D \subseteq \mathbb{N}$ such that $C \subseteq A$, $D \subseteq B$, $C \cap D = \emptyset$, and $C \cup D = A \cup B$.*

Proof. Let $E := (A \times \{0\}) \cup (B \times \{1\}) \subseteq \mathbb{N}^2$. Then E is recursively enumerable. Take a partial recursive selector $f : \mathbb{N} \rightharpoonup \mathbb{N}$ for E. Let $C = f^{-1}(0)$, $D = f^{-1}(1)$. Then C and D have the desired properties. □

Theorem 34 (Post's Theorem)**.** *A set $A \subseteq \mathbb{N}$ is recursive iff A and $\mathbb{N} \setminus A$ are recursively enumerable.*

Proof. The left-to-right direction is clear. Assume that A and $\mathbb{N} \setminus A$ are recursively enumerable. Let

$$E := ((\mathbb{N} \setminus A) \times \{1\}) \cup (A \times \{0\}).$$

By assumption, $E \subseteq \mathbb{N}^2$ is recursively enumerable. Take a partial recursive selector $f : \mathbb{N} \rightharpoonup \mathbb{N}$ for E. Then $D(f) = \mathbb{N}$, so f is a recursive function. Since $f = \chi_A$, we obtain that A is recursive. □

Post's Theorem is easy to explain by the Church-Turing Thesis: Suppose A and $\mathbb{N}\setminus A$ are recursively enumerable and nonempty, and let $f, g : \mathbb{N} \to \mathbb{N}$ computably enumerate these sets. Since the union of these two sets is $\mathbb{N}$, every natural number must appear in exactly one of the sequences $f(0), f(1), f(2), \ldots$ and $g(0), g(1), g(2), \ldots$. Hence the computable sequence

$$f(0), g(0), f(1), g(1), f(2), g(2), \ldots$$

enumerates all of $\mathbb{N}$, and thus provides a computable way to determine whether or not a given n is in A.

The sets $A, B \subseteq \mathbb{N}$ are called *recursively separable* if there is a recursive set $R \subseteq \mathbb{N}$ such that $A \subseteq R$ and $B \cap R = \emptyset$ (so A and B are disjoint). Such an R is said to *recursively separate* A from B.

Theorem 35. *There are disjoint recursively enumerable sets $A, B \subseteq \mathbb{N}$ that cannot be recursively separated.*

Proof. Let $A = \{e : \phi_e(e) \simeq 0\}$, and $B = \{e : \phi_e(e) \simeq 1\}$. Then A and B are disjoint and recursively enumerable.

Assume that $R \subseteq \mathbb{N}$ recursively separates A and B. Let $e_0 \in \mathbb{N}$ satisfy $\chi_R = \phi_{e_0}$. Then

$$\begin{aligned} e_0 \in R &\iff \phi_{e_0}(e_0) \simeq 1 \iff e_0 \in B \implies e_0 \notin R, \\ e_0 \notin R &\iff \phi_{e_0}(e_0) \simeq 0 \iff e_0 \in A \implies e_0 \in R, \end{aligned}$$

a contradiction. □

1.7 Trees and Recursive Inseparability

We recall that $2^{\mathbb{N}}$ is the set of all functions $\beta : \mathbb{N} \to \{0, 1\}$, that is, the set of all characteristic functions of subsets of $\mathbb{N}$, and we think of $\beta \in 2^{\mathbb{N}}$ as the infinite sequence $\beta(0), \beta(1), \beta(2), \ldots$ of zeros and ones. Graphically we represent $2^{\mathbb{N}}$ as the *infinite binary tree*, suggested in the picture below. For example, a sequence $0, 1, 1, 0, 0, \ldots$ represents the branch in the tree that travels from the top down via “left, right, right, left, left, . . . ” , with “0” representing “go left” and “1” representing “go right”.

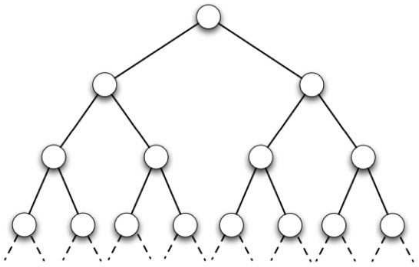


Figure 1.1: The complete infinite binary tree

We also let $2^{<\mathbb{N}}$ be the set of all *finite* sequences $\vec{s} = (s_0, \ldots, s_{n-1})$ such that $s_i \in \{0, 1\}$ for $i < n$; note that this includes an empty sequence for $n = 0$. The elements of $2^{<\mathbb{N}}$ correspond to the nodes in the tree above, the empty sequence corresponding to the top node.

Let $\vec{s} = (s_0, \ldots, s_{n-1}) \in 2^{<\mathbb{N}}$. The *length* of $\vec{s}$ is the number $|\vec{s}| := n$. An *initial segment* of $\vec{s}$ is a finite sequence $(s_0, \ldots, s_{m-1})$ with $m \leq n$. A *binary tree* is a non-empty set $D \subseteq 2^{<\mathbb{N}}$ closed under initial segments.

An *infinite branch* of a binary tree D is a sequence $\beta \in 2^{\mathbb{N}}$ such that for all n, $(\beta(0), \ldots, \beta(n-1)) \in D$. A set $D \subseteq 2^{<\mathbb{N}}$ is said to be *recursive* if the set

$$\{\langle s_0, \ldots, s_{n-1}\rangle : (s_0, \ldots, s_{n-1}) \in D\} \subseteq \mathbb{N}$$

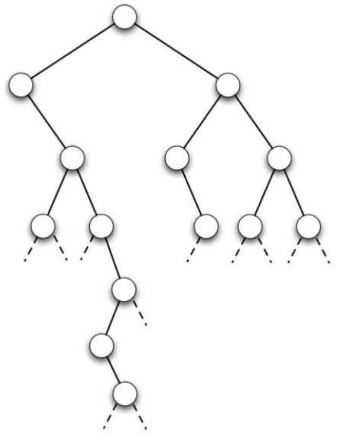


Figure 1.2: A binary tree

is recursive.

Lemma 36. *The full binary tree $2^{<\mathbb{N}}$ is recursive.*

Proof. Let $A := \{\langle s_0, \dots, s_{n-1}\rangle : (s_0, \dots, s_{n-1}) \in 2^{<\mathbb{N}}\}$. Then

$$x \in A \iff x \in B \wedge \forall i \leq \mathrm{lh}(x)\big(i = 0 \vee (x)_i = 0 \vee (x)_i = 1\big).$$

Thus A is recursive. □

Lemma 37 (König)**.** *Every infinite binary tree has an infinite branch.*

Proof. Let D be an infinite binary tree. We define $\beta : \mathbb{N} \to \{0, 1\}$ inductively such that β is an infinite branch of D.

$$\beta(0) = \begin{cases} 0 & \text{if there are infinitely many } \vec{s} \in D \text{ such that } s_0 = 0, \\ 1 & \text{otherwise,} \end{cases}$$

and for $n > 0$

$$\beta(n) = \begin{cases} 0 & \text{if there are infinitely many } \vec{s} \in D \text{ such that} \\ & |\vec{s}| > n,\ s_n = 0 \text{ and } s_i = \beta(i) \text{ for } i < n, \\ 1 & \text{otherwise.} \end{cases}$$

Induction on n shows that for each n there are infinitely many $\vec{s} \in D$ such that $|\vec{s}| > n$ and $\beta(0) = s_0, \dots, \beta(n) = s_n$. It follows that β is an infinite branch of D. (It is actually the "leftmost" such branch.) □

Theorem 38. *There exists an infinite recursive binary tree with no recursive infinite branch.*

Proof. Let $A_0, A_1 \subseteq \mathbb{N}$ be disjoint and recursively inseparable r.e. sets. Thus we have recursive sets $S_0, S_1 \subseteq \mathbb{N}^2$ such that for all m,

$$\begin{aligned} A_0(m) &\iff \exists n S_0(m,n) \\ A_1(m) &\iff \exists n S_1(m,n) \end{aligned}$$

Next, define $D \subseteq 2^{<\mathbb{N}}$ as follows; for $\vec{s} = (s_0, \ldots, s_{m-1}) \in 2^{<\mathbb{N}}$,

$$\vec{s} \in D \iff \forall k < m\, [(\exists n < m\ (k,n) \in S_i) \Rightarrow s_k = i] \text{ for } i = 0, 1.$$

It is easy to check that D is a recursive binary tree. To see that D is infinite we show how to construct, for any m, a sequence of length m in D. The sequence $(s_0, \ldots, s_{m-1})$ defined such that for $k < m$, if there exists an $n < m$ with $(k,n) \in S_1$, then $s_k = 1$ and otherwise $s_k = 0$, satisfies our requirements.

Next we show that for any infinite branch $\beta : \mathbb{N} \to \{0,1\}$ of D, the set $R \subseteq \mathbb{N}$ such that $\chi_R = \beta$, is not recursive. Specifically, we show that R separates A_0 and A_1, and thus cannot be recursive.

Let $k \in A_i$ and take n such that $(k,n) \in S_i$. Consider any sequence $(\beta(0), \ldots, \beta(m-1))$ where $m > k, n$; note that $(\beta(0), \ldots, \beta(m-1)) \in D$ by assumption. Then by definition $\beta(k) = i$, so $k \in A_0 \Rightarrow k \notin R$ and $k \in A_1 \Rightarrow k \in R$, so $R \cap A_0 = \emptyset$ and $A_1 \subseteq R$. If R was recursive this would contradict that A_0 and A_1 are not recursively separable. □

1.8 Indices and Enumeration

In this section $\psi = \{\psi_e^n\}$ is an **index system**, that is, for each n,

$$\{\psi_e^n : e = 0, 1, \ldots\} = \{f : \mathbb{N}^n \rightharpoonup \mathbb{N} : f \text{ is partial recursive}\}.$$

Our earlier work shows that $\{\phi_e^{(n)}\}$ is an index system.

Definition 39. *We say that ψ satisfies* enumeration *if for each n there is an a such that*

$$\psi_a^{n+1}(e, \vec{x}) \simeq \psi_e^n(\vec{x}) \quad \textit{for all } (e, \vec{x}) \in \mathbb{N}^{1+n},$$

equivalently, $\lambda e\vec{x}.\psi_e^n(\vec{x}) : \mathbb{N}^{1+n} \rightharpoonup \mathbb{N}$ is partial recursive, for each n.
We say that ψ satisfies parametrization *if for all m, n there is a recursive $s : \mathbb{N}^{1+n} \to \mathbb{N}$ such that*

$$\psi_{s(e,\vec{x})}^n(\vec{y}) \simeq \psi_e^{m+n}(\vec{x}, \vec{y}) \quad \textit{for all } (e, \vec{x}, \vec{y}) \in \mathbb{N}^{1+m+n}.$$

We say that ψ is acceptable *if for each n there are recursive $f_n, g_n : \mathbb{N} \to \mathbb{N}$ such that for all e*

$$\psi_e^n = \phi_{f_n(e)}^{(n)} \textit{ and } \phi_e^{(n)} = \psi_{g_n(e)}^n$$

The index system $\{\phi_e^{(n)}\}$ is trivially acceptable, and earlier results show that it satisfies enumeration and parametrization. We put $\psi_e := \psi_e^1$.

Proposition 40. *ψ is acceptable if and only if ψ satisfies enumeration and parametrization.*

Proof. Assume ψ is acceptable and let $\{f_n, g_n\}_n$ witness this as in the definition. To prove enumeration, fix n and take c such that $\phi_k^{(n)}(\vec{x}) \simeq \phi_c^{(n+1)}(k, \vec{x})$ for all $(k, \vec{x}) \in \mathbb{N}^{1+n}$. Then for all $e, \vec{x}$,

$$\psi_e^n(\vec{x}) \simeq \phi_{f_n(e)}^{(n)}(\vec{x}) \simeq \phi_c^{(n+1)}(f_n(e), \vec{x}).$$

The right hand side is partial recursive as a function of $(e, \vec{x})$, so we have b such that

$$\phi_c^{(n+1)}(f_n(e), \vec{x}) \simeq \phi_b^{(n+1)}(e, \vec{x}) \simeq \psi_{g_{n+1}(b)}^{n+1}(e, \vec{x}) \quad \text{for all } (e, \vec{x}).$$

So for $a = g_{n+1}(b)$, we have $\psi_e^n(\vec{x}) \simeq \psi_a^{n+1}(e, \vec{x})$ for all $e, \vec{x}$.

To prove parametrization, fix m, n. Then enumeration gives a such that for all $(e, \vec{x}, \vec{y}) \in \mathbb{N}^{1+m+n}$,

$$\begin{aligned}
\psi_e^{m+n}(\vec{x}, \vec{y}) &\simeq \psi_a^{1+m+n}(e, \vec{x}, \vec{y}) \\
&\simeq \phi_{f_{1+m+n}(a)}^{(1+m+n)}(e, \vec{x}, \vec{y}) \\
&\simeq \phi_{s_n^{1+m}(f_{1+m+n}(a), e, \vec{x})}^{(n)}(\vec{y}) \\
&\simeq \psi_{g_n(s_n^{1+m}(f_{1+m+n}(a), e, \vec{x}))}^{n}(\vec{y}),
\end{aligned}$$

so ψ satisfies parametrization.

Conversely, assume ψ satisfies enumeration and parametrization. Fix n, and take a such that $\psi_a^{1+n}(e, \vec{x}) \simeq \psi_e^n(\vec{x})$ for all $(e, \vec{x}) \in \mathbb{N}^{1+n}$. Also, take b such that $\phi_b^{(1+n)}(e, \vec{x}) \simeq \psi_a^{1+n}(e, \vec{x})$ for all $e, \vec{x}$, hence

$$\phi_{s_n^1(b,e)}^{(n)}(\vec{x}) \simeq \psi_a^{1+n}(e, \vec{x}) \simeq \psi_e^n(\vec{x}) \quad \text{for all } e, \vec{x}.$$

Then $f_n : \mathbb{N} \to \mathbb{N}$ given by $f_n(e) = s_n^1(b, e)$ satisfies

$$\psi_e^n = \phi_{f_n(e)}^{(n)}.$$

Likewise we can get a recursive $g_n : \mathbb{N} \to \mathbb{N}$ such that for all e we have

$$\phi_e^{(n)} = \psi_{g_n(e)}^n. \qquad \square$$

We need an effective form of the recursion theorem for acceptable ψ.

Corollary 41. *Assume ψ is acceptable. There is a recursive $\beta : \mathbb{N} \to \mathbb{N}$ so that for any recursive $h : \mathbb{N} \to \mathbb{N}$ and any d with $h = \psi_d$ the number $e = \beta(d)$ satisfies $\psi_e = \psi_{h(e)}$.*

Proof. Repeat the proofs of Kleene's recursion theorems 18, 19, and 20, with the system ψ instead of $\{\phi_e^{(n)}\}$, using that these proofs only rely on enumeration and parametrization. $\square$

Lemma 42 (Padding Lemma)**.** *Assume ψ is acceptable. Then we can effectively construct, for any a, infinitely many b such that $\psi_a = \psi_b$.*

Proof. Given a and finite $D \subseteq \mathbb{N}$ such that $\psi_d = \psi_a$ for all $d \in D$, we shall construct $e \notin D$ with $\psi_e = \psi_a$. Define partial recursive $g : \mathbb{N}^2 \rightharpoonup \mathbb{N}$ by

$$g(e,x) \simeq \begin{cases} \psi_a(x) & \text{if } e \notin D \\ \uparrow & \text{if } e \in D. \end{cases}$$

Enumeration and parametrization give a recursive $f : \mathbb{N} \to \mathbb{N}$ such that for all $(e,x) \in \mathbb{N}^2$,

$$\psi_{f(e)}(x) \simeq g(e,x)$$

The recursion theorem yields e_0, effectively from a, D, with $\psi_{e_0} = \psi_{f(e_0)}$. If $e_o \notin D$, then $\psi_{e_0} = \psi_a$ and we are done. Suppose $e_0 \in D$; then $\psi_a = \psi_{e_0}$, hence $D(\psi_a) = D(\psi_{e_0}) = \emptyset$. So it remains to construct effectively an index $e \notin D$ with $D(\psi_e) = \emptyset$. As before, we get a recursive $h : \mathbb{N} \to \mathbb{N}$ such that

$$\psi_{h(i)}(x) = \begin{cases} 0 & \text{if } i \in D \\ \uparrow & \text{if } i \notin D \end{cases}$$

The recursion theorem gives e_1 with $\psi_{h(e_1)} = \psi_{e_1}$. If $e_1 \in D$, then

$$\psi_a = \psi_{e_1} = \psi_{h(e_1)}, \quad \text{a function with nonempty domain,}$$

contradicting $D(\psi_a) = \emptyset$. So $e_1 \notin D$, hence $D(\psi_{e_1}) = \emptyset$, and we are done. □

Theorem 43. *The index system ψ is acceptable if and only if for each n, there is a recursive bijection $h : \mathbb{N} \to \mathbb{N}$ such that $\psi^n_e = \phi^{(n)}_{h(e)}$ for all e.*

Proof. Such a bijection for each n witnesses the acceptability condition for ψ. Conversely, suppose ψ is acceptable. We shall construct h as in the theorem for $n = 1$, and we leave it to the reader to deal with arbitrary n.

We have recursive functions $f, g : \mathbb{N} \to \mathbb{N}$ such that $\psi_d = \phi_{f(d)}$ and $\phi_e = \psi_{g(e)}$ for all d, e. We construct h in stages where at each stage n we have a partial function $h_n : \mathbb{N} \rightharpoonup \mathbb{N}$ with $|D(h_n)| = n$.

Stage 0**:** Take h_0 with empty domain.

Stage $2n+1$**:** Take the least $d \notin D(h_{2n})$ and use f and the padding lemma for ϕ to obtain effectively an $e \notin \mathrm{Im}(h_{2n})$ such that $\psi_d = \phi_e$. Define h_{2n+1} as an extension of h_{2n} with $D(h_{2n+1}) = D(h_{2n}) \cup \{d\}$ with $h_{2n+1}(d) = e$.

Stage $2n+2$**:** Take the least $e \notin \mathrm{Im}(h_{2n+1})$ and use g and the padding lemma to obtain effectively a $d \notin D(h_{2n+1})$ such that $\psi_d = \phi_e$. Define h_{2n+2} as an extension of h_{2n+1} with $D(h_{2n+2}) = D(h_{2n+1}) \cup \{d\}$ with $h_{2n+2}(d) = e$.

Then $h = \bigcup h_n$ has the desired properties. □

Chapter 2

Hilbert's 10th Problem

2.1 Introduction

Recursively enumerable sets occur outside pure recursion theory, most strikingly in *Hilbert's 10th Problem*, hereafter denoted $H10$:

Question. Is there an algorithm that decides for any given polynomial

$$f(X_1, \ldots, X_n) \in \mathbb{Z}[X_1, \ldots, X_n]$$

whether the equation

$$f(X_1, \ldots, X_n) = 0 \tag{2.1}$$

has a solution in $\mathbb{Z}^n$? A solution of (2.1) in $\mathbb{Z}^n$ (or integer solution of (2.1)) is a tuple $(x_1, \ldots, x_n) \in \mathbb{Z}^n$ such that $f(x_1, \ldots, x_n) = 0$. Likewise we define what we mean by a solution of (2.1) in $\mathbb{N}^n$ (or natural number solution).

The answer to $H10$ is known since 1970: there is no such algorithm. It is still open whether such an algorithm exists when n is fixed to be 2.

Observation. The following are equivalent:

(a) There is an algorithm that decides for any $f(X_1, \ldots, X_n) \in \mathbb{Z}[X_1, \ldots, X_n]$ whether (2.1) has a solution in $\mathbb{Z}^n$.

(b) There is an algorithm that decides for any $f(X_1, \ldots, X_n) \in \mathbb{Z}[X_1, \ldots, X_n]$ whether (2.1) has a solution in $\mathbb{N}^n$.

Proof. Note that (2.1) has a solution in $\mathbb{Z}^n$ if and only if one of the 2^n equations

$$f(q_1 X_1, \ldots, q_n X_n) = 0,$$

with all $q_i \in \{1, -1\}$, has a solution in $\mathbb{N}^n$. This yields $(b) \implies (a)$.

For the other direction we use Lagrange's four-square theorem. By this theorem, (2.1) has a solution in $\mathbb{N}^n$ if and only if the equation

$$f(X_{11}^2 + X_{12}^2 + X_{13}^2 + X_{14}^2, \dots, X_{n1}^2 + X_{n2}^2 + X_{n3}^2 + X_{n4}^2) = 0$$

has a solution in $\mathbb{Z}^n$. □

In the following account of $H10$, we use the Observation above to restrict attention to natural number solutions for equations of type (2.1). Thus from now on in this chapter, u, x, y, z (sometimes with subscripts) range over $\mathbb{N}$. This is in addition to the convention that $a, b, c, d, e, k, l, m, n$ range over $\mathbb{N}$.

