Topics in Logic and Foundations: Spring 2005

Stephen G. Simpson

First Draft: April 29, 2005

This Draft: November 1, 2005

The latest version is available at

`http://www.math.psu.edu/simpson/notes/`.

Please send corrections to `<simpson@math.psu.edu>`.

This is a set of lecture notes from a 15-week graduate course at the Pennsylvania State University taught as Math 574 by Stephen G. Simpson in Spring 2005. The course was intended for students already familiar with the basics of mathematical logic. The course covered some topics which are important in contemporary mathematical logic and foundations but usually omitted from introductory courses.

These notes were typeset by the students in the course: John Ethier, Esteban Gomez-Riviere, David King, Carl Mummert, Michael Rowell, Chenying