Let $A_1, \dots, A_m, X_1, \dots, X_n$ be distinct polynomial indeterminates, $A = (A_1, \dots, A_m)$, $X = (X_1, \dots, X_n)$, and $f(A, X) \in \mathbb{Z}[A, X]$. We consider $f(A, X)$ as defining the family of diophantine equations

$$f(\vec{a}, X) = 0, \qquad \vec{a} \in \mathbb{N}^m.$$

Definition 44. *The* diophantine set *defined by $f(A, X)$ is*

$$\{\vec{a} \in \mathbb{N}^m : \exists \vec{x}\, (f(\vec{a}, \vec{x}) = 0)\}$$

A diophantine set in $\mathbb{N}^m$ *is a diophantine set defined by some polynomial $f(A, X) \in \mathbb{Z}[A, X]$, where $X = (X_1, \dots, X_n)$ and n can be arbitrary.*

Lemma 45. *Diophantine sets are recursively enumerable.*

Proof. Let $f = f(A, X) \in \mathbb{Z}\,[A, X]$ be as before, and take $g, h \in \mathbb{N}\,[A, X]$ such that $f = g - h$. Then

$$\{(\vec{a}, \vec{x}) \in \mathbb{N}^{m+n} : f(\vec{a}, \vec{x}) = 0\} = \{(\vec{a}, \vec{x}) \in \mathbb{N}^{m+n} : g(\vec{a}, \vec{x}) = h(\vec{a}, \vec{x})\},$$

so $\{(\vec{a}, \vec{x}) \in \mathbb{N}^{m+n} : f(\vec{a}, \vec{x}) = 0\}$ is primitive recursive. Thus the set

$$\{\vec{a} \in \mathbb{N}^m : \exists \vec{x}\, (f(\vec{a}, \vec{x}) = 0)\}$$

is recursively enumerable. □

Suppose we have a *nonrecursive* diophantine set $D \subseteq \mathbb{N}^m$. Then, assuming the Church-Turing Thesis, the answer to $H10$ is negative. To see why, take $f(A, X) \in \mathbb{Z}\,[A, X]$ as above such that

$$D = \{\vec{a} \in \mathbb{N}^m : f(\vec{a}, X_1, \dots, X_n) = 0 \text{ has a solution in } \mathbb{N}^n\}$$

Since D is not recursive, the Church-Turing Thesis says that there cannot exist an algorithm deciding for any $\vec{a} \in \mathbb{N}^m$ as input, whether the equation $f(\vec{a}, X) = 0$ has a solution in $\mathbb{N}^n$. In particular, there cannot exist an algorithm as demanded by $H10$.

The following is the key result. It gives much more than just a nonrecursive diophantine set.

Main Theorem [Matiyasevich 1970]. *The diophantine sets in $\mathbb{N}^m$ are exactly the recursively enumerable sets in $\mathbb{N}^m$.*

It follows that there is no algorithm as required in $H10$ (assuming the Church-Turing Thesis). Suggestive results in the direction of Matiyasevich's theorem were obtained by M. Davis, H. Putnam, and J. Robinson in the 50's and 60's. Basically, they reduced the problem to showing that the set

$$\{(a,b) : b = 2^a\} \subseteq \mathbb{N}^2$$

is diophantine. This reduction is given in the next section.

2.2 Reduction to Exponentiation

Lemma 46. *The following sets are diophantine.*

(a) $\{(a,b) : a = b\} \subseteq \mathbb{N}^2$,

(b) $\{(a,b) : a \leq b\} \subseteq \mathbb{N}^2$,

(c) $\{(a,b) : a < b\} \subseteq \mathbb{N}^2$,

(d) $\{(a,b) : a|b\} \subseteq \mathbb{N}^2$,

(e) $\{(a,b,c) : a \equiv b \ mod \ c\} \subseteq \mathbb{N}^3$,

(f) $\{(a,b,c) : c = |a,b|\} \subseteq \mathbb{N}^3$,

(g) $\{a : a \neq 2^n \ for \ all \ n\} \subseteq \mathbb{N}$.

Proof. Just note the following equivalences:

(a) $a = b \iff a - b = 0$;

(b) $a \leq b \iff \exists x \ a + x = b$;

(c) $a < b \iff \exists x \ a + x + 1 = b$;

(d) $a|b \iff \exists x \ ax = b$;

(e) $a \equiv b \text{ mod } c \iff \exists x \exists y \ a - b = cx - cy$;

(f) $c = |a,b| \iff 2c = (a+b)(a+b+1) + 2a$;

(g) $a \neq 2^n$ for all $n \iff \exists x \exists y \ (2x+3)y = a$.

□

Lemma 47. *Being diophantine is preserved under certain operations:*

(a) If $D, E \subseteq \mathbb{N}^m$ are diophantine, so are $D \cup E, D \cap E \subseteq \mathbb{N}^m$.

(b) If $D \subseteq \mathbb{N}^{m+n}$ is diophantine, so is $\pi(D) \subseteq \mathbb{N}^m$ where $\pi : \mathbb{N}^{m+n} \to \mathbb{N}^m$ is defined by $\pi(x_1, \ldots, x_m, y_1, \ldots, y_n) = (x_1, \ldots, x_m)$.

(c) A set $D \subseteq \mathbb{N}^m$ is diophantine if and only if D is existentially definable *in the ordered semiring* $(\mathbb{N};\ 0,\ 1,\ +,\ \cdot,\ \leq)$.

Proof.

(a) Let $D, E \subseteq \mathbb{N}^m$ be diophantine, and take $f, g \in \mathbb{Z}\,[A; X]$ such that

$$\begin{aligned} D &= \{\vec{a} :\ \exists \vec{x}\ f(\vec{a}, \vec{x}) = 0\}, \\ E &= \{\vec{a} :\ \exists \vec{x}\ g(\vec{a}, \vec{x}) = 0\}. \end{aligned}$$

Then clearly

$$\begin{aligned} D \cup E &= \{\vec{a} :\ \exists \vec{x}\ f(\vec{a}, \vec{x}) g(\vec{a}, \vec{x}) = 0\}, \\ D \cap E &= \{\vec{a} :\ \exists \vec{x}\ f(\vec{a}, \vec{x})^2 + g(\vec{a}, \vec{x})^2 = 0\}. \end{aligned}$$

(b) Let $D \subseteq \mathbb{N}^{m+n}$ be diophantine, and take $f \in \mathbb{Z}\,[A, B, X]$ such that

$$D = \{(\vec{a}, \vec{b}) :\ \exists \vec{x}\ f(\vec{a}, \vec{b}, \vec{x}) = 0\}$$

Then $\pi(D) = \{\vec{a} :\ \exists \vec{b}\, \exists \vec{x}\ f(\vec{a}, \vec{b}, \vec{x}) = 0\}$.

(c) For $(\Rightarrow)$ use that for each $f \in \mathbb{Z}[A, X]$ there exist $g, h \in \mathbb{N}\,[A, X]$ such that $f = g - h$.

For the $(\Leftarrow)$ direction, note that for any quantifier free formula in the language of $(\mathbb{N}; 0, 1, +, \cdot, <)$ you can get rid of inequalities, negations, disjunctions, and conjunctions at the cost of introducing extra existentially quantified variables by means of the following equivalences:

$$\begin{aligned} x \neq y &\iff \exists z\, (x + z + 1 = y\ \vee\ y + z + 1 = x)\,, \\ x = 0 \vee y = 0 &\iff xy = 0, \\ x = 0 \wedge y = 0 &\iff x^2 + y^2 = 0, \\ x < y &\iff \exists z\, (x + z + 1 = y)\,. \end{aligned}$$

□

Definition 48. *A* diophantine function $f : \mathbb{N}^m \to \mathbb{N}$ *is a function whose graph is a diophantine set in* $\mathbb{N}^{m+1}$.

For example, addition and multiplication are diophantine functions on $\mathbb{N}^2$, and so is the monus function. Also $\gcd : \mathbb{N}^2 \to \mathbb{N}$ is diophantine:

$$\gcd(a, b) = c \iff \exists x, y\big((ax - by = c \vee by - ax = c)\ \wedge\ c|a \wedge c|b\big).$$

Lemma 49. *Let $f : \mathbb{N}^m \to \mathbb{N}$, $R \subseteq \mathbb{N}^m$ and $g_1, \dots, g_m : \mathbb{N}^n \to \mathbb{N}$ be diophantine. Then the function $f(g_1, \dots, g_m) : \mathbb{N}^n \to \mathbb{N}$, and the relation $R(g_1, \dots, g_m) \subseteq \mathbb{N}^n$ defined by*

$$R(g_1, \dots, g_m)(\vec{a}) \iff R(g_1(\vec{a}), \dots, g_m(\vec{a}))$$

are diophantine.

Proof. Exercise. □

Some terminology. A *polynomial zero set* in $\mathbb{N}^n$ is a set $\{\vec{x} \in \mathbb{N}^n : f(\vec{x}) = 0\}$ where $f(X) \in \mathbb{Z}[X]$. A *projection* of a set $R \subseteq \mathbb{N}^n$ is a set $S = \pi(R) \subseteq \mathbb{N}^m$ with $m \leq n$ and $\pi : \mathbb{N}^n \to \mathbb{N}^m$ given by $\pi(x_1, \dots, x_n) = (x_1, \dots, x_m)$.

Thus polynomial zero sets are primitive recursive, and the diophantine sets are just the projections of polynomial zero sets.

The *bounded universal quantification* of a set $R \subseteq \mathbb{N}^{n+2}$, denoted $\forall^{\leq} R$, is the subset of $\mathbb{N}^{n+1}$ defined by: $(\forall^{\leq} R)(\vec{x}, y) \iff \forall i_{\leq y}\, R(\vec{x}, y, i)$. Let Σ be the smallest subset of $\bigcup_m \mathcal{P}(\mathbb{N}^m)$ that contains all polynomial zero sets and is closed under taking projections and bounded universal quantifications.

Proposition 50. $\Sigma = \bigcup_m \{S : S$ *is a recursively enumerable set in* $\mathbb{N}^m\}$.

Before we can prove this we need to know that Σ is closed under some elementary operations like taking cartesian products and intersections. This is basically an exercise in predicate logic, which we now carry out.

Instead of $\mathbb{N}$ and its ordering we may consider in this exercise any nonempty set A with a binary relation $\vartriangle$ on A. Till further notice we let a, b, x, y, sometimes with indices, range over A, with vector notation like $\vec{a}$ used to denote a tuple $(a_1, \dots, a_n)$ of the appropriate length n. For $R \subseteq A^{n+2}$ we define the *bounded universal quantification* $\forall^{\vartriangle} R \subseteq A^{n+1}$ of R by

$$(\forall^{\vartriangle} R)(\vec{a}, b) :\iff \forall y_{\vartriangle b} R(\vec{a}, b, y) :\iff \forall y\big(\, y \vartriangle b \;\Rightarrow\; R(\vec{a}, b, y)\big),$$

For $n \geq m$, let $\pi^n_m : A^n \to A^m$ be given by $\pi^n_m(x_1, \dots, x_n) = (x_1, \dots, x_m)$. Taking cartesian products commutes with projecting:

$$D \subseteq A^d,\ m \leq n,\ S \subseteq A^n \implies D \times \pi^n_m(S) = \pi^{d+n}_{d+m}(D \times S).$$

It also commutes with bounded universal quantification:

$$D \subseteq A^d,\ R \subseteq A^{n+2} \implies \forall^{\vartriangle}(D \times R) = D \times \forall^{\vartriangle} R.$$

For $\lambda : \{1, \dots, m\} \to \{1, \dots, n\}$ and $S \subseteq A^m$ we define

$$\lambda(S) := \{(x_1, \dots, x_n) : (x_{\lambda(1)}, \dots, x_{\lambda(m)}) \in S\} \subseteq A^n.$$

Let for each n a collection $\mathcal{C}_n$ of subsets of A^n be given such that:

- $A^0 \in \mathcal{C}_0$ and $\Delta := \{(x, y) : x = y\} \in \mathcal{C}_2$;

- $C, D \in \mathcal{C}_n \implies C \cap D \in \mathcal{C}_n$;
- $\lambda : \{1, \dots, m\} \to \{1, \dots, n\}$ is injective, $C \in \mathcal{C}_m \;\Rightarrow\; \lambda(C) \;\in\; \mathcal{C}_n$.

With $m = 0$ in the last condition we obtain $A^n \in \mathcal{C}$ for all n. More generally, the inclusion map $\{1, \dots, m\} \to \{1, \dots, m+n\}$ shows that if $C \in \mathcal{C}_m$, then $C \times A^n \in \mathcal{C}_n$. Using also permutations of coordinates we see that if $D \in \mathcal{C}_n$, then $A^m \times D \in \mathcal{C}_{m+n}$, and thus by intersecting $C \times A^n$ and $A^m \times D$,

$$C \in \mathcal{C}_m,\ D \in \mathcal{C}_n \implies C \times D \in \mathcal{C}_{m+n}.$$

Also, if $1 \le i < j \le n$, then $\Delta^n_{i,j} := \{(x_1, \dots, x_n) :\ x_i = x_j\} \in \mathcal{C}_n$.

An example of a family $(\mathcal{C}_n)$ with the properties above is obtained by taking $\mathcal{C}_n$ to be the set of polynomial zerosets in $\mathbb{N}^n$, with $(A, \triangle) = (\mathbb{N}, \le)$.

Let $\mathcal{C} := \bigcup_n \mathcal{C}_n$, a subset of the disjoint union $\bigcup_n \mathcal{P}(A^n)$, and define $\Sigma(\mathcal{C})$ to be the smallest subset of $\bigcup_n \mathcal{P}(A^n)$ that includes $\mathcal{C}$, and is closed under projection and bounded universal quantification. (In the example above for $(A, \triangle) = (\mathbb{N}, \le)$ this gives $\Sigma(\mathcal{C}) = \Sigma$.) Below we show that the conditions imposed on $\mathcal{C}$ are inherited by $\Sigma(\mathcal{C})$. The main issue is to deal with the fact that bounded universal quantification does not commute with some coordinate reindexings given by maps $\lambda\colon \{1, \dots, m\} \to \{1, \dots, n\}$.

Lemma 51. *If $D \in \mathcal{C}$, and $S \in \Sigma(\mathcal{C})$, then $D \times S \in \Sigma(\mathcal{C})$.*

Proof. Let $D \in \mathcal{C}_d$. Define the subset Σ' of $\Sigma(\mathcal{C})$ to have as elements the $S \in \Sigma(\mathcal{C})$ such that $D \times S \in \Sigma(\mathcal{C})$. Note that $\mathcal{C} \subseteq \Sigma'$. Since projection and bounded universal quantification commute with the operation of taking the cartesian product with D as first factor, it follows that Σ' is closed under projection and bounded universal quantification. Thus $\Sigma' = \Sigma(\mathcal{C})$. □

Lemma 52. *If $S \in \Sigma(\mathcal{C})$, $S \subseteq A^m$, then:*

(1) *$C \cap S \in \Sigma(\mathcal{C})$ for all $C \in \mathcal{C}_m$;*

(2) *$\lambda(S) \in \Sigma(\mathcal{C})$ for all injective $\lambda : \{1, \dots, m\} \to \{1, \dots, n\}$.*

Proof. Define the subset Σ' of $\Sigma(\mathcal{C})$ to have as its elements the $S \subseteq A^m$, $m = 0, 1, 2 \dots$, such that $S \in \Sigma(\mathcal{C})$ and $D \cap \lambda(S) \in \Sigma(\mathcal{C})$ for all $D \in \mathcal{C}_n$ and injective $\lambda : \{1, \dots, m\} \to \{1, \dots, n\}$. Then $\mathcal{C} \subseteq \Sigma'$. Next we show that Σ' is closed under projection. Let $m \le n$ and $S \subseteq A^n$, $S \in \Sigma'$. To show that $\pi^n_m S \in \Sigma'$, let $\lambda : \{1, \dots, m\} \to \{1, \dots, k\}$ be injective and $D \in \mathcal{C}_k$. Then for $\vec{a} = (a_1, \dots, a_k) \in A^k$ and with $\vec{x} = (x_1, \dots, x_{n-m})$ ranging over A^{n-m},

$$\begin{aligned}
\big(D \cap \lambda(\pi^n_m S)\big)(\vec{a}) \;&\Leftrightarrow D(\vec{a}) \ \&\ (\pi^n_m S)(a_{\lambda(1)}, \dots, a_{\lambda(m)}) \\
&\Leftrightarrow D(\vec{a}) \ \&\ \exists \vec{x}\ S(a_{\lambda(1)}, \dots, a_{\lambda(m)}, \vec{x}) \\
&\Leftrightarrow \exists \vec{x}\big(D(\vec{a}) \ \&\ S(a_{\lambda(1)}, \dots, a_{\lambda(m)}, \vec{x})\big) \\
&\Leftrightarrow \exists \vec{x}\big(D(\vec{a}) \ \&\ (\mu S)(\vec{a}, \vec{x})\big),\ \text{where}
\end{aligned}$$

$$\mu : \{1, \dots, n\} \to \{1, \dots, k+n-m\} \text{ is given by}$$

$$\mu(j) = \lambda(j) \text{ for } 1 \le j \le m, \quad \mu(m+j) = k+j \text{ for } 1 \le j \le n-m,$$

Since $S \in \Sigma'$, the set $(D \times A^{n-m}) \cap \mu S$ belongs to $\Sigma(C)$, and so does

$$D \cap \lambda(\pi^n_m S) \;=\; \pi^{k+n-m}_k\big((D \times A^{n-m}) \cap \mu S\big).$$

Thus Σ' is closed under projection. It remains to show that Σ' is closed under bounded universal quantification, so let $R \subseteq A^{m+2}$, $R \in \Sigma'$. To get $\forall^{\vartriangle} R \in \Sigma'$, let $\lambda : \{1, \dots, m+1\} \to \{1, \dots, n\}$ be injective and $D \in \mathcal{C}_n$. Then for $S = D \cap \lambda(\forall^{\vartriangle} R)$ and $\vec{a} = (a_1, \dots, a_n)$,

$$\begin{aligned}
S(\vec{a}) \;&\Leftrightarrow D(\vec{a}) \;\&\; (\forall^{\vartriangle} R)(a_{\lambda(1)}, \dots, a_{\lambda(m+1)})\\
&\Leftrightarrow D(\vec{a}) \;\&\; \forall y_{\vartriangle a_{\lambda(m+1)}} R(a_{\lambda(1)}, \dots, a_{\lambda(m+1)}, y)\\
&\Leftrightarrow \exists x\big(D(\vec{a}) \;\&\; x = a_{\lambda(m+1)} \& \; \forall y_{\vartriangle x} R(a_{\lambda(1)}, \dots, a_{\lambda(m)}, x, y)\big)\\
&\Leftrightarrow \exists x \forall y_{\vartriangle x}\big(D(\vec{a}) \;\&\; x = a_{\lambda(m+1)} \;\&\; (\mu R)(\vec{a}, x, y)\big), \text{ where}\\
\mu \;&: \{1, \dots, m+2\} \to \{1, \dots, n+2\} \text{ is given by}\\
\mu(j) = \lambda(j) &\text{ for } 1 \le j \le m, \quad \mu(m+1) = n+1, \;\; \mu(m+2) = n+2.
\end{aligned}$$

Arguing as before this gives $S \in \Sigma(\mathcal{C})$, and so Σ' is closed under bounded universal quantification. □

Corollary 53. *If $S, T \in \Sigma(\mathcal{C})$, then $S \times T \in \Sigma(\mathcal{C})$. If $S, T \in \Sigma(\mathcal{C})$ and also $S, T \subseteq A^n$, then $S \cap T \in \Sigma(\mathcal{C})$.*

Proof. Let $S \subseteq A^n$ and $S \in \Sigma(\mathcal{C})$. Define Σ' to be the subset of $\Sigma(\mathcal{C})$ whose elements are the $T \in \Sigma(\mathcal{C})$ with $S \times T \in \Sigma(\mathcal{C})$. It follows from Lemma 51 and (2) of Lemma 52 that $\mathcal{C} \subseteq \Sigma'$. Also, projection and bounded universal quantification commute with taking cartesian products with S as first factor, so Σ' is closed under projection and bounded universal quantification. Hence $\Sigma' = \Sigma(\mathcal{C})$, which gives the first claim of the corollary. Now, with also $T \subseteq A^n$ and $T \in \Sigma(\mathcal{C})$, we have

$$S \cap T \;=\; \pi^{2n}_n\big(\Delta^{2n}_{1,n+1} \cap \dots \cap \Delta^{2n}_{n,2n} \cap (S \times T)\big),$$

and so $S \cap T \in \Sigma(\mathcal{C})$ by (1) of Lemma 52. □

We have now shown that $\Sigma(\mathcal{C})$ inherits the conditions that we imposed on $\mathcal{C}$. In particular, with $(A, \vartriangle) = (\mathbb{N}, \le)$, it follows that Σ is closed under cartesian products and intersections. We now return to the setting where variables like a, b, x, y range over $\mathbb{N}$ and give the proof of Proposition 50:

Proof. For $\subseteq$, note that the polynomial zero sets are primitive recursive, and hence recursively enumerable. Now use that the class of recursively enumerable sets is closed under taking projections and bounded universal quantifications.

For $\supseteq$ we use the familiar inductive definition of "recursive function" to show that these functions are in Σ, that is, their graphs are. It is clear that this holds for the initial functions (addition, multiplication, ...). Assume inductively that the graphs of $f : \mathbb{N}^n \to \mathbb{N}$ and $g_1, \ldots, g_n : \mathbb{N}^m \to \mathbb{N}$ belong to Σ. Then with $\vec{a} = (a_1, \ldots, a_m)$ and $\vec{x} = (x_1, \ldots, x_n)$ we have the equivalence

$$f(g_1, \ldots, g_n)(\vec{a}) = b \iff \exists \vec{x}\big(g_1(\vec{a}) = x_1 \ \& \ \ldots \& \ g_n(\vec{a}) = x_n \ \& \ f(\vec{x}) = b\big),$$

which shows that the graph of $f(g_1, \ldots, g_n)$ belongs to Σ. Next, assume that the graph of $g : \mathbb{N}^{n+1} \to \mathbb{N}$ is in Σ, that for all $\vec{x} \in \mathbb{N}^n$ there is y such that $g(\vec{x}, y) = 0$ and that $f : \mathbb{N}^n \to \mathbb{N}$ is given by $f(\vec{x}) = \mu y \ g(\vec{x}, y) = 0$. Then

$$f(\vec{x}) = y \iff g(\vec{x}, y) = 0 \ \& \ \forall i_{<y} \exists z \ g(\vec{x}, y) = z + 1,$$

so the graph of f belongs to Σ. It follows that every recursive set $S \subseteq \mathbb{N}^m$ is in Σ. It remains to use that recursively enumerable sets are projections of recursive sets. □

It follows from Proposition 50 that to obtain the Main Theorem, it suffices to show that if $D \subseteq \mathbb{N}^{m+2}$ is diophantine, so is $\forall^{\leq} D \subseteq \mathbb{N}^{m+1}$. Thus we have to eliminate a *bounded universal* quantifier in favour of existential quantifiers. We shall be able to do this, but at the cost of introducing non-polynomial functions like 2^x. This reduces our task to showing that these functions are diophantine.

In order to eliminate bounded universal quantifiers, it helps to introduce the functions qu, rem : $\mathbb{N}^2 \to \mathbb{N}$ (*quotient and remainder*), defined by

$$\begin{gathered} x = \mathrm{qu}(x, y) \cdot y + \mathrm{rem}(x, y), \text{ where} \\ \mathrm{rem}(x, y) < y \ \text{ if } y \neq 0, \quad \mathrm{rem}(x, 0) = \mathrm{qu}(x, 0) = x. \end{gathered}$$

It is easy to check that qu and rem are diophantine. We also need the functions $x \mapsto x! : \ \mathbb{N} \to \mathbb{N}$ and $(x, y) \mapsto \binom{x}{y} : \ \mathbb{N}^2 \to \mathbb{N}$, where

$$\begin{aligned} \binom{x}{y} &= \frac{x(x-1)\ldots(x-y+1)}{y!} \\ &= \begin{cases} \frac{x!}{y!(x-y)!} & \text{if } x \geq y, \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

Lemma 54. *If the function 2^x is diophantine, so are x^y, $\binom{x}{y}$ and $x!$.*

Proof. Trivially, $2^{xy} \equiv x \mod 2^{xy} - x$, so $(2^{xy})^y \equiv x^y \mod 2^{xy} - x$, that is,

$$2^{xy^2} \equiv x^y \mod 2^{xy} - x.$$

As $x^y < 2^{xy} - x$ for $y > 1$, we have $x^y = \mathrm{rem}(2^{xy^2}, 2^{xy} - x)$ for $y > 1$. This expression shows that if 2^x is diophantine, so is x^y.

For $\binom{x}{y}$, note that $2^x = (1+1)^x = \sum_{y=0}^{x} \binom{x}{y}$, so $\binom{x}{y} < 2^x$ for $x > 0$. Also, we have $(1+d)^x = \sum_{y=0}^{x} \binom{x}{y} d^y$, which gives the base d expansion of $(1+d)^x$ when $d \geq 2^x$. In other words, for all $x > 0, 0 \leq y \leq x$ and all n:

$$\binom{x}{y} = n \iff \exists c, d \left[d = 2^x, n < d, c < d^y, (1+d)^x \equiv c + nd^y \mod d^{y+1}\right].$$

This equivalence shows that if 2^x is diophantine, then $\binom{x}{y}$ is diophantine.

For $x!$, note that if $y > x > 1$ then

$$\begin{aligned}\binom{y}{x} &= \frac{y(y-1)\dots(y-x+1)}{x!} \\ &= \frac{y^x(1-\frac{1}{y})\dots(1-\frac{x-1}{y})}{x!}.\end{aligned}$$

An easy induction on n shows that for $0 \le \epsilon_1, \dots, \epsilon_n \le 1$ we have

$$\prod_{i=1}^{n}(1-\epsilon_i) \ge 1 - \sum_{i=1}^{n} \epsilon_i.$$

For $x > 1$ and $y \ge x^2$ this gives

$$\begin{aligned}x!\binom{y}{x} \ \le \ y^x &= x!\binom{y}{x}\frac{1}{(1-\frac{1}{y})\dots(1-\frac{x-1}{y})} \\ &\le x!\binom{y}{x}\left(1+\frac{x^2}{y}\right) \\ &= x!\binom{y}{x} + \frac{x^2x!}{y}\binom{y}{x}.\end{aligned}$$

Hence, if $\frac{x^2x!}{y} < 1$, then $x! = \mathrm{qu}(y^x, \binom{y}{x})$. But $\frac{x^2x!}{y} < 1$ for $y \ge 2x^{x+2}$, and thus $x! = \mathrm{qu}(y^x, \binom{y}{x})$ for $y = 2x^{x+2}$, $x > 1$. Therefore, $x!$ is diophantine if 2^x is. □

We are going to reduce the proof of the Main Theorem to showing that 2^x is a diophantine function. This reduction can be viewed as an elaborated version of Gödel's coding lemma, which we review here for the reader's convenience, since its proof is suggestive of what comes later.

Lemma 55. *Let $b_1, \dots, b_n \le B \in \mathbb{N}$. Then there is a unique $\beta \in \mathbb{N}$ such that*

$$\beta < \prod_{i=1}^{n}(1+i\cdot n!B) \quad \textit{and } b_i = \mathrm{rem}(\beta, 1+i\cdot n!B) \textit{ for } i=1,\dots,n.$$

Proof. The numbers $1+i\cdot n!\cdot B$ for $i=1,\dots,n$ are pairwise coprime: if p were a common prime factor of $1+i\cdot n!B$ and $1+j\cdot n!B$ with $1 \le i < j \le n$, then $p \mid (j-i)\cdot n!B$, so $p \mid n!$ or $p \mid B$, a contradiction.

Therefore, by the Chinese Remainder Theorem, there exists $\beta \in \mathbb{N}$ such that $\beta \equiv b_i \mod 1+i\cdot n!B$ for $i=1,\dots,n$ and such a β is uniquely determined modulo $\prod_{i=1}^{n}(1+i\cdot n!\cdot B)$.

Since $b_i < 1+i\cdot n!\cdot B$ for $i=1,\dots,n$, the unique $\beta < \prod_{i=1}^{n}(1+i\cdot n!\cdot B)$ with $\beta \equiv b_i \mod 1+i\cdot n!\cdot B$ for $i=1,\dots,n$ has the desired property. □

It will be convenient to fix some notation in the remainder of this section:

$$\begin{aligned}&\vec{a} := (a_1, \dots, a_m), \quad \vec{b} := (b_1, \dots, b_n), \quad \vec{v} := (v_1, \dots, v_n),\\ &\vec{v} \le y \ :\Longleftrightarrow \ v_1 \le y, \dots, v_n \le y.\end{aligned}$$

We showed that, for each m, the recursively enumerable sets in $\mathbb{N}^m$ are exactly the subsets of $\mathbb{N}^m$ that are in Σ. Hence, to obtain the Main Theorem, it suffices to prove: *If $D \subseteq \mathbb{N}^{m+2}$ is diophantine, then so is $\forall^{\le} D \subseteq \mathbb{N}^{m+1}$.*

We first make a small further reduction. Let the diophantine set $D \subseteq \mathbb{N}^{m+2}$ be given by the equivalence

$$D(\vec{a}, x, u) \iff \exists \vec{v}\ F(\vec{a}, x, u, \vec{v}) = 0$$

where $F(A, X, U, V) \in \mathbb{Z}[A, X, U, V]$, $A = (A_1, \dots, A_m)$, $V = (V_1, \dots, V_n)$ and where $A_1, \dots, A_m, X, U, V_1, \dots, V_n$ are distinct polynomial indeterminates. Then for all $\vec{a}, x$ we have:

$$\begin{aligned}(\forall^{\le} D)(\vec{a}, x) &\iff \forall u_{\le x}\ D(\vec{a}, x, u)\\ &\iff \forall u_{\le x}\ \exists \vec{v}\ F(\vec{a}, x, u, \vec{v}) = 0\\ &\iff \exists y\ \forall u_{\le x}\ \exists \vec{v} \le y\ F(\vec{a}, x, u, \vec{v}) = 0.\end{aligned}$$

Thus it remains to show that for any $F(A, X, U, V) \in \mathbb{Z}[A, X, U, V]$, the set

$$\{(\vec{a}, x, y) : \ \forall u_{\le x}\ \exists \vec{v}_{\le y}\ F(\vec{a}, x, u, \vec{v}) = 0\} \ \subseteq \ \mathbb{N}^{m+2}$$

is diophantine. The next result goes under the name of *Bounded Quantifier Theorem*. Besides the polynomial indeterminates in (A, X, U, V), we let Y be an extra indeterminate which we allow also to occur in F:

Theorem 56. *Suppose $F(A, X, Y, U, V)$ in $\mathbb{Z}[A, X, Y, U, V]$ and $G(A, X, Y)$ in $\mathbb{N}[A, X, Y]$ are polynomials such that for all $\vec{a}, x, y$ and all $u \le x$ and $\vec{v} \le y$,*

$$G(\vec{a}, x, y) > |F(\vec{a}, x, y, u, \vec{v})| + 2x + y + 2.$$

Then the following equivalence holds for all $\vec{a}, x, y$, with $g := G(\vec{a}, x, y)$:

$$\begin{gathered}\forall u_{\le x}\ \exists \vec{v}_{\le y}\ F(\vec{a}, x, y, u, \vec{v}) = 0\\ \Longleftrightarrow\\ \exists \vec{b} \left[\tbinom{b_1}{y+1} \equiv \cdots \equiv \tbinom{b_n}{y+1} \equiv F(\vec{a}, x, y, g! - 1, \vec{b}) \equiv 0 \mod \tbinom{g!-1}{x+1} \right]\end{gathered}$$

Before we start the proof, note that for any $F \in \mathbb{Z}[A, X, Y, U, V]$ we obtain $G(A, X, Y) \in \mathbb{N}[A, X, Y]$ as in the hypothesis of the lemma as follows: Let $F^* \in \mathbb{N}[A, X, Y, U, V]$ be obtained from F by replacing each coefficient with its absolute value and put

$$G(A, X, Y) := F^*(A, X, Y, X, Y, \dots, Y) + 2X + Y + 3.$$

Proof. We shall refer to the equivalence in the Bounded Quantifier Theorem as the *BQ-equivalence*. Let $\vec{a}$, x, and y be given and put $g = G(\vec{a}, x, y)$. Then

$$\begin{aligned}\binom{g!-1}{x+1} &= \frac{(g!-1)(g!-2)\dots(g!-x-1)}{1 \cdot 2 \cdot \dots \cdot (x+1)} \\ &= \frac{g!-1}{1} \cdot \frac{g!-2}{2} \cdot \dots \cdot \frac{g!-x-1}{x+1} \\ &= \left(\frac{g!}{1} - 1\right)\left(\frac{g!}{2} - 1\right)\dots\left(\frac{g!}{x+1} - 1\right).\end{aligned}$$

Since $g > x+2$, all $x+1$ factors in this last product are in $\mathbb{N}^{>1}$.

Claim 1. Each prime factor of $\binom{g!-1}{x+1}$ is greater than g.

This is because $g \geq 2x+2$, so every prime $p \leq g$ divides $\frac{g!}{u+1}$ for each $u \leq x$, so no prime $p \leq g$ divides $\binom{g!-1}{x+1}$.

Claim 2. The factors $\frac{g!}{1} - 1, \frac{g!}{2} - 1, \dots, \frac{g!}{x+1} - 1$ are pairwise coprime.

To see why, suppose p is a prime factor of $\frac{g!}{i} - 1$, and $\frac{g!}{j} - 1$, with $1 \leq i < j \leq x+1$. Then $p > g$, but $p \mid g! - i$, $p \mid g! - j$, so $p \mid j - i$, contradicting $j - i \leq g$.

For each $u \leq x$, take a prime factor p_u of $\frac{g!}{u+1} - 1$. Then

$$p_u > g, \quad \frac{g!}{u+1} - 1 \equiv 0 \mod p_u,$$

so $g! \equiv u+1 \mod p_u$, hence $g! - 1 \equiv u \mod p_u$. Thus for all $\vec{b}$ and all $u \leq x$:

$$F(\vec{a}, x, y, g!-1, \vec{b}) \equiv F(\vec{a}, x, y, u, \mathrm{rem}(b_1, p_u), \dots, \mathrm{rem}(b_n, p_u)) \mod p_u.$$

Now, assume that for each $u \leq x$ we have natural numbers $v_{u1}, v_{u2}, \dots, v_{un} \leq y$ such that

$$F(\vec{a}, x, y, u, v_{u1}, \dots, v_{un}) = 0.$$

For $i = 1, \dots, n$, the Chinese Remainder Theorem provides $b_i < \binom{g!-1}{x+1}$ such that $b_i \equiv v_{ui} \mod \frac{g!}{u+1} - 1$ for all $u \leq x$. In other words,

$$\frac{g!}{u+1} - 1 \mid b_i(b_i - 1) \cdots (b_i - y) \text{ for } i = 1, \dots, n \text{ and all } u \leq x.$$

Then by Claim 2, $\binom{g!-1}{x+1} \mid b_i(b_i-1)\dots(b_i-y)$. By Claim 1 and $g > y+1$, all prime factors of $\binom{g!-1}{x+1}$ are greater than $y+1$. So,

$$\binom{g!-1}{x+1} \mid \frac{b_i(b_i-1)\dots(b_i-y)}{(y+1)!} = \binom{b_i}{y+1}.$$

Hence,

$$\binom{b_1}{y+1} \equiv \dots \equiv \binom{b_n}{y+1} \equiv 0 \mod \binom{g!-1}{x+1}.$$

For $u \le x$, $\left(\frac{g!}{u+1}-1\right)(u+1) = g! - (u+1) = (g!-1) - u$, so

$$\begin{aligned} g!-1 &\equiv u \mod \frac{g!}{u+1}-1, \quad \text{hence} \\ F(\vec{a},x,y,g!-1,\vec{b}) &\equiv F(\vec{a},x,y,u,\vec{v}_u) \\ &\equiv 0 \mod \frac{g!}{u+1}-1. \end{aligned}$$

Thus, by the second claim, $F(\vec{a},x,y,g!-1,\vec{b}) \equiv 0 \mod \binom{g!-1}{x+1}$. This proves the forward direction of the BQ-equivalence.

For the converse, let $\vec{b}$ be such that

$$\binom{b_1}{y+1} \equiv \cdots \equiv \binom{b_n}{y+1} \equiv F(a,x,y,g!-1,\vec{b}) \equiv 0 \mod \binom{g!-1}{x+1}.$$

Then for $1 \le i \le n$, $u \le x$:

$$\binom{b_i}{y+1} \equiv 0 \mod p_u, \qquad \text{so} \qquad p_u \mid b_i(b_i-1)\dots(b_i-y),$$

hence $p_u | b_i - k$ for some $k \le y$, which gives $\operatorname{rem}(b_i,p_u) \le y$. Hence

$$|F(\vec{a},x,y,u,\operatorname{rem}(b_1,p_u),\dots,\operatorname{rem}(b_n,p_u))| < g < p_u \mid \binom{g!-1}{x+1},$$

and thus $F(\vec{a},x,y,u,\operatorname{rem}(b_1,p_u),\dots,\operatorname{rem}(b_n,p_u)) = 0$, as desired. □

Lemma 54 and The Bounded Quantifier Theorem reduce the proof of the Main Theorem to showing that the function 2^x is diophantine. This will be achieved by exploiting subtle properties of solutions of Pell equations. In the next section we derive the relevant facts about these equations. We finish this section with a digression on exponentially diophantine equations. Readers who want to take the shortest path to a proof of the Main Theorem can skip this.

Exponentially diophantine sets. Some results above are conditional on the function 2^x being diophantine. A nice way to eliminate the conditional nature of these results without proving outright that 2^x is diophantine is as follows. Call a set $E \subseteq \mathbb{N}^m$ *exponentially diophantine* if for some polynomial

$$\begin{gathered} F(A,X,Y) \in \mathbb{Z}[A,X,Y], \text{ with} \\ A=(A_1,\dots,A_m),\ X=(X_1,\dots,X_n),\ Y=(Y_1,\dots,Y_n), \end{gathered}$$

the following equivalence holds for all $\vec{a} \in \mathbb{N}^m$:

$$E(\vec{a}) \iff \exists \vec{x}\ F(\vec{a},\vec{x},2^{\vec{x}}) = 0,$$

where $2^{\vec{x}} := (2^{x_1},\dots,2^{x_n})$ for $\vec{x} = (x_1,\dots,x_n) \in \mathbb{N}^n$. Clearly, if $E \subseteq \mathbb{N}^m$ is diophantine, then E is exponentially diophantine. Note that Lemma 45 goes through with "exponentially diophantine" instead of "diophantine". So does Lemma 47, where in part (c) the ordered semiring $(\mathbb{N}; 0, 1, +, \cdot, \leq)$ is replaced by the ordered exponential semiring $(\mathbb{N}; 0, 1, +, \cdot, \exp, \leq)$ with $\exp(x) := 2^x$. A function $f : \mathbb{N}^m \to \mathbb{N}$ is said to be exponentially diophantine if its graph (as a subset of $\mathbb{N}^{m+1}$) is exponentially diophantine. Then Lemma 49 goes through with "exponentially diophantine" instead of "diophantine". The *proof* of Lemma 54 shows (unconditionally) that the functions x^y, $\binom{x}{y}$, and $x!$ are exponentially diophantine. We are now approaching a proof of the following unconditional result.

Theorem 57. *For each set* $S \subseteq \mathbb{N}^m$,

$$S \textit{ is recursively enumerable} \iff S \textit{ is exponentially diophantine.}$$

As stated, this is of course a consequence of the Main Theorem to be proved in the next section, but it seems worth deriving this exponential version using just the facts we have already established. Moreover, this exponential version holds in an enhanced form that we briefly touch on at the end of this digression. One might try to prove Theorem 57 by an exponential analogue of the BQ-Theorem. However, the proof of that theorem doesn't go through, since it can happen that $x \equiv y \mod m$, but $2^x \not\equiv 2^y \mod m$.

Instead we shall focus on improving Proposition 50, which says that any given recursively enumerable set can be obtained from a polynomial zero set by a finite sequence of projections and bounded universal quantifications. The main point is to show that bounded universal quantification is needed only once in such a description of a given recursively enumerable set. This is an old result due to Martin Davis, the *Davis Normal Form*:

Proposition 58. *Given any recursively enumerable set* $S \subseteq \mathbb{N}^d$, *there is a diophantine set* $D \subseteq \mathbb{N}^{d+m+2}$ *with the property that for all* $\vec{a} \in \mathbb{N}^d$,

$$S(\vec{a}) \iff \exists x_1 \dots \exists x_m \exists x \, \forall u_{\leq x} \, D(\vec{a}, x_1, \dots, x_m, x, u).$$

Proof. We first show how to interchange the order of bounded universal quantification and existential quantification. Let $S \subseteq \mathbb{N}^{m+3}$. Using the Chinese Remainder Theorem to code finite sequences as in the proof of Lemma 55 we have for all $(\vec{a}, b) \in \mathbb{N}^{m+1}$,

$$\begin{gathered} \forall u_{\leq b} \, \exists x \, S(\vec{a}, b, u, x) \\ \iff \\ \exists y, z \, \forall u_{\leq b} \, S\big(\vec{a}, b, u, \operatorname{rem}(y, 1 + (u+1)z)\big). \end{gathered}$$

Next, we show how to contract two successive bounded universal quantifications into a single bounded universal quantification, at the cost of an extra existential quantification. Let $L, R : \mathbb{N} \to \mathbb{N}$ be the functions such that $n = |L(n), R(n)|$ for all n. It is easy to check that L, R are diophantine. Let now $S \subseteq \mathbb{N}^{m+4}$.

Then for all $(\vec{a}, b, c) \in \mathbb{N}^{m+2}$:

$$\forall u_{\leq b}\ \forall v_{\leq c}\ S(\vec{a}, b, c, u, v)$$

$$\Longleftrightarrow$$

$$\forall x_{\leq |b,c|}\ \big[(L(x) \leq b\ \&\ R(x) \leq c)\ \Rightarrow\ S(\vec{a}, b, c, L(x), R(x))\big]$$

$$\Longleftrightarrow$$

$$\exists y\ \forall x_{\leq y}\big[y = |b,c|\ \&\ \big\{\big(L(x) \leq b\ \&\ R(x) \leq c\big)\ \Rightarrow\ S(\vec{a}, b, c, L(x), R(x))\big\}\big]$$

Now an inductive argument using Proposition 50 yields the desired result. □

The BQ-theorem in combination with Proposition 58 and the facts mentioned earlier on exponentially diophantine sets and functions show that any recursively enumerable set $S \subseteq \mathbb{N}^d$ is exponentially diophantine. This concludes the proof of Theorem 57.

2.3 Pell Equations

Below we assume that $d \in \mathbb{N}$ is not a square, that is,

$$d \notin \mathrm{sq}(\mathbb{N}) := \{n^2 : n = 0, 1, 2, \dots\} = \{0, 1, 4, 9, \dots\}.$$

So $2, 3, 5, 6, 7, 8, 10, \dots$ are the possible values of d. Then

$$X^2 - dY^2 = 1 \qquad\qquad \text{(P)}$$

is called a *Pell equation*. Note that (P) has the solution $(1, 0)$, which we call the *trivial* solution. Given any solution (a, b), we have the factorization

$$a^2 - db^2 = (a + b\sqrt{d})(a - b\sqrt{d}) = 1,$$

which yields $(a + b\sqrt{d})^n(a - b\sqrt{d})^n = 1$, and taking a', b' such that

$$(a + b\sqrt{d})^n = a' + b'\sqrt{d}, \quad (a - b\sqrt{d})^n = a' - b'\sqrt{d},$$

we see that (a', b') is also a solution of (P).

Example. Consider the case $d = 2$. Then $X^2 - 2Y^2 = 1$ has the non-trivial solution $(3, 2)$, and

$$(3 + 2\sqrt{2})^2 = 9 + 12\sqrt{2} + 8 = 17 + 12\sqrt{2},$$
$$(3 - 2\sqrt{2})^2 = 9 - 12\sqrt{2} + 8 = 17 - 12\sqrt{2},$$

yielding another solution $(17, 12)$.

A key fact in connection with $H10$ is that the solutions of (P) in $\mathbb{N}^2$ form a sequence $(x_n, y_n), n = 0, 1, 2, \dots$ where x_n and y_n grow roughly as an exponential function of n. Our short term goal is to establish this fact.

Lemma 59. *$G := \{x \pm y\sqrt{d} : x^2 - dy^2 = 1\}$ is a subgroup of the multiplicative group $\mathbb{R}^{>0}$ of positive real numbers.*

Proof. First, $1 = 1 + 0\sqrt{d}$, so $1 \in G$. We have $x^2 - dy^2 = (x + y\sqrt{d})(x - y\sqrt{d})$, so if $x^2 - dy^2 = 1$, then $x - y\sqrt{d} = \frac{1}{x+y\sqrt{d}}$. Also,

$$\begin{array}{rcl}(x_1 + y_1\sqrt{d})(x_2 + y_2\sqrt{d}) & = & (x_1x_2 + dy_1y_2) + (x_1y_2 + x_2y_1)\sqrt{d} \\ (x_1 - y_1\sqrt{d})(x_2 - y_2\sqrt{d}) & = & (x_1x_2 + dy_1y_2) - (x_1y_2 + x_2y_1)\sqrt{d}\end{array}$$

Hence, if $x_1^2 - dy_1^2 = x_2^2 - dy_2^2 = 1$, then $x^2 - dy^2 = 1$ where

$$x := x_1x_2 + dy_1y_2, \quad y := \pm(x_1y_2 + x_2y_1).$$

□

Note that the $x + y\sqrt{d} \in G$ are ≥ 1 and the $x - y\sqrt{d} \in G$ are ≤ 1.

Corollary 60. *Suppose* (P) *has a nontrivial solution in $\mathbb{N}^2$. Let (x_1, y_1) be such a solution with minimal $x_1 > 1$. Then the solutions in $\mathbb{N}^2$ of* (P) *are exactly the $(x_n, y_n) \in \mathbb{N}^2$ given by*

$$(x_1 + y_1\sqrt{d})^n = x_n + y_n\sqrt{d}, \qquad n = 0, 1, 2, \ldots$$

Proof. Put $r := x_1 + y_1\sqrt{d}$, so r is the least element of G which is greater than 1. We have $1 < r < r^2 < \ldots$ and $r^n \to \infty$ as $n \to \infty$. Let (x, y) be a solution of (P). Take the unique n such that $r^n \leq x + y\sqrt{d} < r^{n+1}$. Then $1 \leq (x + y\sqrt{d})r^{-n} < r$. By the previous lemma, $(x + y\sqrt{d})r^{-n} \in G$, so $(x + y\sqrt{d})r^{-n} = 1$, that is, $x + y\sqrt{d} = r^n$, so $x = x_n$, $y = y_n$. □

If there is a non-trivial solution of (P) in $\mathbb{N}^2$, then (x_1, y_1) as in Corollary 60 is called *the minimal* solution of (P). Note that in this corollary, $n = 0$ yields the trivial solution, and $n = 1$ the minimal solution.

We record two rules for solutions of (P), assuming there is a non-trivial solution of (P) in $\mathbb{N}^2$. First, the **addition rules**:

$$x_{m+n} = x_mx_n + dy_my_n, \quad y_{m+n} = x_my_n + x_ny_m,$$

which follow from the definitions of x_n and y_n, and the **doubling rules**, which follow from the addition rules:

$$x_{2n} = 2x_n^2 - 1, \quad y_{2n} = 2x_ny_n.$$

Our next goal is Lagrange's theorem that non-trivial solutions of (P) in $\mathbb{N}^2$ do exist. First a lemma.

Lemma 61. *Let $\xi \in \mathbb{R}^{>1}$ be irrational. Then there are infinitely many (x, y) with $y > 0$ such that $\left|\xi - \frac{x}{y}\right| < \frac{1}{y^2}$.*

The obvious bound would be $\frac{1}{y}$, so the lemma says that there are much better rational approximations to ξ, in terms of the size of the denominator. One can show that replacing y^2 in the lemma by y^r with a fixed real number $r > 2$ would result in a false statement for $\xi = \sqrt{2}$.

Proof. Let $n \geq 1$; we claim that there are $x, y > 0$ such that $y \leq n$ and

$$\left|\xi - \frac{x}{y}\right| < \frac{1}{ny} \leq \frac{1}{y^2}.$$

For $y = 0, 1, 2, \ldots, n$, we have $y\xi = x + \epsilon(y)$, where $x \in \mathbb{N}$ and $0 \leq \epsilon(y) < 1$. Divide the interval $[0, 1)$ into the n subintervals

$$[0, 1/n), [1/n, 2/n), \ldots, [(n-1)/n, 1).$$

By the pigeonhole principle, two of the $n+1$ numbers $\epsilon(0), \epsilon(1), \ldots, \epsilon(n)$ must lie in the same subinterval. Thus we have $y_1 < y_2$, both in $\{0, 1, \ldots, n\}$ such that $|\epsilon(y_1) - \epsilon(y_2)| < 1/n$. This gives $x_1 \leq x_2$ such that

$$y_1\xi = x_1 + \epsilon(y_1), \qquad y_2\xi = x_2 + \epsilon(y_2).$$

Now set $y = y_2 - y_1$ and $x = x_2 - x_1$. We get $y\xi = x + \epsilon(y_2) - \epsilon(y_1)$. So

$$|y\xi - x| < \frac{1}{n}, \text{ and thus } \left|\xi - \frac{x}{y}\right| < \frac{1}{ny}.$$

We have $x > 0$ because $\xi > 1$ and $\frac{1}{ny} < 1$. For each $n \geq 1$, take $x(n), y(n) \in \mathbb{N}^{>0}$ such that $y(n) \leq n$ and

$$\left|\xi - \frac{x(n)}{y(n)}\right| < \frac{1}{ny} \leq \frac{1}{y^2}.$$

As n goes to infinity, we have $\frac{x(n)}{y(n)} \to \xi$, so the set

$$\{(x(n), y(n)) : n = 1, 2, 3, \ldots\}$$

is infinite. □

Theorem 62. *The equation* (P) *has a nontrivial solution in* $\mathbb{N}^2$.

This theorem is due to Lagrange, but we give Dirichlet's proof.

Proof. If $x, y \in \mathbb{N}^{>0}$ and $\left|\sqrt{d} - x/y\right| < 1/y^2$, then

$$0 < \left|d - \frac{x^2}{y^2}\right| = \left|\sqrt{d} - \frac{x}{y}\right| \left|\sqrt{d} + \frac{x}{y}\right| < \frac{1}{y^2}\left(1 + 2\sqrt{d}\right),$$

so

$$0 < \left|x^2 - dy^2\right| < 1 + 2\sqrt{d}.$$

By the previous lemma we have therefore a nonzero integer λ with $|\lambda| < 1+2\sqrt{d}$, such that the equation

$$X^2 - dY^2 = \lambda, \qquad (\mathrm{P}_\lambda)$$

has infinitely many solutions (x, y) with $x, y \in \mathbb{N}^{>0}$. Consider two such solutions of (P_λ), (x_1, y_1) and (x_2, y_2). Then:

$$\begin{array}{rcccl} x_1^2 - dy_1^2 & = & (x_1 + y_1\sqrt{d})(x_1 - y_1\sqrt{d}) & = & \lambda, \\ x_2^2 - dy_2^2 & = & (x_2 + y_2\sqrt{d})(x_2 - y_2\sqrt{d}) & = & \lambda, \end{array}$$

so

$$\left(\frac{x_1 + y_1\sqrt{d}}{x_2 + y_2\sqrt{d}}\right)\left(\frac{x_1 - y_1\sqrt{d}}{x_2 - y_2\sqrt{d}}\right) = 1.$$

Now we remove square roots from denominators:

$$\begin{aligned} \frac{x_1 + y_1\sqrt{d}}{x_2 + y_2\sqrt{d}} &= \frac{(x_1 + y_1\sqrt{d})(x_2 - y_2\sqrt{d})}{(x_2 + y_2\sqrt{d})(x_2 - y_2\sqrt{d})} = \frac{(x_1x_2 - dy_1y_2) + (x_2y_1 - x_1y_2)\sqrt{d}}{\lambda}, \\ \frac{x_1 - y_1\sqrt{d}}{x_2 - y_2\sqrt{d}} &= \frac{(x_1 - y_1\sqrt{d})(x_2 + y_2\sqrt{d})}{(x_2 - y_2\sqrt{d})(x_2 + y_2\sqrt{d})} = \frac{(x_1x_2 - dy_1y_2) - (x_2y_1 - x_1y_2)\sqrt{d}}{\lambda}. \end{aligned}$$

Thus, setting

$$s := \left|\frac{x_1x_2 - dy_1y_2}{\lambda}\right|, \quad t := \left|\frac{x_2y_1 - x_1y_2}{\lambda}\right|,$$

we get $s^2 - dt^2 = 1$. The problem is that s and t might not be integers. However, (P_λ) has infinitely many solutions (x, y), so by the pigeonhole principle we can choose (x_1, y_1) and (x_2, y_2) as above so that $x_1 \equiv x_2 \mod \lambda$ and $y_1 \equiv y_2 \mod \lambda$, and $(x_1, y_1) \neq (x_2, y_2)$. Then

$$x_1x_2 - dy_1y_2 \equiv x_1^2 - dy_1^2 \equiv 0 \mod \lambda,$$

and $x_2y_1 - x_1y_2 \equiv 0 \mod \lambda$, so $s, t \in \mathbb{N}$.

It remains to show that (s, t) is not the trivial solution to (P). Suppose towards a contradiction that $s = 1$ and $t = 0$. Then $x_2y_1 - x_1y_2 = 0$, so $x_2/x_1 = y_2/y_1$. We can assume $x_2 > x_1$; set $r = x_2/x_1 > 1$. Then

$$x_2^2 - dy_2^2 = r^2(x_1^2 - dy_1^2) = r^2\lambda \neq \lambda,$$

a contradiction. So (s, t) is a nontrivial solution of (P) in $\mathbb{N}^2$. □

Although this proof appears non-constructive, one can actually derive from it an explicit upper bound on the minimal solution: our use of the pigeonhole principle does not really need infinitely many solutions to (P_λ), but only more than λ^2, and we have $|\lambda| < 1 + 2\sqrt{d}$.

The minimal solution of a Pell equation $X^2 - dY^2 = 0$ varies rather wildly with d. For our purpose we can restrict attention to a situation where the minimal solution is obvious: let $a \geq 2$ and consider a Pell equation

$$X^2 - (a^2-1)Y^2 = 1, \qquad\qquad (\mathrm{P}_a).$$

Its minimal solution is $(a, 1)$. Now define $x_a(n), y_a(n) \in \mathbb{N}$ by

$$x_a(n) + y_a(n)\sqrt{a^2-1} = (a + \sqrt{a^2-1})^n.$$

By previous results, $\{(x_a(n), y_a(n)) : n = 0, 1, 2, \ldots\}$ is the set of all solutions in $\mathbb{N}^2$. We now list several rules, most of which follow trivially from the earlier addition and doubling rules. First of all, we have the **addition rules**:

$$\begin{aligned} x_a(m+n) &= x_a(m)x_a(n) + (a^2-1)y_a(m)y_a(n), \\ y_a(m+n) &= x_a(m)y_a(n) + x_a(n)y_a(m), \end{aligned}$$

next, the **doubling formulae**:

$$\begin{aligned} x_a(2n) &= 2x_a(n)^2 - 1, \\ y_a(2n) &= 2x_a(n)y_a(n), \end{aligned}$$

the **simultaneous recursion equations**:

$$\begin{array}{rclrcl} x_a(0) &=& 1, & x_a(n+1) &=& ax_a(n) + (a^2-1)y_a(n), \\ y_a(0) &=& 0, & y_a(n+1) &=& x_a(n) + ay_a(n), \end{array}$$

and the **course-of-values** recursion:

$$\begin{array}{rclrclrcl} x_a(0) &=& 1, & x_a(1) &=& a, & x_a(n+2) &=& 2ax_a(n+1) - x_a(n), \\ y_a(0) &=& 0, & y_a(1) &=& 1, & y_a(n+2) &=& 2ay_a(n+1) - y_a(n). \end{array}$$

One can prove the course-of-values recursion, for example, as follows:

$$\begin{aligned} x_a(n+2) &= ax_a(n+1) + (a^2-1)y_a(n+1) \\ &= ax_a(n+1) + (a^2-1)\left[x_a(n) + ay_a(n)\right] \\ &= ax_a(n+1) + (a^2-1)x_a(n) + a\left[x_a(n+1) - ax_a(n)\right] \\ &= 2ax_a(n+1) - x_a(n). \end{aligned}$$

In addition we have a result about **growth**:

$$\begin{array}{rclcll} (2a-1)^n &\leq& y_a(n+1) &<& (2a)^n & \\ 2n &\leq& y_a(n) & & & \text{for } n \geq 2 \end{array}$$

Exercise. Prove the growth result.

Finally, we state some **congruence rules**: for $a, b \geq 2$ we have

$$\begin{aligned} y_a(n) &\equiv n \mod (a-1) \\ y_a(n) &\equiv y_b(n) \mod (a-b) \end{aligned}$$

Proof. For the first result, we have $y_a(0) = 0 \equiv 0 \mod (a-1)$ and $y_a(1) = 1 \equiv 1 \mod (a-1)$. Thus we can apply induction:

$$\begin{aligned} y_a(n+2) &\equiv 2y_a(n+1) - y_a(n) \mod (a-1) \\ &\equiv 2(n+1) - n \\ &\equiv n+2 \mod (a-1). \end{aligned}$$

For the second result, we note that it holds for $n = 0, 1$, and also that $a \equiv b \mod (a-b)$. Thus,

$$y_a(n+2) \equiv 2by_a(n+1) - y_a(n) \mod (a-b),$$

which shows that $y_a(n)$ satisfies the same recursion as $y_b(n)$ modulo $(a-b)$. □

Lemma 63 (Periodicity)**.** *Suppose $m, n > 0$ and $y_a(n) \equiv 0 \mod m$. Then for all k, l, if $k \equiv l \mod 2n$, then $y_a(k) \equiv y_a(l) \mod m$.*

Thus the remainder $\mathrm{rem}(y_a(k), m)$ is periodic as a function of k, with period dividing $2n$, where the period is the least $p \in \mathbb{N}^{>0}$ such that

$$\mathrm{rem}(y_a(k), m) = \mathrm{rem}(y_a(k+lp), m) \quad \text{for all } k, l.$$

Proof. Just observe the following equalities and congruences:

$$\begin{aligned} y_a(k+2n) &= x_a(k)y_a(2n) + x_a(2n)y_a(k) \\ &= 2x_a(k)x_a(n)y_a(n) + (2x_a(n)^2 - 1)y_a(k) \\ &\equiv (2x_a(n)^2 - 1)y_a(k) \mod m \\ &\equiv \big[2\big((a^2-1)y_a(n)^2 + 1\big) - 1\big]y_a(k) \mod m \\ &\equiv y_a(k) \mod m. \end{aligned}$$

□

We note also that for any $m > 0$, there is an $n > 0$ so that $y_a(n) \equiv 0 \mod m$. This is because for $d := m^2(a^2-1)$ the Pell equation $X^2 - dY^2 = 1$ has a non-trivial solution, say (x, y). Then (x, my) is a nontrivial solution to (P_a).

Next we prove some "step-down" lemmas. Roughly speaking, these provide more precise information about the period in the sequence of remainders $\mathrm{rem}(y_a(k), m)$ when m has a special form. In addition, they provide information about n from information about $y_a(n)$.

Lemma 64 (First Step-Down Lemma)**.** *Let $m, n > 0$. Then:*

$$\begin{aligned} y_a(m) | y_a(n) &\iff m | n, \\ y_a(m)^2 | y_a(n) &\iff m y_a(m) | n. \end{aligned}$$

Proof. Assume first that $m|n$, so $n = mk$, $k > 0$. Then

$$\begin{aligned} x_a(n) + y_a(n)\sqrt{a^2-1} &= (a+\sqrt{a^2-1})^{mk} \\ &= \left(\left(a+\sqrt{a^2-1}\right)^m\right)^k \\ &= \left(x_a(m) + y_a(m)\sqrt{a^2-1}\right)^k . \end{aligned}$$

Expanding the right hand side by the binomial theorem and comparing coefficients of $\sqrt{a^2-1}$, we get

$$(*) \qquad y_a(n) = kx_a(m)^{k-1}y_a(m) + y_a(m)^3 \cdot \text{(integer)}.$$

In particular, (*) yields that $y_a(m)$ divides $y_a(n)$, proving the backwards direction of the implication.

Conversely, assume that $y_a(m)|y_a(n)$. Towards a contradiction, let $m \nmid n$. Then $n = qm + r$ with $q, r \in \mathbb{N}$ and $0 < r < m$. By the addition formula,

$$y_a(n) = y_a(qm+r) = x_a(qm)y_a(r) + x_a(r)y_a(qm).$$

From $y_a(m)|y_a(qm)$, we obtain $y_a(m)|x_a(qm)y_a(r)$. But $y_a(qm)$ and $x_a(qm)$ are coprime, so $y_a(m)$ and $x_a(qm)$ are coprime. Hence $y_a(m)|y_a(r)$, which contradicts $0 < y_a(r) < y_a(m)$. This proves the first equivalence.

For the second, first note that both sides of the equivalence imply $m|n$, so we can assume $n = mk$, with $k \in \mathbb{N}^{>0}$. Then we again apply (*):

$$\begin{aligned} y_a(m)^2|y_a(n) &\iff y_a(m)^2|kx_a(m)^{k-1}y_a(m) \\ &\iff y_a(m)|kx_a(m)^{k-1} \\ &\iff y_a(m)|k \\ &\iff my_a(m)|n. \end{aligned}$$

□

Lemma 65 (Second Step-Down Lemma)**.** *Let $n \geq 1$ and suppose that $y_a(k) \equiv y_a(l) \mod x_a(n)$. Then $k \equiv \pm l \mod 2n$.*

Proof. Let $i, j \in \mathbb{N}$. Then the following congruences hold:

$$\begin{array}{lclll} y_a(2n+i) & \equiv & -y_a(i) & \mod x_a(n) & \\ y_a(2n-i) & \equiv & y_a(i) & \mod x_a(n) & \text{for} \quad 0 \leq i \leq n \\ y_a(4n+i) & \equiv & y_a(i) & \mod x_a(n) & \\ y_a(4n-i) & \equiv & -y_a(i) & \mod x_a(n) & \text{for} \quad 0 \leq i \leq n \end{array}$$

We prove the first of these congruences, leave the second as an exercise, and note that the third and fourth follow easily from applications of the first and second. The proof of the first congruence is as follows.

$$\begin{aligned} y_a(2n+i) &= x_a(2n)y_a(i) + x_a(i)y_a(2n) \\ &\equiv x_a(2n)y_a(i) \mod x_a(n) \\ &\equiv (2x_a(n)^2 - 1)y_a(i) \mod x_a(n) \\ &\equiv -y_a(i) \mod x_a(n) \end{aligned}$$

The second line here uses $x_a(n)|y_a(2n)$, which follows from the doubling formula. Next we have two incongruences to the effect that the congruences above are in some sense best possible:

$$\begin{aligned} y_a(i) &\not\equiv \pm y_a(j) \mod x_a(n) \quad \text{for } 0 \le i < j \le n, \\ y_a(i) &\not\equiv -y_a(i) \mod x_a(n) \quad \text{for } 0 < i \le n. \end{aligned}$$

To prove these incongruences we assume $a > 2$, and leave the somewhat special case $a = 2$ to the reader. Then

$$x_a(n) = \sqrt{(a^2-1)y_a(n)^2+1} > y_a(n)\sqrt{a^2-1} > 2y_a(n)$$

The first incongruence now follows because for $0 \le i < j \le n$,

$$x_a(n) > 2y_a(n) \ge y_a(j) \pm y_a(i).$$

The second incongruence holds because if $0 < i \le n$, then

$$x_a(n) > 2y_a(n) \ge 2y_a(i) > 0.$$

Now put $i = \text{rem}(k, 4n)$, $j = \text{rem}(l, 4n)$. Then

$$y_a(i) \equiv y_a(j) \mod x_a(n).$$

If $i = j$ then $k \equiv l \mod 4n$, so $k \equiv l \mod 2n$. If $i \ne j$, then a close look at the four initial congruences and the two incongruences above yields $i \equiv -j \mod 2n$, so $k \equiv -l \mod 2n$. □

2.4 Proof of the Main Theorem

We are now ready to show that $\{(a, n, y) : a \ge 2, y = y_a(n)\} \subseteq \mathbb{N}^3$ is diophantine. In the next result and its proof, all variables range over $\mathbb{N}$.

Theorem 66. *Let $a, y, n \ge 2$. Then the following are equivalent*

- *$y = y_a(n)$;*
- *there are x, u, v, s, t, b such that*

(i) $2n \le y$

(ii) $x^2 - (a^2-1)y^2 = 1$

(iii) $v \ge 1$ & $u^2 - (a^2-1)v^2 = 1$

(iv) $b \ge 2$ & $s^2 - (b^2-1)t^2 = 1$

(v) $b \equiv a \mod u$

(vi) $b \equiv 1 \mod y$

(vii) $t \equiv y \mod u$

(viii) $t \equiv n \mod y$

(ix) $y^2|v$.

Proof. Let x, u, v, s, t, b be as in (i)–(ix). Then take k, l, m such that

$$y = y_a(k), \qquad u = x_a(l), \qquad t = y_b(m).$$

We shall prove that $k = n$, so $y = y_a(n)$. By (iii), $l \geq 1$. By the congruence rules we have $y_b(m) \equiv y_a(m) \mod b - a$, and so by (v),

$$y_b(m) \equiv y_a(m) \mod x_a(l).$$

By (vii) we have $y_b(m) \equiv y_a(k) \mod x_a(l)$, so

$$y_a(m) \equiv y_a(k) \mod x_a(l)$$

By the second step down lemma,

$$m \equiv \pm k \mod 2l$$

By (ix), $y_a(k)^2 | y_a(l)$, so by the first step down lemma, $ky_a(k)|l$, so $y|l$ and thus

$$m \equiv \pm k \mod y \tag{2.2}$$

By the congruence rules, $y_b(m) \equiv m \mod (b-1)$, so

$$\begin{aligned} y_b(m) &\equiv m \mod y && \text{by (vi)}, \\ y_b(m) &\equiv n \mod y && \text{by (viii)} \\ n &\equiv m \mod y && \text{and thus by (2.2)}, \\ n &\equiv \pm k \mod y && \end{aligned} \tag{2.3}$$

By (i) and an earlier result on the bounds on solutions of P_a we have $2n \leq y$, $2k \leq y$, which in view of (2.3) yields $n = k$, so $y = y_a(n)$, as desired.

For the converse, assume $y_a(n) = y$. Then (ii) holds with $x = x_a(n)$. Take

$$l := ny_a(n), \qquad u := x_a(l), \qquad v := y_a(l).$$

Then (iii) holds. By the first step down lemma,

$$y^2 = (y_a(n))^2 | y_a(l) = v \geq 1$$

and (ix) holds. By (iii), u, v are coprime, and $y|v$, so u, y are coprime. By the Chinese Remainder Theorem we get $b \geq 2$ such that

$$\begin{aligned} b &\equiv a \mod u \\ b &\equiv 1 \mod y \end{aligned}$$

Then (v) and (vi) hold. Now take $s := x_b(n)$, $t := y_b(n)$. Then (iv) holds, and the congruence rules yield (vii) and (viii). □

Corollary 67. *The function $x \mapsto 2^x$ is diophantine.*

Proof. Let $a \geq 2$. Then

$$(2a-1)^n \leq y_a(n+1) \leq (2a)^n, \qquad (4a-1)^n \leq y_{2a}(n+1) \leq (4a)^n.$$

Therefore

$$2^n\left(1-\frac{1}{4a}\right)^n = \frac{(4a-1)^n}{(2a)^n} \leq \frac{y_{2a}(n+1)}{y_a(n+1)} \leq \frac{(4a)^n}{(2a-1)^n} = 2^n\left(1+\frac{1}{2a-1}\right)^n.$$

For fixed n, with $a \to \infty$,

$$\left(1+\frac{1}{2a-1}\right)^n \to 1, \qquad \left(1-\frac{1}{4a}\right)^n \to 1.$$

So for some $B(n) \in \mathbb{N}^{>0}$, whenever $a > B(n)$, then

$$\left|\frac{y_{2a}(n+1)}{y_a(n+1)} - 2^n\right| < \frac{1}{2},$$

hence for all $a > B(n)$ and all m,

$$2^n = m \iff 2\left|y_{2a}(n+1) - my_a(n+1)\right| < y_a(n+1).$$

So it only remains to show that there is a *diophantine* function $B : \mathbb{N} \to \mathbb{N}^{>0}$ such that for all n and all $a > B(n)$,

$$\left|\frac{y_{2a}(n+1)}{y_a(n+1)} - 2^n\right| < \frac{1}{2}.$$

We have

$$2^n\left(1-\frac{1}{4a}\right)^n \geq 2^n\left(1-\frac{n}{4a}\right) = 2^n - \frac{n2^n}{4a},$$

so if $4a > 2n2^n$, then $2^n\left(1-\frac{1}{4a}\right)^n > 2^n - \frac{1}{2}$. Also,

$$2^n\left(1+\frac{1}{2a-1}\right)^n \leq 2^n\left(1+\frac{n2^n}{2a-1}\right) = 2^n + \frac{n2^{2n}}{2a-1},$$

so if $2a-1 > 2n2^{2n}$, then $2^n\left(1+\frac{1}{2a-1}\right)^n < 2^n + \frac{1}{2}$. As $5^n \leq y_3(n+1) \leq 6^n$, we can take $B(n) := 2ny_3(n+1)+1$. Then $B(n) > 2n2^{2n}$, so B is a diophantine function as desired. □

The main theorem has now been established. It gives much more than a negative solution to H10.

Exercise. Show that there is no algorithm determining for any system

$$P_1(X_1,\ldots,X_n) = \cdots = P_m(X_1,\ldots,X_n) = 0$$

with all $P_i \in \mathbb{Z}[X_1, \dots, X_n]$ of degree ≤ 2, whether the system has a solution in $\mathbb{Z}^n$. Using this, show that there is no algorithm determining for any given equation

$$P(X_1, \dots, X_n) = 0$$

with $P \in \mathbb{Z}[X_1, \dots, X_n]$ of degree ≤ 4 whether the equation has a solution in $\mathbb{Z}^n$.

The answer to H10 is not even known for polynomial equations in two variables. There is, however, an algorithm to decide for any given $P \in \mathbb{Z}[X, Y]$ whether $P(X, Y) = 0$ has *infinitely many* integer solutions. This follows from a deep theorem due to C.L. Siegel (1929), with some extra work. There is also an algorithm based on other work by Siegel to decide whether a given polynomial equation $P(X_1, \dots, X_n) = 0$ with $P \in \mathbb{Z}[X_1, \dots, X_n]$ of degree 2, has an integer solution.

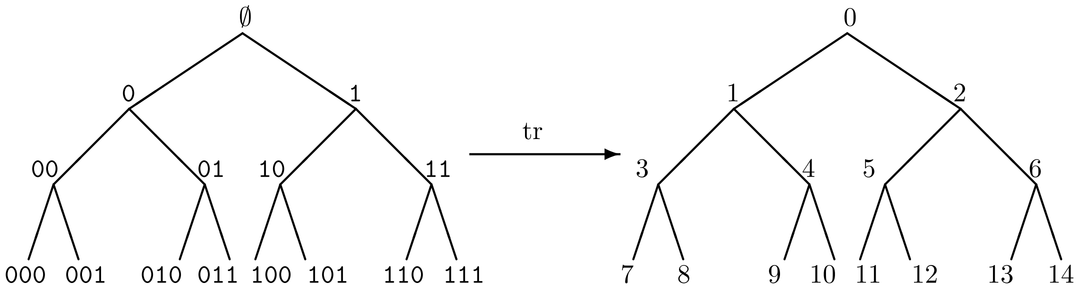
It is not known whether H10 has a positive answer when we allow *rational* solutions instead of integer solutions. On the other hand, H10 does have a positive answer when we allow real solutions, or p-adic solutions; this follows from the stronger result that the elementary theory of the real field is decidable (Tarski), and likewise for the elementary theory of each p-adic field (Ax–Kochen & Ersov).

Chapter 3

Kolmogorov Complexity

This is a theory about randomness using concepts of recursion theory. The idea is that a long finite string of 0's and 1's is random if any program to compute the string is about as long as the string itself.

Instead of natural numbers it is more convenient to work with finite strings of 0's and 1's, that is, elements of $2^{<\mathbb{N}}$, as the inputs and outputs of algorithms. We let $x, y, z, \pi, \rho, \sigma, \tau$ range over $2^{<\mathbb{N}}$. For $\sigma = (\sigma_0, \dots, \sigma_{n-1})$ we defined its length $|\sigma| = n$; the concatenation of x, y is denoted xy. The *tree coding* of $2^{<\mathbb{N}}$ is the bijection $\mathrm{tr} : 2^{<\mathbb{N}} \to \mathbb{N}$ indicated in the picture below:



In words, tr is the unique bijection $2^{<\mathbb{N}} \to \mathbb{N}$ such that $\mathrm{tr}(\sigma) < \mathrm{tr}(\tau)$ if $|\sigma| < |\tau|$, or $|\sigma| = |\tau|$ and σ precedes τ lexicographically. Via this bijection the notion of partial recursive function $\mathbb{N}^d \rightharpoonup \mathbb{N}$ can be transferred to a notion of partial recursive function $(2^{<\mathbb{N}})^d \rightharpoonup 2^{<\mathbb{N}}$. Precisely, a partial function $f : (2^{<\mathbb{N}})^d \rightharpoonup 2^{<\mathbb{N}}$ is declared to be partial recursive if the partial function $g : \mathbb{N}^d \rightharpoonup \mathbb{N}$ such that

$$g(\mathrm{tr}(x_1), \dots, \mathrm{tr}(x_d)) \simeq \mathrm{tr}(f(x_1, \dots x_d))$$

is partial recursive. Note: $2^{|\sigma|} - 1 \ \leq \ \mathrm{tr}(\sigma) \ \leq \ 2^{|\sigma|+1} - 2$.

3.1 Plain Kolmogorov Complexity

Let $f : 2^{<\mathbb{N}} \rightharpoonup 2^{<\mathbb{N}}$ be a partial recursive function. The (plain) Kolmogorov complexity of σ with respect to f is

$$C_f(\sigma) = \begin{cases} \min\{|\rho| : f(\rho) = \sigma\} & \text{if there is } \rho \in Df \text{ such that } f(\rho) = \sigma, \\ \infty & \text{otherwise.} \end{cases}$$

We say that σ is Kolmogorov random with respect to f if $C_f(\sigma) \geq |\sigma|$.

Let $U : 2^{<\mathbb{N}} \times 2^{<\mathbb{N}} \rightharpoonup 2^{<\mathbb{N}}$ be a universal partial recursive function. That is, U is partial recursive and

$$\{U(\pi, \cdot) : \pi \in 2^{<\mathbb{N}}\} = \{f : 2^{<\mathbb{N}} \rightharpoonup 2^{<\mathbb{N}} : \ f \text{ is partial recursive}\}.$$

Define the partial recursive function $u : 2^{<\mathbb{N}} \rightharpoonup 2^{<\mathbb{N}}$ by

$$\begin{aligned} u(0^{|\pi|}1\pi\rho) &\simeq U(\pi, \rho), \\ u(\sigma) \uparrow & \text{ if } \sigma \text{ does not have the form } 0^{|\pi|}1\pi\rho. \end{aligned}$$

Then $1 \leq C_u(\sigma) < \infty$ for all σ: the left inequality holds because $u(\emptyset) \uparrow$, and the right inequality holds because $U(\pi, \cdot) = \mathrm{id}_{2^{<\mathbb{N}}}$ for some π. This complexity function C_u is as good as any, up to an additive constant:

Lemma 68. *For all partial recursive $f : 2^{<\mathbb{N}} \rightharpoonup 2^{<\mathbb{N}}$ and all σ,*

$$C_u(\sigma) \leq C_f(\sigma) + \mathrm{const}(f).$$

Here and later $\mathrm{const}(f)$ denotes a constant in $\mathbb{N}$ depending only on f. Likewise, const denotes an absolute constant in $\mathbb{N}$. (Of course, the possible values of const may vary with the context.)

Proof. Let π be such that $f = U(\pi, \cdot)$. If $\rho \in Df$ and $f(\rho) = \sigma$, then $u(0^{|\pi|}1\pi\rho) = \sigma$ and $|0^{|\pi|}1\pi\rho| = |\rho| + 2|\pi| + 1$. Thus for all σ:

$$C_u(\sigma) \leq C_f(\sigma) + 2|\pi| + 1.$$

□

Definition 69. $C := C_u : 2^{<\mathbb{N}} \to \mathbb{N}$.

This complexity function C depends on the initial choice of U, but only in an inessential way: Let $U' : \ 2^{<\mathbb{N}} \times 2^{<\mathbb{N}} \rightharpoonup 2^{<\mathbb{N}}$ also be a universal partial recursive function, with $C' : 2^{<\mathbb{N}} \to \mathbb{N}$ as its associated complexity function. Then there is a constant $c \in \mathbb{N}$ such that $|C(\sigma) - C'(\sigma)| \leq c$ for all σ. Our interest is in the behaviour of C on sufficiently long strings, and we regard C and C' as equivalent for this purpose.

Lemma 70. *The following hold:*

(a) $C(x) \leq |x| + \mathrm{const}$;

(b) $C(xx) \le |x| + \text{const}$;

(c) if $h : 2^{<\mathbb{N}} \to 2^{<\mathbb{N}}$ *is recursive, then for all* x

$$C(h(x)) \le C(x) + \text{const}(h).$$

Proof. For $f = \text{id}_{2^{<\mathbb{N}}}$ we have $C_f(x) = |x|$. So $C(x) \le |x| + \text{const}$. This gives (a). Let $h : 2^{<\mathbb{N}} \to 2^{<\mathbb{N}}$ be recursive, and put $f := h \circ u$. Then

$$\begin{array}{rcl} C(h(x)) & \le & C_f(h(x)) + \text{const}(f) \\ & \le & C_u(x) + \text{const}(f) \\ & = & C(x) + \text{const}(h). \end{array}$$

This yields (c), and (b) is a special case of (c). □

Call x *random* if $C(x) \ge |x|$. For each n, there is a random x with $|x| = n$. This follows from the *pigeon hole principle*:

$$|\{\tau : |\tau| < n\}| = 2^0 + 2^1 + \cdots + 2^{n-1} = 2^n - 1,$$

and $|\{\tau : |\tau| = n\}| = 2^n$, so the partial map

$$\sigma \mapsto u(\sigma) : \{\sigma : |\sigma| < n\} \;\rightharpoonup\; \{x : |x| = n\}$$

cannot be surjective. More precisely, we have the following result.

Lemma 71. *For all* n *and all* $k \in \{0, \ldots, n\}$,

$$|\{x : \;|x| = n,\; C(x) \ge n - k\}| \;>\; 2^n(1 - \frac{1}{2^k}).$$

Proof. We have $|\{\tau : |\tau| < n - k\}| = 2^0 + 2^1 + \cdots + 2^{n-k-1} = 2^{n-k} - 1$, so $|\{x : |x| = n,\; C(x) \ge n - k\}| \;\ge\; 2^n - (2^{n-k} - 1) \;>\; 2^n(1 - \frac{1}{2^k})$. □

Thus $C(x) \ge 90$ for more than 99.9% of the strings x of length 100.

Lemma 72 (Martin-Löf)**.** *Let* $k \in \mathbb{N}$ *be given. Then any sufficiently long* z *has an initial segment* x *with* $C(x) < |x| - k$.

Proof. Let $f : 2^{<\mathbb{N}} \to 2^{<\mathbb{N}}$ be the recursive function satisfying $f(\sigma) = \rho\sigma$, where $\text{tr}(\rho) = |\sigma|$. Take $\Delta \in \mathbb{N}$ such that $C < C_f + \Delta$. Let z be a string satisfying $|z| > 2^{k+\Delta+2}$, and let ρ be the initial segment of z with $|\rho| = k + \Delta$. Let $r = \text{tr}(\rho)$; then $2^{|\rho|} - 1 \le r < 2^{|\rho|+1} - 1$, hence $|\rho| + r < k + \Delta + 2^{k+\Delta+1} - 1 \le |z|$.

Take $x = z|_{|\rho|+r}$ (the initial segment of z of length $|\rho| + r$). Then $x = \rho\sigma$, with $|\sigma| = r$, so $f(\sigma) = \rho\sigma = x$. Hence,

$$\begin{aligned} C(x) < C_f(x) + \Delta \le |\sigma| + \Delta &= (|\rho| - (\Delta + k)) + |\sigma| + \Delta \\ &= |\rho| + |\sigma| - k = |x| - k. \end{aligned}$$

□

Corollary 73. *For all sufficiently long random z, there are x, y such that $z = xy$, but $C(x) + C(y) < C(z)$.*

Proof. Take $k \in \mathbb{N}$ such that for all σ, $C(\sigma) \leq |\sigma| + k$. Let z be random and so long that it has an initial segment x with $C(x) < |x| - k$. Take y with $z = xy$. Then $C(x) + C(y) < |x| - k + |y| + k = |x| + |y| = |z| \leq C(z)$. □

Call a string σ *compressible* if $C(\sigma) < |\sigma|$. Then the Martin-Löf lemma says that any sufficiently long string has compressible initial segments, and any sufficiently long random string can be decomposed into two strings the sum of whose complexities is strictly less than that of the original string. Thus, the decomposition itself stores information additional to that contained in the strings separately.

The logarithms below are with respect to base 2, so $\log 2^n = n$.

Lemma 74. $C(xy) \leq C(x) + C(y) + 2\log C(x) + \text{const}$.

Proof. Let $f : 2^{<\mathbb{N}} \rightharpoonup 2^{<\mathbb{N}}$ be the partial recursive function that takes as inputs the strings $0^{|\rho|}1\rho\sigma\tau$, where $|\sigma| = \text{tr}(\rho)$, with output $f(0^{|\rho|}1\rho\sigma\tau) := u(\sigma)u(\tau)$.

Let x,y be given. Take σ,τ be of minimum length such that $u(\sigma) = x$ and $u(\tau) = y$. Then $C(x) = |\sigma| > 0$, and $C(y) = |\tau| > 0$. Take ρ such that $|\sigma| = \text{tr}(\rho)$. Then $2^{|\rho|-1} \leq 2^{|\rho|} - 1 \leq |\sigma| = C(x)$, so $\log C(x) \geq |\rho| - 1$. Then $f(0^{|\rho|}1\rho\sigma\tau) = xy$, and

$$\begin{aligned} |0^{|\rho|}1\rho\sigma\tau| = |\sigma| + |\tau| + 2|\rho| + 1 &\leq C(x) + C(y) + 2(\log(C(x)) - 1) + 1 \\ &= C(x) + C(y) + 2\log(C(x)) - 1, \end{aligned}$$

hence

$$C(xy) \leq C_f(xy) + \text{const}(f) \leq C(x) + C(y) + 2\log C(x) + \text{const}.$$

□

The next result has a striking consequence for the incompleteness of any given axiomatic framework for mathematics, as we discuss after the proof.

Proposition 75. *The set $A = \{x : C(x) \geq \frac{|x|}{2}\}$ of half-random strings is immune; that is, A has no infinite recursively enumerable subset. In particular, A is not recursively enumerable, and thus $C : 2^{<\mathbb{N}} \to \mathbb{N}$ is not recursive.*

Proof. Suppose B is an infinite r.e. subset of A. We can find a recursive function $h : 2^{<\mathbb{N}} \to 2^{<\mathbb{N}}$ such that for all σ, $h(\sigma) \in B$ and $|h(\sigma)| \geq \text{tr}(\sigma)$. Then we have $C(h(\sigma)) \geq \frac{|h(\sigma)|}{2} \geq \frac{\text{tr}(\sigma)}{2}$, and $C(h(\sigma)) \leq C(\sigma) + \text{const} \leq |\sigma| + \text{const}$. Thus, $\frac{\text{tr}(\sigma)}{2} \leq |\sigma| + \text{const}$, a contradiction.

□

By Lemma 71, almost all strings are half random: the probability that a string of length n is half random goes to 1 very rapidly as $n \to \infty$. As a consequence of Proposition 75, however, there are only finitely many σ for which ZFC can prove that σ is half random, assuming of course that ZFC is consistent. (Here ZFC is the standard axiomatic set theory in which essentially all of mathematics can be formally derived, including the entire contents of this course.) The same remains true for any recursively axiomatized consistent extension of ZFC.

3.2 Prefix-Free Sets

We say that σ and τ are *comparable* if σ is an initial segment of τ or τ is an initial segment of σ. We call a set $A \subseteq 2^{<\mathbb{N}}$ *prefix-free* if any two distinct $\sigma, \tau \in A$ are incomparable. Finite strings are going to be treated as initial segments of infinite strings of zeros and ones, and in the rest of this chapter we let α, β be such infinite strings, that is $\alpha, \beta : \mathbb{N} \to \{0,1\}$; we consider α, β also as points of the so-called *Cantor space* $2^{\mathbb{N}}$. We let $\alpha|n := (\alpha(0), \ldots, \alpha(n-1)) \in 2^{<\mathbb{N}}$ be the initial segment of α of length n.

We make $2^{\mathbb{N}}$ a topological space by giving it the product topology, where $\{0,1\}$ is given the discrete topology. So $2^{\mathbb{N}}$ has the clopen sets

$$N_\sigma := \{\alpha : \ \alpha(i) = \sigma(i) \text{ for all } i < |\sigma|\}$$

as a basis for its topology. Note that $N_\emptyset = 2^{\mathbb{N}}$, and that σ and τ are incomparable iff $N_\sigma \cap N_\tau = \emptyset$.

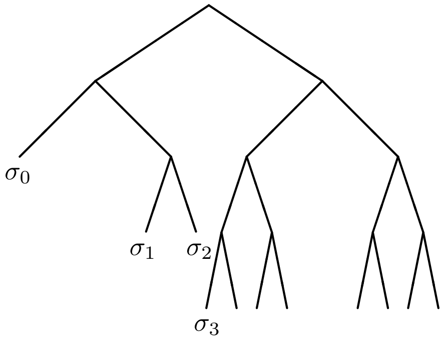
Let μ be the probability measure on the σ-algebra of Borel subsets of $2^{\mathbb{N}}$ (the σ-algebra on $2^{\mathbb{N}}$ generated by the N_σ) given by $\mu(N_\sigma) = 2^{-|\sigma|}$.

Lemma 76. *Suppose $A \subseteq 2^{<\mathbb{N}}$ is prefix-free. Then $\sum_{\sigma \in A} \frac{1}{2^{|\sigma|}} \leq 1$.*

Proof. $\sum_{\sigma \in A} \frac{1}{2^{|\sigma|}} = \sum_{\sigma \in A} \mu(N_\sigma) = \mu(\bigcup_{\sigma \in A} N_\sigma) \leq \mu(2^{\mathbb{N}}) = 1.$ □

Theorem 77 (Kraft-Chaitin)**.** *Let $(d_i)_{i \in \mathbb{N}}$ be a sequence of natural numbers such that $\sum_{i=0}^{\infty} \frac{1}{2^{d_i}} \leq 1$. Then there is an injective sequence $(\sigma_i)_{i \in \mathbb{N}}$ such that $|\sigma_i| = d_i$ for all i and $\{\sigma_i : i \in \mathbb{N}\}$ is prefix-free.*

Proof. We can assume that $d_0 \leq d_1 \leq d_2 \leq \ldots$. Then we choose $\sigma_0, \sigma_1, \sigma_2, \ldots$ successively by always choosing the left-most possibility. For example, when d_0, d_1, d_2, d_3 are $2, 3, 3, 4$, this is illustrated by the following picture.



The assumption that $\sum_{i=0}^{\infty} \frac{1}{2^{d_i}} \leq 1$ guarantees that we can always continue this process. We make this precise as follows. Take $\sigma_0 = 0 \dots 0$, with $|\sigma_0| = d_0$. By the induction hypothesis, suppose that at stage n, we have pairwise incomparable $\sigma_0, \dots, \sigma_n$ satisfying $|\sigma_i| = d_i$ for $i = 0, \dots, n$ and such that all α with $\mathrm{tr}(\alpha|d_n) \leq \mathrm{tr}(\sigma_n)$ belong to $N_{\sigma_0} \cup \dots \cup N_{\sigma_n}$. Then there is a σ with $|\sigma| = d_n$ and $\mathrm{tr}(\sigma) > \mathrm{tr}(\sigma_n)$; otherwise, all α would belong to $N_{\sigma_0} \cup \dots \cup N_{\sigma_n}$, yielding $1 = \mu(N_{\sigma_0} \cup \dots \cup N_{\sigma_n}) = \sum_{i=0}^{n} \frac{1}{2^{d_i}} < \sum_{i=0}^{\infty} \frac{1}{2^{d_i}} \leq 1$, a contradiction.

Let σ satisfy $|\sigma| = d_n$ and $\mathrm{tr}(\sigma) = \mathrm{tr}(\sigma_n) + 1$. Define $\sigma_{n+1} := \sigma 0 \dots 0$ with $|\sigma_{n+1}| = d_{n+1}$. It is easy to verify that $\sigma_0, \dots, \sigma_{n+1}$ are pairwise incomparable, and all α with $\mathrm{tr}(\alpha|d_{n+1}) \leq \mathrm{tr}(\sigma_{n+1})$ belong to $N_{\sigma_0} \cup \dots \cup N_{\sigma_{n+1}}$. □

This proof contains an inconstructive step, namely the initial rearrangement of the d_i in increasing order. It is desirable to have a more constructive proof which avoids this.

Constructive proof. We make the inductive assumption that at stage n, we have $\sigma_0, \dots, \sigma_n$ with $|\sigma_i| = d_i$ for $i = 0, \dots, n$, and for $0 \leq i < j \leq n$, σ_i and σ_j are incomparable and that in addition we have $\tau_1^n, \dots, \tau_p^n$, $p \geq 1$, such that

$$|\tau_1^n| < \dots < |\tau_p^n|, \qquad 2^{\mathbb{N}} = N_{\sigma_0} \sqcup \dots \sqcup N_{\sigma_n} \sqcup N_{\tau_1^n} \sqcup \dots \sqcup N_{\tau_p^n},$$

where the square union signs indicate a *disjoint* union. Here p depends on n. For $n = 0$, the inductive assumption is realized as follows: note that $d_0 > 0$ and put $\sigma_0 = 0 \dots 0$, with $|\sigma_0| = d_0$. If $d_0 = 1$, set $p = 1$, and $\tau_1^0 = 1$; if $d_0 \geq 2$, set $p = d_0$, and $\tau_i^0 = 0 \dots 01$, with $|\tau_i^0| = i$.

Given stage n, we construct σ_{n+1} at stage $n+1$ as follows. We are given d_{n+1}. If $|\tau_i^n| = d_{n+1}$, $1 \leq i \leq p$, then we set $\sigma_{n+1} := \tau_i^n$, so

$$2^{\mathbb{N}} = N_{\sigma_0} \sqcup \dots \sqcup N_{\sigma_n} \sqcup N_{\sigma_{n+1}} \sqcup N_{\tau_1^n} \sqcup \dots \sqcup N_{\tau_{i-1}^n} \sqcup N_{\tau_{i+1}^n} \sqcup \dots \sqcup N_{\tau_p^n}.$$

We put

$$\tau_j^{n+1} = \begin{cases} \tau_j^n & \text{if } 1 \leq j \leq i-1 \\ \tau_{j+1}^n & \text{if } i+1 \leq j \leq p. \end{cases}$$

Now suppose that for all $i \in \{1, \dots, p\}$, $|\tau_i^n| \neq d_{n+1}$. It is easy to check that $|\tau_1^n| < d_{n+1}$. Take $i \in \{1, \dots, p\}$ maximal such that $|\tau_i^n| < d_{n+1}$. Set $k := d_{n+1} - |\tau_i^n|$. Take $\rho_1, \dots, \rho_k, \sigma_{n+1}$ such that $\rho_j = \tau_i^n 0 \dots 01$, with $|\rho_j| = |\tau_i^n| + j$ (so $|\rho_k| = d_{n+1}$) for $1 \leq j \leq k$, and $\sigma_{n+1} = \tau_i^n 0 \dots 0$, with $|\sigma_{n+1}| = d_{n+1}$. Then

$$N_{\tau_i^n} = N_{\rho_1} \sqcup \dots \sqcup N_{\rho_k} \sqcup N_{\sigma_{n+1}}.$$

Thus replacing τ_i^n in the sequence $\tau_1^n, \dots, \tau_p^n$ by $\rho_1, \dots, \rho_k$ yields a sequence $\tau_1^{n+1}, \dots, \tau_{p+k-1}^{n+1}$ with the desired properties. □

This constructive proof of the Kraft-Chaitin Theorem yields:

Proposition 78. *Let $i \mapsto d_i : \mathbb{N} \to \mathbb{N}$ be recursive such that $\sum_{i=0}^{\infty} 2^{-d_i} \leq 1$. Then there is a recursive injective map $i \mapsto \sigma_i : \mathbb{N} \to 2^{<\mathbb{N}}$ such that $|\sigma_i| = d_i$ for all i and $\{\sigma_i : i \in \mathbb{N}\}$ is prefix-free.*

Here a function $f : \mathbb{N} \to 2^{<\mathbb{N}}$ is said to be recursive if $\mathrm{tr}(f) : \mathbb{N} \to \mathbb{N}$ is recursive.

Exercise. Show that for each $Y \subseteq 2^{<\mathbb{N}}$ there is a prefix-free $X \subseteq Y$ such that

$$\bigcup_{\sigma \in Y} N_\sigma = \bigcup_{\sigma \in X} N_\sigma .$$

3.3 Prefix-Free Complexity

A partial recursive function $f : 2^{<\mathbb{N}} \rightharpoonup 2^{<\mathbb{N}}$ is said to be *prefix-free* if its domain $D(f) \subseteq 2^{<\mathbb{N}}$ is prefix-free. In order to construct an effective enumeration of

$$\{f : 2^{<\mathbb{N}} \rightharpoonup 2^{<\mathbb{N}} | \ f \text{ is partial recursive and prefix-free}\},$$

we first associate to each partial recursive $f : \mathbb{N} \rightharpoonup \mathbb{N}$ the partial recursive $\hat{f} : 2^{<\mathbb{N}} \rightharpoonup 2^{<\mathbb{N}}$ by $\mathrm{tr}(\hat{f}(\sigma)) \simeq f(\mathrm{tr}(\sigma))$. Next, recall our enumeration (ϕ_e) of $\{f : \mathbb{N} \rightharpoonup \mathbb{N} | \ f \text{ is partial recursive}\}$. We defined $\phi_{e,s} : \mathbb{N} \rightharpoonup \mathbb{N}$ for $s \in \mathbb{N}$ by

$$\phi_{e,s}(x) \simeq \begin{cases} \phi_e(x) & \text{if } \exists z \leq s \ \ T_1(e, x, z) \\ \uparrow & \text{otherwise} \end{cases}$$

with $D(\phi_{e,s}) \subseteq \{0, \dots, s\}, \phi_e = \bigcup_s \phi_{e,s}$. This yields an enumeration $(\hat{\phi}_e)$ of

$$\{f : 2^{<\mathbb{N}} \rightharpoonup 2^{<\mathbb{N}} | \ f \text{ is partial recursive}\}$$

and for each s a function $\hat{\phi}_{e,s} : 2^{<\mathbb{N}} \rightharpoonup 2^{<\mathbb{N}}$ with

$$D(\hat{\phi}_{e,s}) \subseteq \{\sigma : \mathrm{tr}(\sigma) \in \{0, \dots, s\}\},$$

such that $\hat{\phi}_e = \bigcup_s \hat{\phi}_{e,s}$ for each e. Define $\psi_{e,s} : 2^{<\mathbb{N}} \rightharpoonup 2^{<\mathbb{N}}$ by recursion on s :

$$\begin{aligned} \psi_{e,0} &= \hat{\phi}_{e,0} \\ \psi_{e,s+1} &= \begin{cases} \hat{\phi}_{e,s+1} & \text{if } \hat{\phi}_{e,s+1} \text{ is prefix-free} \\ \psi_{e,s} & \text{otherwise} \end{cases} \end{aligned}$$

So $\phi_{e,0} \subseteq \psi_{e,1} \subseteq \psi_{e,2} \subseteq \dots$. Put

$$\psi_e := \bigcup_s \psi_{e,s} : 2^{<\mathbb{N}} \rightharpoonup 2^{<\mathbb{N}}.$$

A bit of thought shows:

(a) each ψ_e is partial recursive and prefix-free;

(b) (ψ_e) is an enumeration of

$$\{f : 2^{<\mathbb{N}} \rightharpoonup 2^{<\mathbb{N}} \, | \, f \text{ is partial recursive and prefix-free}\};$$

(c) there is a recursive $\iota : \mathbb{N} \to \mathbb{N}$ such that $\psi_e = \hat{\phi}_{\iota(e)}$ for all e.

Define $\Psi : 2^{<\mathbb{N}} \rightharpoonup 2^{<\mathbb{N}}$ by

$$\begin{aligned} \Psi(0^e 1\sigma) &\simeq \psi_e(\sigma), \\ \Psi(\rho)\uparrow &\quad \text{if } \rho \neq 0^e 1\sigma \text{ for all } e \text{ and } \sigma. \end{aligned}$$

Note: Ψ is partial recursive, prefix-free, surjective, and $C_\Psi(\sigma) \in \mathbb{N}^{>0}$ for all σ.

Lemma 79. *Let $f : 2^{<\mathbb{N}} \rightharpoonup 2^{<\mathbb{N}}$ be prefix-free and partial recursive. Then*

$$C_\Psi \leq C_f + \mathrm{const}(f)$$

Proof. Take e such that $f = \psi_e$. Then $C_\Psi \leq C_f + (e+1)$. □

Definition 80. *The prefix-free complexity of a string σ is*

$$K(\sigma) := C_\Psi(\sigma) \in \mathbb{N}^{>0}.$$

Lemma 81. *There are infinitely many σ such that $K(\sigma) > |\sigma| + \log|\sigma|$.*

Proof. Suppose $K(\sigma) \leq |\sigma| + \log|\sigma|$ for all but finitely many σ. Then, with σ ranging over the *nonempty* strings, we have

$$\begin{aligned} \sum_\sigma 2^{-K(\sigma)} &\geq \sum_\sigma 2^{-(|\sigma|+\log|\sigma|)} + \mathrm{const} \\ &= \sum_{n=1}^{\infty} \sum_{|\sigma|=n} 2^{-(n+\log n)} + \mathrm{const} \\ &= \sum_{n=1}^{\infty} 2^n 2^{-(n+\log\, n)} + \mathrm{const} \\ &= \sum_{n=1}^{\infty} \frac{1}{n} + \mathrm{const} = \infty. \end{aligned}$$

Since Ψ is surjective, we have an injective map $\sigma \mapsto \sigma' : 2^{<\mathbb{N}} \to D(\Psi)$ such that $\Psi(\sigma') = \sigma$ and $K(\sigma) = |\sigma'|$. for all σ. Since Ψ is prefix-free,

$$1 \geq \sum_{\rho \in D(\Psi)} 2^{-|\rho|} \geq \sum_\sigma 2^{-|\sigma'|} = \sum_\sigma 2^{-K(\sigma)} = \infty,$$

a contradiction. □

The only fact about log that we used is that $\sum_{n=1}^{\infty} 2^{-\log n} = \infty$. The same proof yields the following generalization:

Lemma 82. *If $f : \mathbb{N}^{>0} \to \mathbb{R}^{>0}$ satisfies $\sum_{n=1}^{\infty} 2^{-f(n)} = \infty$, then there are infinitely many σ such that $K(\sigma) > |\sigma| + f(|\sigma|)$.*

Let $n^* \in 2^{<\mathbb{N}}$ be the string such that $\mathrm{tr}(n^*) = n$. For example, if σ has length $1000 \approx 2^{10}$, then $|\sigma|^*$ is the string whose tree code is 1000, thus $|\sigma|^*$ has length about 10. More generally, $\mathrm{length}(|\sigma|^*) \approx \log(\mathrm{length}(\sigma))$.

Proposition 83. $K(\sigma) < |\sigma| + 2\log|\sigma| + \text{const}$ *for nonempty* σ.

Proof. It suffices to find a prefix-free partial recursive $f : 2^{<\mathbb{N}} \rightharpoonup 2^{<\mathbb{N}}$ and a $c \in \mathbb{N}$ such that for all nonempty σ there exists ρ with

$$f(\rho) \simeq \sigma, \qquad |\rho| < |\sigma| + 2\log|\sigma| + c.$$

We shall construct such f via the effective version of the Kraft-Chaitin Theorem, using $\sum_{k=0} \frac{1}{10(k+1)^2} \leq 1$. Let $n_k = \lfloor \frac{2^{k+1}}{10(k+1)^2} \rfloor \in \mathbb{N}$. Note that $n_k \to \infty$ as $k \to \infty$. Define $d_i \in \mathbb{N}$ by

$$\begin{array}{rll} d_i = 1 & \text{for} & 0 \leq i < n_0, \\ d_i = 2 & \text{for} & n_0 \leq i < n_0 + n_1, \\ d_i = 3 & \text{for} & n_0 + n_1 \leq i < n_0 + n_1 + n_2, \\ & \vdots & \\ d_i = k+1 & \text{for} & n_0 + n_1 + \cdots + n_{k-1} \leq i < n_0 + n_1 + \cdots + n_k. \end{array}$$

Then

$$\begin{aligned} \sum_{i=0}^{\infty} 2^{-d_i} &= n_0 2^{-1} + n_1 2^{-2} + \dots \\ &= \sum_{k=0}^{\infty} n_k 2^{-(k+1)} \\ &\leq \sum_{k=0}^{\infty} \frac{2^{k+1}}{10(k+1)^2} \cdot 2^{-(k+1)} \\ &= \sum_{k=0}^{\infty} \frac{1}{10(k+1)^2} \\ &\leq 1. \end{aligned}$$

The function $i \mapsto d_i : \mathbb{N} \to \mathbb{N}$ is recursive, so by the Kraft-Chaitin Theorem we have an injective recursive map $i \mapsto \rho_i : \mathbb{N} \to 2^{<\mathbb{N}}$ such that $|\rho_i| = d_i$ for all i, and $\{\rho_i : i \in \mathbb{N}\}$ is prefix-free. Now define $f : \{\rho_i : i \in \mathbb{N}\} \to 2^{<\mathbb{N}}$ by $f(\rho_n) = n^*$. Note that $f : 2^{<\mathbb{N}} \rightharpoonup 2^{<\mathbb{N}}$ is prefix-free, partial recursive, and surjective.

To show f has the above property, note first that there are n_k strings in $\{\rho_i : i \in \mathbb{N}\}$ of length $k+1$. For $m > 0$ and $k = \lfloor m + 2\log m + c \rfloor$, $c \in \mathbb{N}$,

$$
\begin{aligned}
n_k &\geq \frac{2^{k+1}}{10(k+1)^2} - 1 \\
&\geq \frac{2^{m+2\log m+c}}{10(m+2\log m+c+1)^2} - 1 \\
&\geq \frac{2^m \cdot 2^{\log m^2} \cdot 2^c}{10(m+2\log m+c+1)^2} - 1 \\
&\geq \frac{2^m \cdot m^2 \cdot 2^c}{10(m+2\log m+c+1)^2} - 1 \\
&\geq 2^{m+1} \text{ for } c = 10.
\end{aligned}
$$

Therefore, if $0 < |\sigma| \leq m$, we obtain ρ_i such that

$$|\rho_i| \leq m + 2\log m + 11, \text{ and } f(\rho_i) = \sigma.$$

□

Theorem 84. $K(\sigma) \leq |\sigma| + K(|\sigma|^*) + \text{const}.$

Proof. Since Ψ is prefix-free, each string σ has at most initial segment $\rho \in D(\Psi)$. So we can define a prefix-free $f : 2^{<\mathbb{N}} \to 2^{<\mathbb{N}}$ by

$$f(\sigma) \simeq \tau \iff \exists \rho\ (\sigma = \rho\tau \text{ and } \Psi(\rho) \simeq |\tau|^*).$$

Then f is partial recursive, so $K(\tau) = C_\Psi(\tau) \leq C_f(\tau) + \text{const}.$

We claim that $C_f(\tau) \leq |\tau| + K(|\tau|^*)$. To see this, take ρ of minimal length such that $\Psi(\rho) \simeq |\tau|^*$. Then $|\rho| = K(|\tau|^*)$, and $f(\rho\tau) \simeq \tau$. Hence,

$$C_f(\tau) \leq |\rho\tau| = |\tau| + |\rho| = |\tau| + K(|\tau|^*).$$

From this claim, we obtain

$$K(\tau) \leq C_f(\tau) + \text{const} \leq |\tau| + K(|\tau|^*) + \text{const}\,.$$

□

3.4 Information Content Measures

Here and below we let k range over $\mathbb{N}$.

Definition 85. *An* ***information content measure*** *(icm) is a partial function* $I : 2^{<\mathbb{N}} \rightharpoonup \mathbb{N}$ *such that*

(a) $\sum_{\sigma \in D(I)} 2^{-I(\sigma)} \leq 1$*;*

(b) the set $\{(\sigma, k) : \sigma \in D(I), I(\sigma) \leq k\} \subseteq 2^{<\mathbb{N}} \times \mathbb{N}$ *is recursively enumerable.*

The next two results establish a close connection between partial recursive prefix-free functions and icm's, almost a one-to-one correspondence.

Lemma 86. *Let $\psi : 2^{<\mathbb{N}} \rightharpoonup 2^{<\mathbb{N}}$ be prefix-free partial recursive. Then the partial function $I_\psi : 2^{<\mathbb{N}} \rightharpoonup \mathbb{N}$ with $D(I_\psi) = \operatorname{Im}(\psi)$, and $I_\psi(\sigma) = C_\psi(\sigma)$ for $\sigma \in \operatorname{Im}(\psi)$ is an icm.*

Proof.

$$\sum_{\sigma \in D(I_\psi)} 2^{-I_\psi(\sigma)} = \sum_{\sigma \in \operatorname{Im}(\psi)} 2^{-C_\psi(\sigma)} \leq \sum_{\rho \in D(\psi)} 2^{-|\rho|} \leq 1.$$

Also, $\sigma \in D(I_\psi), I_\psi(\sigma) \leq k \iff \exists \rho\big(|\rho| \leq k, \psi(\rho) \simeq \sigma\big)$. Therefore, the set

$$\{(\sigma, k) : \sigma \in D(I_\psi), I_\psi(\sigma) \leq k\} \subseteq 2^{<\mathbb{N}} \times \mathbb{N}$$

is recursively enumerable. □

For $\psi = \Psi$, this yields $I_\psi = K$, so K is an icm. Conversely,

Lemma 87. *Let I be an icm with infinite domain $D(I)$. Then there is a prefix-free partial recursive map $\psi :\ 2^{<\mathbb{N}} \rightharpoonup 2^{<\mathbb{N}}$ such that*

(a) $D(I) = \operatorname{Im}(\psi)$;

(b) $I(\sigma) + 1 = C_\psi(\sigma)$ for all $\sigma \in D(I)$.

Proof. Take a recursive map $i \mapsto (\sigma_i, k_i) : \mathbb{N} \to 2^{<\mathbb{N}} \times \mathbb{N}$ with image

$$\{(\sigma, k) : \sigma \in D(I),\ I(\sigma) \leq k\}.$$

We construct a set $A \subseteq \{(\sigma, k+1) : \sigma \in D(I), I(\sigma) \leq k\}$ in stages.

- *Stage* 0*:* Put $(\sigma_0, k_0 + 1)$ in A.
- *Stage* $s+1$*:* Compute (σ_{s+1}, k_{s+1}), see whether $k_{s+1} < k_i$ for all $i \leq s$ with $\sigma_i = \sigma_{s+1}$, and if so, put $(\sigma_{s+1}, k_{s+1} + 1)$ in A.

Note: $\sigma \in D(I) \iff \exists d(\sigma, d) \in A$: the direction $\Leftarrow$ is clear, and $\Rightarrow$ holds in the more precise form that if $\sigma \in D(I)$, then $\min\{d : (\sigma, d) \in A\} = I(\sigma) + 1$.

Put $A(\sigma) := \{d : (\sigma, d) \in A\}$. Then, for $\sigma \in D(I)$,

$$\sum_{d \in A(\sigma)} 2^{-d} \leq 2^{-(I(\sigma)+1)} + 2^{-(I(\sigma)+2)} + \cdots \leq 2 \cdot 2^{-(I(\sigma)+1)} = 2^{-I(\sigma)}, \text{ so}$$

$$\sum_{(\sigma,d) \in A} 2^{-d} = \sum_{\sigma \in D(I)} \sum_{d \in A(\sigma)} 2^{-d} \leq \sum_{\sigma \in D(I)} 2^{-I(\sigma)} \leq 1.$$

By construction, A is r.e. and infinite, so there is an injective recursive map

$$i \mapsto (\sigma_i', d_i) : \mathbb{N} \to 2^{<\mathbb{N}} \times \mathbb{N}$$

with image A. By the last inequality we have $\sum_{i=0}^{\infty} 2^{-d_i} \leq 1$, so the recursive version of Kraft-Chaitin yields a recursive injective map $i \mapsto \rho_i : \mathbb{N} \to 2^{<\mathbb{N}}$ such that $|\rho_i| = d_i$ for all i, and $\{\rho_i : i \in \mathbb{N}\}$ is prefix-free. Now define $\psi : 2^{<\mathbb{N}} \rightharpoonup \mathbb{N}$ to have domain $D(\psi) = \{\rho_i : i \in \mathbb{N}\}$ and $\psi(\rho_i) = \sigma_i'$ for all i. Then ψ is prefix-free partial recursive and $\mathrm{Im}(\psi) = D(I)$. It remains to check that (b) is satisfied:

Let $\sigma \in D(I) = \mathrm{Im}(\psi)$. Then

$$C_\psi(\sigma) = \min\{d_i : i \in \mathbb{N}, \sigma = \sigma_i'\} = \min\{d : (\sigma, d) \in A\} = I(\sigma) + 1.$$

□

The next lemma follows from K being an icm with $D(K) = 2^{<\mathbb{N}}$.

Lemma 88. *The set*

$$\{(m, n, k) : \sum_{|\sigma|=m} 2^{-K(\sigma)} \geq \frac{n}{2^k}\} \subseteq \mathbb{N}^3$$

is recursively enumerable.

Proof. Just note that $\sum_{|\sigma|=m} 2^{-K(\sigma)} \geq \frac{n}{2^k}$ if and only if there is a tuple $(n_\sigma)_{|\sigma|=m}$ of natural numbers such that

$$\sum_{|\sigma|=m} 2^{-n_\sigma} \geq \frac{n}{2^k} \quad \text{and } K(\sigma) \leq n_\sigma \text{ for each } \sigma \text{ with } |\sigma| = m.$$

□

Lemma 89. *There is an enumeration $(I_e)_{e \in \mathbb{N}}$ of the set of icms such that*

$$\{(\sigma, e, k) \in 2^{<\mathbb{N}} \times \mathbb{N} \times \mathbb{N} : \sigma \in D(I_e),\ I_e(\sigma) \leq k\}$$

is recursively enumerable.

Proof. Note that we have a recursive bijection

$$(\sigma, n) \mapsto |\,\mathrm{tr}(\sigma), n| \ :\ 2^{<\mathbb{N}} \times \mathbb{N} \to \mathbb{N}.$$

Using this bijection the sets W_e and their finite approximations $W_{e,s}$ yield an enumeration (V_e) of the set of recursively enumerable subsets of $2^{<\mathbb{N}} \times \mathbb{N}$ and a family $(V_{e,s})$ of finite subsets of $2^{<\mathbb{N}} \times \mathbb{N}$ with the following properties:

(a) $V_{e,s} \subseteq \{(\sigma, n) : |\,\mathrm{tr}(\sigma), n| \leq s\}$;

(b) $|V_{e,0}| \leq 1$, $V_{e,s} \subseteq V_{e,s+1}$, and $\bigcup_s V_{e,s} = V_e$;

(c) $\{(\sigma, n, e, s) \in 2^{<\mathbb{N}} \times \mathbb{N}^3 : (\sigma, n) \in V_{e,s}\}$ is recursive.

Define $\eta_{e,s} : 2^{<\mathbb{N}} \rightharpoonup \mathbb{N}$ by $\eta_{e,s}(\sigma) := \mu n\big((\sigma, n) \in V_{e,s}\big)$, so $D(\eta_{e,s})$ is finite, $D(\eta_{e,s}) \subseteq D(\eta_{e,s+1})$, and $\eta_{e,s+1}(\sigma) \le \eta_{e,s}(\sigma)$ for all $\sigma \in D(\eta_{e,s})$. Moreover, the partial map

$$(e, s, \sigma) \mapsto \eta_{e,s}(\sigma) : \ \mathbb{N}^2 \times 2^{<\mathbb{N}} \rightharpoonup \mathbb{N}$$

is partial recursive and has recursive domain. Next, define $I_{e,s} : 2^{<\mathbb{N}} \rightharpoonup \mathbb{N}$ by recursion on s: $I_{e,0} := \eta_{e,0}$, $I_{e,s+1} = \eta_{e,s+1}$ if $\eta_{e,s+1}$ is an icm, and $I_{e,s+1} = I_{e,s}$ otherwise. Identifying functions with their graphs, it is easy to check that $I_{e,s}$ is a finite icm, $D(I_{e,s}) \subseteq D(I_{e,s+1})$, and $I_{e,s+1}(\sigma) \le I_{e,s}(\sigma)$ for all $\sigma \in D(I_{e,s})$, and

$$(e, s, \sigma) \mapsto I_{e,s}(\sigma) : \ \mathbb{N}^2 \times 2^{<\mathbb{N}} \rightharpoonup \mathbb{N}$$

is partial recursive and has recursive domain. In particular, the set

$$\{(\sigma, n, e, s) \in 2^{<\mathbb{N}} \times \mathbb{N}^3 : I_{e,s}(\sigma) \simeq n\}$$

is recursive. Define $I_e : 2^{<\mathbb{N}} \rightharpoonup \mathbb{N}$ by

$$D(I_e) = \bigcup_s D(I_{e,s}), \qquad I_e(\sigma) = \min\{I_{e,s}(\sigma) : \sigma \in D(I_{e,s})\}.$$

Given any icm I, take e such that $V_e = \{(\sigma, k) : \sigma \in D(I),\ I(\sigma) \le k\}$, and note that then $I = I_e$. It follows easily that (I_e) is an enumeration of the set of icms as required by the lemma. □

Fix an enumeration (I_e) of the set of icms as in the lemma above. This allows us to define $\hat{K} : \ 2^{<\mathbb{N}} \to \mathbb{N}$ by

$$\hat{K}(\sigma) := \min\{I_e(\sigma) + e + 1 : I_e(\sigma) \downarrow\}$$

Lemma 90. *$\hat{K}$ is an icm.*

Proof. Condition (a) in the definition of *icm* holds for $\hat{K}$:

$$\begin{aligned}\sum_\sigma 2^{-\hat{K}(\sigma)} &\le \sum_{e=0}^{\infty} \sum_{\sigma \in D(I_e)} 2^{-(I_e(\sigma)+e+1)} \\ &= \sum_{e=0}^{\infty} 2^{-(e+1)} \sum_{\sigma \in D(I_e)} 2^{-I_e(\sigma)} \\ &\le \sum_{e=0}^{\infty} 2^{-(e+1)} = 1.\end{aligned}$$

Condition (b) is satisfied by $\hat{K}$, since

$$\hat{K}(\sigma) \le k \iff \exists e(\sigma \in D(I_e),\ I_e(\sigma) + e + 1 \le k).$$

□

This icm $\hat{K}$ is also called the *minimal icm*. We can use it to prove results about K because of the following fact.

Corollary 91. $|K - \hat{K}| \leq \text{const}$.

Proof. We already know that $\hat{K} \leq K + \text{const}$. Since $\hat{K}$ is an icm, Lemma 87 gives a prefix-free partial recursive $\psi : 2^{<\mathbb{N}} \rightharpoonup 2^{<\mathbb{N}}$ such that $C_\psi = \hat{K} + 1$. Hence, $K(\sigma) \leq C_\psi(\sigma) + \text{const} = \hat{K}(\sigma) + 1 + \text{const}$. □

Theorem 92 (Counting Theorem)**.** *We have*

(a) $|\{\sigma : |\sigma| = n,\ K(\sigma) < n + K(n^*) - k\}| \leq 2^{n-k+\text{const}}$*;*

(b) $n + K(n^*) + \text{const} \leq \max\{K(\sigma) : |\sigma| = n\} \leq n + K(n^*) + \text{const}$*.*

Proof. Assuming (a) for the moment, we show how (b) follows. Let $c \in \mathbb{N}$ be a value of const for which (a) holds. Then

$$|\{\sigma : |\sigma| = n,\ K(\sigma) < n + K(n^*) - c - 1\}| \leq 2^{n-1} < 2^n.$$

So we get σ with $|\sigma| = n$ such that $K(\sigma) \geq n + K(n^*) - c - 1$. This gives the lower bound in (b). The upper bound is given by Theorem 84.

To prove (b), note first that

$$\begin{aligned} 1 &\geq \sum_\sigma 2^{-K(\sigma)} \\ &= \sum_{n=0}^{\infty} \sum_{|\sigma|=n} 2^{-K(\sigma)} \\ &= \sum_\rho \sum_{|\sigma|=\text{tr}(\rho)} 2^{-K(\sigma)}. \end{aligned}$$

Take $I(\rho) \in [1, \infty)$ with $2^{-I(\rho)} = \sum_{|\sigma|=\text{tr}(\rho)} 2^{-K(\sigma)}$, that is,

$$I(\rho) = -\log \left(\sum_{|\sigma|=\text{tr}(\rho)} 2^{-K(\sigma)} \right).$$

Then $\sum_\rho 2^{-I(\rho)} \leq 1$, so I is like an icm, except that I may have values not in $\mathbb{N}$. So we modify I to $\hat{I} : 2^{<\mathbb{N}} \to \mathbb{N}^{>0}$ by $\hat{I}(\rho) := \lceil I(\rho) \rceil$.

Claim. $\hat{I}$ is an information content measure.

To prove this claim, note that $\sum_\rho 2^{-\hat{I}(\rho)} \leq \sum_\rho 2^{-I(\rho)} \leq 1$, and

$$\begin{aligned} \hat{I}(\rho) \leq k &\iff I(\rho) \leq k \\ &\iff 2^{-I(\rho)} \geq 2^{-k} \\ &\iff \sum_{|\sigma|=\text{tr}(\rho)} 2^{-K(\sigma)} \geq 2^{-k}. \end{aligned}$$

Using for example lemma 88 it follows that the set

$$\{(\rho,k):\ \hat{I}(\rho)\leq k\}\ =\ \{(\rho,k):\ \sum_{|\sigma|=\mathrm{tr}(\rho)} 2^{-K(\sigma)}\geq 2^{-k}\}$$

is recursively enumerable. This finishes the proof of the claim.

This claim and earlier results yield a $c\in\mathbb{N}$ such that $K(\sigma)\leq\hat{I}(\sigma)+c$ for all σ. Then

$$\begin{aligned} K(n^*) &\leq \hat{I}(n^*)+c \\ &\leq I(n^*)+c+1 \\ &= -\log\left(\sum_{|\sigma|=n}2^{-K(\sigma)}\right)+c+1. \end{aligned}$$

Therefore,

$$2^{-K(n^*)+c+1}\geq\sum_{|\sigma|=n}2^{-K(\sigma)}.$$

Let $\mathcal{F}:=\{\sigma:|\sigma|=n,\ K(\sigma)<n+K(n^*)-k\}$, and assume $|\mathcal{F}|\geq 2^{n-k+c+1}$. Then by the above,

$$\begin{aligned} 2^{-K(n^*)+c+1} &\geq \sum_{|\sigma|=n}2^{-K(\sigma)} \\ &\geq \sum_{\sigma\in\mathcal{F}}2^{-K(\sigma)} \\ &> 2^{n-k+c+1}\cdot 2^{-n-K(n^*)+k} \\ &= 2^{-K(n^*)+c+1}, \end{aligned}$$

a contradiction. □

Likewise, we obtain the following.

Proposition 93. $|\{\sigma:|\sigma|=n,\ K(\sigma)<n-k\}|\leq 2^{n-K(n^*)-k+\mathrm{const}}$.

Proof. Take c as in the previous proof. Let $\mathcal{F}:=\{\sigma:|\sigma|=n,K(\sigma)<n-k\}$. Assume $|\mathcal{F}|>2^{n-K(n^*)-k+c+1}$. Then

$$\begin{aligned} 2^{-K(n^*)+c+1} &\geq \sum_{|\sigma|=n}2^{-K(\sigma)} \\ &\geq \sum_{\sigma\in\mathcal{F}}2^{-K(\sigma)} \\ &> 2^{n-K(n^*)-k+c+1}\cdot 2^{-(n-k)} \\ &= 2^{-K(n^*)+c+1}, \end{aligned}$$

a contradiction. □

3.5 Martin-Löf Randomness

A set $T \subseteq \mathbb{N} \times 2^{<\mathbb{N}}$ is called a **Martin-Löf test** if T is recursively enumerable and for all n,

$$\mu\left(\bigcup_{\sigma \in T(n)} N_\sigma\right) \leq 2^{-n} \quad \text{where } T(n) = \{\sigma : (n, \sigma) \in T\}.$$

A set $B \subseteq 2^{\mathbb{N}}$ is said to have *effective measure zero* if there is a Martin-Löf test T such that $B \subseteq \bigcap_n \left(\bigcup_{\sigma \in T(n)} N_\sigma\right)$. Note that $\mu(B) = 0$ for such B, and that every Martin-Löf test T determines a largest subset of $2^{\mathbb{N}}$ of effective measure zero *witnessed* by T, namely

$$N_T \ := \bigcap_n \left(\bigcup_{\sigma \in T(n)} N_\sigma\right).$$

We say that α *passes the Martin-Löf test* T if $\alpha \notin N_T$. We say that α is *1-random* if it passes every Martin-Löf test. We are going to show that there is Martin-Löf test such that passing it is equivalent to passing all Martin-Löf tests. Such a Martin-Löf test is said to be *universal*. Note that if T is a universal Martin-Löf test, then N_T is the largest subset of $2^{\mathbb{N}}$ of effective measure zero; in particular, this set is independent of the choice of universal Martin-Löf test. This is somewhat remarkable, since for, say, Lebesgue measure on $\mathbb{R}$, there is obviously no largest subset of $\mathbb{R}$ of measure zero.

Theorem 94 (**Martin-Löf**). *There exists a universal Martin-Löf test.*

Proof. Let $R \subseteq \mathbb{N} \times \mathbb{N} \times 2^{<\mathbb{N}}$ be recursively enumerable such that $(R(e))_{e \in \mathbb{N}}$ enumerates

$$\{S \subseteq \mathbb{N} \times 2^{<\mathbb{N}} : S \text{ is recursively enumerable}\}$$

and such that for all e there exists infinitely many e' with $R(e) = R(e')$. Take a recursive function $f : \mathbb{N} \to \mathbb{N} \times \mathbb{N} \times 2^{<\mathbb{N}}$ such that Image$(f) = R$ and define

$$R_t = \{f(0), f(1), \ldots, f(t)\} \subseteq R, \quad t \in \mathbb{N}.$$

Then $R_0 \subseteq R_1 \subseteq \ldots$, $R = \bigcup_{t \in \mathbb{N}} R_t$ and

$$R_t(e, n, \sigma) \Longleftrightarrow \exists s \leq t \, [f(s) = (e, n, \sigma)],$$

so the set

$$\{(t, e, n, \sigma) : R_t(e, n, \sigma)\} \subseteq \mathbb{N} \times \mathbb{N} \times \mathbb{N} \times 2^{<\mathbb{N}}$$

is recursive.

Define $T \subseteq \mathbb{N} \times 2^{<\mathbb{N}}$ by

$$T(k, \sigma) \iff \exists n_{\geq k+1} \, \exists t \left[R_t(n, n, \sigma) \;\&\; \mu\left(\bigcup_{\tau \in R_t(n,n)} N_\tau\right) \leq 2^{-n}\right].$$

It is is rather clear from the Church-Turing Thesis that T is recursively enumerable. Note also that $T(0) \supseteq T(1) \supseteq T(2) \supseteq \dots$. Put $B^n_t := \bigcup_{\tau \in R_t(n,n)} N_\tau$, so $B^n_0 \subseteq B^n_1 \subseteq \dots$. Moreover,

$$\begin{aligned}
\mu\left(\bigcup_{\sigma \in T(k)} N_\sigma\right) &= \mu\left(\bigcup_{n \geq k+1} \bigcup_{\mu(B^n_t) \leq 2^{-n}} B^n_t\right) \\
&\leq \sum_{n \geq k+1} \mu\left(\bigcup_{\mu(B^n_t) \leq 2^{-n}} B^n_t\right) \\
&\leq \sum_{n \geq k+1} 2^{-n} \\
&= 2^{-k}.
\end{aligned}$$

Thus T is a Martin-Löf test. To see that T is universal, let $S \subseteq \mathbb{N} \times 2^{<\mathbb{N}}$ be any Martin-Löf test and take m such that $R(m) = S$. Fix any k and take $n \geq k+1$ such that $R(n) = R(m) = S$. Then

$$N_S = \bigcap_{l \in \mathbb{N}} \bigcup_{\sigma \in S(l)} N_\sigma = \bigcap_{l \in \mathbb{N}} \bigcup_{\sigma \in R(n,l)} N_\sigma \subseteq \bigcup_{\sigma \in R(n,n)} N_\sigma.$$

Now $B^n_t \subseteq \bigcup_{\sigma \in R(n,n)} N_\sigma$ for each t, and $R(n)$ is a Martin-Löf test, so

$$\mu\left(B^n_t\right) \leq \mu\left(\bigcup_{\sigma \in R(n,n)} N_\sigma\right) \leq 2^{-n}.$$

In particular, $\sigma \in R(n,n) \Rightarrow \sigma \in T(k)$. Hence, $\bigcap_l \bigcup_{\sigma \in S(l)} N_\sigma \subseteq \bigcup_{\sigma \in T(k)} N_\sigma$. Since k was arbitrary, we obtain

$$N_S = \bigcap_l \bigcup_{\sigma \in S(l)} N_\sigma \subseteq \bigcap_k \bigcup_{\sigma \in T(k)} N_\sigma = N_T,$$

so N_T is a largest set of effective measure zero. □

Corollary 95. *The set* $\{\alpha : \alpha$ *is 1-random*$\}$ *has measure 1. In other words, almost all* α *are 1-random.*

Exercise. Show that no recursive α is 1-random.

We say that α is *LGC-random* if there exists $c \in \mathbb{N}$ such that $K(\alpha|n) \geq n - c$ for all n. Note that $\alpha|n = \alpha(0), \alpha(1), \dots, \alpha(n-1)$ is the initial segment of α of length n.

It wouldn't be interesting to replace K by C in this definition, as by Martin-Löf's compressability lemma, given any α and any and $c \in \mathbb{N}$, there exists infinitely many n such that $C(\alpha|n) < n - c$.

Theorem 96 (Schnorr).

$$\alpha \textit{ is LGC-random} \iff \alpha \textit{ is 1-random.}$$

Proof. (of $\Leftarrow$) Define $T \subseteq \mathbb{N} \times 2^{<\mathbb{N}}$ by $T(k,\sigma) \iff K(\sigma) \leq |\sigma| - k$.

Claim. T is a Martin-Löf test.

To prove this claim note first that T is recursively enumerable. Moreover,

$$\begin{aligned} \mu\left(\bigcup_{\sigma \in T(k)} N_\sigma\right) &\leq \sum_{K(\sigma) \leq |\sigma| - k} 2^{-|\sigma|} \\ &\leq \sum_\sigma 2^{-K(\sigma)-k} \\ &= 2^{-k} \sum_\sigma 2^{-K(\sigma)} \\ &\leq 2^{-k}. \end{aligned}$$

Thus T is a Martin-Löf test. Now suppose α is 1-random. Then α passes T, so we obtain k such that $\alpha \notin \bigcup_{\sigma \in T(k)} N_\sigma$, that is,

$$K(\alpha|n) > |\sigma| - k \text{ for all } n,$$

so α is LGC-random. □

For the converse we need:

Lemma 97. *Let $T \subseteq \mathbb{N} \times 2^{<\mathbb{N}}$ be a ML-test. Then there exists an ML-test $\hat{T} \subseteq \mathbb{N} \times 2^{<\mathbb{N}}$ such that for all k,*

(a)
$$\bigcup_{\sigma \in T(k)} N_\sigma = \bigcup_{\sigma \in \hat{T}(k)} N_\sigma,$$

(b) $\hat{T}(k) \subseteq 2^{<\mathbb{N}}$ is prefix-free.

Proof. Let $f : \mathbb{N} \to \mathbb{N} \times 2^{<\mathbb{N}}$ be recursive and such that $\mathrm{Im}(f) = T$. In addition, we let s, k range over $\mathbb{N}$ for the remainder of the proof. Now, we define, recursively, a sequence

$$\hat{T}_0 \subseteq \hat{T}_1 \subseteq \hat{T}_3 \subseteq \ldots$$

of finite subsets of $\mathbb{N} \times 2^{<\mathbb{N}}$ such that $\hat{T}_s(k)$ is prefix-free for all s, k. Take $\hat{T}_0 = \{f(0)\}$, and assume inductively that $\hat{T}_0 \subseteq \ldots \subseteq \hat{T}_s$ are finite subsets of $\mathbb{N} \times 2^{<\mathbb{N}}$ with $\hat{T}_s(k)$ prefix-free for all k. Then $\hat{T}_{s+1}$ is obtained as follows: let $f(s+1) = (k, \sigma)$, and let

$$(k, \tau_1), \ldots, (k, \tau_n)$$

be the distinct elements of $\hat{T}_s$ whose first component is k. Note that it is possible to find $\sigma_1, \ldots, \sigma_m$ such that

$$N_\sigma \setminus (N_{\tau_1} \sqcup \ldots \sqcup N_{\tau_n}) = N_{\sigma_1} \sqcup \ldots \sqcup N_{\sigma_m}$$

where the unions are disjoint, and then we take

$$\hat{T}_{s+1} = \hat{T}_s \cup \{(k, \sigma_1), \ldots, (k, \sigma_m)\} .$$

For $\hat{T}$ we take the union of this sequence of sets:

$$\hat{T} = \bigcup_{s \in \mathbb{N}} \hat{T}_s.$$

Then $\hat{T}$ is an ML-test satisfying (a) and (b). □

Note also the following property of ML-tests; if T is a ML-test and $\sigma \in T(n)$, then $2^{-|\sigma|} \leq 2^{-n}$, and hence $|\sigma| \geq n$. We now continue with the other direction of Schnorr's theorem.

Proof of ⇒ in Schnorr's theorem. We prove the contrapositive. Suppose α is not 1-random; take a ML-test T such that α fails T. That is,

$$\alpha \in \bigcap_n \bigcup_{\sigma \in T(n)} N_\sigma$$

By Lemma 97, we can assume that $T(k)$ is prefix-free for all k. Define

$$I : \ 2^{<\mathbb{N}} \rightharpoonup \mathbb{N} \ \text{ by } D(I) := \{\sigma : \sigma \in T(2n) \text{ for some } n \geq 1\}, \text{ and}$$
$$I(\sigma) := \min\{|\sigma| - n : \ \sigma \in T(2n) \text{ for some } n \geq 1\} \ \text{ for } \sigma \in D(I).$$

Then

$$\begin{aligned}\sum_{\sigma \in D(I)} 2^{-I(\sigma)} \quad &\leq \quad \sum_{n \geq 1} \sum_{\sigma \in T(2n)} 2^{-(|\sigma| - n)} \ \leq \ \sum_{n \geq 1} 2^n \sum_{\sigma \in T(2n)} 2^{-|\sigma|} \\ &\leq \quad \sum_{n \geq 1} 2^n 2^{-2n} \ = \ \sum_{n \geq 1} 2^{-n} \ = \ 1.\end{aligned}$$

Thus one of the two conditions for I to be be an icm is satisfied; we leave it as an exercise to check that the other condition is satisfied as well. So I is an icm.

By earlier results we have, for $\sigma \in D(I)$

$$K(\sigma) \ \leq \ I(\sigma) + \text{const}(I) \ = \ \min\left(\{|\sigma| - n : \sigma \in T(2n), n \geq 1\}\right) + \text{const}(I).$$

Now let $c \in \mathbb{N}$; it remains to show that $K(\alpha|k) < k - c$ for some k. Consider any $n \geq 1$. Because $\alpha \in \bigcup_{\sigma \in T(2n)} N_\sigma$, we can find $\sigma \in T(2n)$ such that σ is an initial segment of α. Set $k = |\sigma|$, so $\alpha|k = \sigma$, hence $k = |\sigma| \geq 2n$, $\sigma \in D(I)$, and

$$K(\alpha|k) \ = \ K(\sigma) \leq |\sigma| - n + \text{const}(I) \ = \ k - n + \text{const}(I) \ < \ k - c$$

provided $n > \text{const}(I) + c$. Now choose n to satisfy this inequality. □

3.6 Solovay Randomness

The definition of Solovay randomness proceeds from a corresponding notion of a Solovay test.

Definition 98. *A* Solovay test *is a recursively enumerable set* $S \subseteq 2^{<\mathbb{N}}$ *such that*

$$\sum_{\sigma \in S} 2^{-|\sigma|} < \infty.$$

Given a Solovay test $S \subseteq 2^{<\mathbb{N}}$ we say that α *passes* S if only finitely many initial segments of α lie in S. In other words, α fails S iff α has infinitely many *initial segments* in S, iff $\alpha \in N_\sigma$ for infinitely many $\sigma \in S$. We say that α is *Solovay random* if α passes all Solovay tests.

Theorem 99. α *is Solovay random* $\iff$ α *is 1-random.*

Proof. For the ($\Rightarrow$) direction, take a universal ML-test T and define

$$S := \bigcup_n T(n) \subseteq 2^{<\mathbb{N}}, \text{ so}$$

$$\sum_{\sigma \in S} 2^{-|\sigma|} \leq \sum_n \sum_{\sigma \in T(n)} 2^{-|\sigma|} \leq \sum_n 2^{-n} \leq 2.$$

Using the equivalence

$$\sigma \in S \iff \exists n\,((\sigma, n) \in T)$$

it is clear that S is recursively enumerable, so S is a Solovay test.

Let α be Solovay random. Then α passes S, so there are only finitely many initial segments of α in S. Take m such that $\alpha|n \notin S$ for all $n > m$. So for $n > m$ we have $\alpha \notin \bigcup_{\sigma \in T(n)} N_\sigma$, because $|\sigma| \geq n$ for all $\sigma \in T(n)$. Thus,

$$\alpha \notin \bigcap_n \bigcup_{\sigma \in T(n)} N_\sigma,$$

that is, α passes T.

As to the direction ($\Leftarrow$), let α be 1-random, and let S be a Solovay test. We are going to show that α passes S. Removing finitely many elements from S, we arrange that

$$\sum_{\sigma \in S} 2^{-|\sigma|} \leq 1.$$

Define $U_k \subseteq 2^{\mathbb{N}}$ by

$$U_k := \left\{\beta \mid \text{there are at least } 2^k \text{ many } \sigma \in S \text{ with } \beta \in N_\sigma\right\}.$$

Then clearly β passes S if and only if $\beta \notin \bigcap_k U_k$. Next, take $T \subseteq \mathbb{N} \times 2^{<\mathbb{N}}$ such that $(k, \sigma) \in T$ if and only if there are at least 2^k many initial segments of σ in

S. Then $\beta \in U_k$ is equivalent to there being at least 2^k many initial segments of β in S, and thus equivalent to $\beta \in \bigcup_{\sigma \in T(k)} N_\sigma$. So $U_k = \bigcup_{\sigma \in T(k)} N_\sigma$. Hence

$$\beta \text{ passes } S \iff \beta \notin \bigcap_k \bigcup_{\sigma \in T(k)} N_\sigma.$$

So, assuming T is a ML-test, we have that α passes T, which implies that α passes S, as promised.

To show that T is, indeed, a ML-test, we leave it as an exercise to check that T is recursively enumerable. It remains to show that $\mu(U_k) \leq 2^{-k}$ for every k. Take finite subsets S_t of S, with $t \in \mathbb{N}$, such that

$$S_0 \subseteq S_1 \subseteq S_2 \subseteq \dots, \text{ and } S = \bigcup_t S_t.$$

Put

$$U_k^t := \left\{\beta \mid \text{there are at least } 2^k \text{ many initial segments of } \beta \text{ in } S_t\right\},$$

so

$$U_k^0 \subseteq U_k^1 \subseteq U_k^2 \subseteq \dots, \text{ and } U_k = \bigcup_t U_k^t.$$

It suffices to show that for all t, k,

$$\mu(U_k^t) \leq 2^{-k}.$$

Fix t and let $\sigma_1, \dots, \sigma_n$ be the distinct elements of S_t. Then

$$\begin{aligned} &\int_\beta \left(\chi_{N_{\sigma_1}}(\beta) + \dots + \chi_{N_{\sigma_n}}(\beta)\right)\, d\mu(\beta) \\ \geq\ &\int_{\beta \in U_k^t} \left(\chi_{N_{\sigma_1}}(\beta) + \dots + \chi_{N_{\sigma_n}}(\beta)\right)\, d\mu(\beta) \ \geq\ 2^k \mu(U_k^t) \end{aligned}$$

and also,

$$\begin{aligned} \int_\beta \left(\chi_{N_{\sigma_1}}(\beta) + \dots + \chi_{N_{\sigma_n}}(\beta)\right)\, d\mu(\beta) \ &= \ \sum_{i=1}^n \int_\beta \chi_{N_{\sigma_i}}(\beta)\, d\mu(\beta) \\ &= \ \sum_{i=1}^n \mu(N_{\sigma_i}) \ = \ \sum_{i=1}^n 2^{-|\sigma_i|} \ \leq \ \sum_{\sigma \in S} 2^{-|\sigma|} \ \leq \ 1. \end{aligned}$$

Thus, $\mu(U_k^t) \leq 2^{-k}$ as required. □

As a consequence of the Theorem and the argument for the $(\Rightarrow)$ direction, the set $S = \bigcup_n T(n)$ with T a universal ML-test is a universal Solovay test in the sense that passing S is equivalent to passing all Solovay tests.

We use the following notation for the initial segment relation on $2^{<\mathbb{N}}$:

$$\rho \subseteq \sigma \ :\iff \ \rho \text{ is an initial segment of } \sigma$$

Theorem 100 (Miller and Yu)**.**

$$\alpha \textit{ is 1-random } \iff \sum_n 2^{n-K(\alpha|n)} < \infty$$

Proof. ($\Leftarrow$) We prove the contrapositive; assume α is not 1-random, so α is not LGC-random, hence $K(\alpha|n) < n$ for infinitely many n, and thus $n - K(\alpha|n) > 0$ for infinitely many n. Therefore,

$$\sum_n 2^{n-K(\alpha|n)} = \infty.$$

To prove the other ($\Rightarrow$) direction, first note that

$$\begin{aligned} \sum_{|\sigma|=m} \sum_{n \leq m} 2^{n-K(\sigma|n)} &= \sum_{|\sigma|=m} \sum_{\rho \subseteq \sigma} 2^{|\rho|-K(\rho)} \\ &= \sum_{|\rho| \leq m} 2^{m-|\rho|} \cdot 2^{|\rho|-K(\rho)} \ = \ 2^m \sum_{|\rho|\leq m} 2^{-K(\rho)} \ \leq \ 2^m \end{aligned}$$

where we use that K is an icm. Let $p \in \mathbb{N}^{\geq 1}$ in what follows. Then

$$\left| \left\{ \sigma \ : \ |\sigma| = m, \sum_{n \leq m} 2^{n-K(\sigma|n)} \geq p \right\} \right| \leq \frac{2^m}{p},$$

so for each m,

$$\mu\left(\left\{ \alpha \ : \ \sum_{n \leq m} 2^{n-K(\alpha|n)} \geq p \right\} \right) \leq \frac{1}{p}.$$

We introduce the following subsets of $2^{\mathbb{N}}$:

$$U_{p,m} := \left\{ \alpha \ : \ \sum_{n \leq m} 2^{n-K(\alpha|n)} \geq p \right\},$$

so

$$U_{p,0} \subseteq U_{p,1} \subseteq U_{p,2} \subseteq \dots$$

In addition, define $U_p := \bigcup_{m \in \mathbb{N}} U_{p,m}$; note that $\mu(U_p) \leq \frac{1}{p}$.

Now, let $T \subseteq \mathbb{N} \times 2^{<\mathbb{N}}$ be defined by the equivalence

$$T(k, \sigma) \ :\iff \sum_{n \leq |\sigma|} 2^{n-K(\sigma|n)} \geq 2^k.$$

Along the lines of the proof of Lemma 88 we see that T is recursively enumerable, and so to prove that T is a ML-test we have only to show that for all k

$$\mu\left(\bigcup_{\sigma \in T(k)} N_\sigma \right) \leq 2^{-k}.$$

To see why this holds, note the following equivalences

$$\begin{aligned} \alpha \in \bigcup_{\sigma \in T(k)} N_\sigma &\Leftrightarrow \exists m \, (\alpha|m \in T(k)) \\ &\Leftrightarrow \exists m \left(\sum_{n \le m} 2^{n-K(\alpha|n)} \ge 2^k \right) \\ &\Leftrightarrow \alpha \in U_{2^k}, \end{aligned}$$

and so $\bigcup_{\sigma \in T(k)} N_\sigma = U_{2^k}$. Hence

$$\mu \left(\bigcup_{\sigma \in T(k)} N_\sigma \right) \; = \; \mu \left(U_{2^k} \right) \; \le \; \frac{1}{2^k},$$

which proves that T is a ML-test. Assume now that α is 1-random. Then we get k such that $\alpha \notin U_{2^k}$, and so $\alpha \notin U_{2^k,m}$ for each m, that is, $\sum_{n \le m} 2^{n-K(\alpha|n)} < 2^k$ for each m, and thus $\sum_n 2^{n-K(\alpha|n)} \le 2^k < \infty$. □

We proved earlier that there are infinitely many σ such that

$$K(\sigma) > |\sigma| + \log |\sigma|.$$

The following corollary is much stronger.

Corollary 101. *Suppose $f : \mathbb{N} \to \mathbb{R}$ satisfies $\sum_n 2^{-f(n)} = \infty$. Given any 1-random α, there are infinitely many n such that $K(\alpha|n) > n + f(n)$.*

Proof. Let α be 1-random. If $K(\alpha|n) \le n + f(n)$ for almost all n, then also $n - K(\alpha|n) \ge -f(n)$ for almost all n. So

$$\sum_n 2^{n-K(\alpha|n)} \ge \sum_n 2^{-f(n)} + \text{const} = \infty,$$

which is a contradiction. □

As an example for the corollary just proved, let f take the value $f(n) = \log n$ for $n > 0$. Then

$$\sum_n 2^{-f(n)} \ge \sum_{n=1}^{\infty} \frac{1}{n} = \infty,$$

so if α is 1-random, then there are infinitely many n with $K(\alpha|n) > n + \log n$.

3.7 Martingales

A *martingale* is a function $f : 2^{<\mathbb{N}} \to [0, \infty)$ such that for all $\sigma \in 2^{<\mathbb{N}}$,

$$f(\sigma) \; = \; \frac{f(\sigma 0) + f(\sigma 1)}{2}$$

Observation 102. *If f and g are martingales, then so are $f+g$ and rf for $r \in [0,\infty)$. Note also that if f is a martingale with $f(\emptyset)=0$, then $f(\sigma)=0$ for all σ.*

Observation 103. *Let f be a martingale and $\sigma \leq \tau$; then $2^{-|\sigma|}f(\sigma) \geq 2^{-|\tau|}f(\tau)$, that is, $f(\sigma) \geq (1/2^{|\tau|-|\sigma|})f(\tau)$. For $\tau = \sigma 0$ and $\tau = \sigma 1$ the inequality reduces to $f(\sigma) \geq \frac{1}{2}f(\tau)$, which holds trivially. The general case follows by induction on $|\tau| - |\sigma|$.*

Theorem 104 (First Kolmogorov Inequality)**.** *Let f be a martingale, σ a string, and let $X \subseteq \{\tau : \sigma \leq \tau\}$ be prefix-free. Then*

$$2^{-|\sigma|}f(\sigma) \geq \sum_{\tau \in X} 2^{-|\tau|}f(\tau).$$

Proof. We can assume that X is finite. The theorem holds for $X = \emptyset$ and when $|X| = 1$. Let $n \geq 1$, assume inductively that the inequality holds for all X of size at most n, and suppose that $|X| = n+1$. Let $\pi \in 2^{<\mathbb{N}}$ be the unique longest common initial segment of the elements of X, so $\sigma \subseteq \pi$. For $i = 0, 1$, set $X_i := \{\tau \in X : \pi i \leq \tau\}$, and note that X_i has size at most n. So $X = X_0 \cup X_1$ and by induction we have

$$\begin{aligned}
\sum_{\tau \in X} 2^{-|\tau|}f(\tau) &= \sum_{\tau \in X_0} 2^{-|\tau|}f(\tau) + \sum_{\tau \in X_1} 2^{-|\tau|}f(\tau) \\
&\leq 2^{-|\pi 0|}f(\pi 0) + 2^{-|\pi 1|}f(\pi 1) \\
&= 2^{-|\pi|}\frac{f(\pi 0) + f(\pi 1)}{2} \\
&= 2^{-|\pi|}f(\pi) \geq 2^{-|\sigma|}f(\sigma),
\end{aligned}$$

which concludes the proof. □

Recall from the section on prefix-free sets that for each $Y \subseteq 2^{<\mathbb{N}}$ there is a prefix-free $X \subseteq Y$ such that

$$\bigcup_{\tau \in Y} N_\tau = \bigcup_{\tau \in X} N_\tau.$$

Theorem 105 (Second Kolmogorov Inequality)**.** *Let f be a martingale, $r \in (0,\infty)$, and put $S^r(f) := \{\tau : f(\tau) \geq r\}$. Then*

$$\mu\left(\bigcup_{\tau \in S^r(f)} N_\tau\right) \leq \frac{f(\emptyset)}{r}.$$

Proof. Take a prefix-free $X \subseteq S^r(f)$ such that

$$\bigcup_{\tau \in S^r(f)} N_\tau = \bigcup_{\tau \in X} N_\tau.$$

Then

$$\mu\left(\bigcup_{\tau\in S^r(f)} N_\tau\right) = \mu\left(\bigcup_{\tau\in X} N_\tau\right) = \sum_{\tau\in X} 2^{-|\tau|}.$$

Applying the first Kolmogorov inequality gives

$$r\mu\left(\bigcup_{\tau\in S^r(f)} N_\tau\right) = \sum_{\tau\in X} 2^{-|\tau|}r \leq \sum_{\tau\in X} 2^{-|\tau|}f(\tau) \leq f(\emptyset).$$

□

Definition 106. *A martingale f is said to be* ***effective*** *if there is a recursive $F : 2^{<\mathbb{N}} \times \mathbb{N} \to \mathbb{Q}^{\geq 0}$ such that, for each σ, we have*

$$F(\sigma,0) \leq F(\sigma,1) \leq F(\sigma,2) \leq \dots \quad \textit{and } F(\sigma,n) \to f(\sigma) \textit{ as } n \to \infty.$$

Here we code a nonnegative rational q as the unique pair $(a,b) \in \mathbb{N}^2$ such that a and b are coprime, $b \neq 0$ and $q = a/b$.

Definition 107. *A martingale f* ***succeeds*** *on α if $\limsup_{n\to\infty} f(\alpha|n) = \infty$, that is, $f(\alpha|n)$ takes arbitrarily large values.*

These definitions allow us to characterize 1-randomness as follows.

Theorem 108 (Schnorr)**.**

$$\alpha \textit{ is } 1\textit{-random} \iff \textit{no effective martingale succeeds on } \alpha.$$

Proof. We do the forward direction here; the backward direction is proved in a stronger form in the next proposition.

Let f be an effective martingale given by F. Take $q \in \mathbb{Q}^{>0}$ such that the martingale $g = qf$ satisfies $g(\emptyset) \leq 1$. Define $T \subseteq \mathbb{N} \times 2^{<\mathbb{N}}$ by

$$T(k,\sigma) \Leftrightarrow g(\sigma) > 2^k.$$

It is easy to see that T is recursively enumerable. In addition,

$$\mu\left(\bigcup_{\sigma\in T(k)} N_\sigma\right) \leq \mu\left(\bigcup_{\sigma\in S^{2^k}(g)} N_\sigma\right) \leq \frac{g(\emptyset)}{2^k} \leq \frac{1}{2^k},$$

so T is an ML-test. Now we have the following equivalences:

$$\begin{aligned}\alpha \text{ fails } T &\iff \alpha \in \bigcap_k \bigcup_{\sigma\in T(k)} N_\sigma \\ &\iff \text{for every } k \text{ there is } n \text{ with } g(\alpha|n) > 2^k \\ &\iff \limsup_{n\to\infty} g(\alpha|n) = \infty \\ &\iff \limsup_{n\to\infty} f(\alpha|n) = \infty.\end{aligned}$$

This concludes the proof of ⇒. For the converse, see below. □

Proposition 109. *Let T be an ML-test. Then there is an effective martingale d_T such that for all α, if α fails T, then $\lim_{k\to\infty} d_T(\alpha|k) = \infty$.*

Proof. By Lemma 97 we can assume that $T(n)$ is prefix-free for all n. Define the martingale d_σ by

$$d_\sigma(\tau) = \begin{cases} 1 & \text{if } \sigma \subseteq \tau \\ 2^{|\tau|-|\sigma|} & \text{if } \tau \subseteq \sigma \\ 0 & \text{if } \sigma, \tau \text{ are incomparable.} \end{cases}$$

Define $d_T : 2^{<\mathbb{N}} \to [0, \infty]$ by

$$d_T(\tau) = \sum_{\sigma \in \cup_n T(n)} d_\sigma(\tau).$$

Note that

$$d_T(\tau) = \frac{d_T(\tau 0) + d_T(\tau 1)}{2},$$

so d_T is a martingale if $d_T(\emptyset) \neq \infty$ (in which case all values of d_T are finite). We have

$$\begin{aligned} d_T(\emptyset) &= \sum_{\sigma \in \bigcup_n T(n)} d_\sigma(\emptyset) \\ &= \sum_{\sigma \in \bigcup_n T(n)} 2^{-|\sigma|} \\ &= \sum_n \sum_{\sigma \in T(n)} 2^{-|\sigma|} \\ &\leq \sum_n 2^{-n} \\ &= 2 < \infty, \end{aligned}$$

so d_T is a martingale. The map $(\sigma, \tau) \mapsto d_\sigma(\tau)$ into $\mathbb{Q}^{\geq 0}$ is recursive and $\bigcup_n T(n)$ is recursively enumerable, from which it follows easily that the infinite sum d_T of the d_σ is an effective martingale.

Suppose α fails T, that is,

$$\alpha \in \bigcap_n \bigcup_{\sigma \in T(n)} N_\sigma.$$

Then for all n there is $k \geq n$ with $\alpha|k \in T(n)$. We claim that, given any m, we have $d_T(\alpha|k) \geq m$ for all sufficiently large k. To see why this claim is true, let m be given. Beginnning with $n = 0$, take k_0 so that $\alpha|k_0 \in T(0)$. Then (using $n = k_0 + 1$) take $k_1 > k_0$ with $\alpha|k_1 \in T(k_0 + 1)$. Proceeding in this fashion, we get an increasing sequence $k_0 < k_1 < \ldots$ with $\alpha|k_i \in T(k_{i-1} + 1)$ for each i. Thus for $k \geq k_m + 1$ each $\alpha|k_i$ with $i < m$ contributes 1 to the sum forming $d_T(\alpha|k)$, so $d_T(\alpha|k) \geq m$ for all $k \geq k_m + 1$. This proves the claim. □

The $\Leftarrow$ direction of Schnorr's theorem now follows by considering the contrapositive.

Corollary 110. *There is a universal effective martingale d, that is, d is an effective martingale such that for all α the following are equivalent:*

(i) *α is 1-random,*

(ii) $\limsup_{n\to\infty} d(\alpha|n) < \infty$,

(iii) $\liminf_{n\to\infty} d(\alpha|n) < \infty$.

Proof. Let T be a universal ML-test, and let $d = d_T$ be the corresponding effective martingale. Then clearly (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) by Schnorr's theorem, and (iii) $\Rightarrow$ α passes T, by the proposition, giving (i). $\square$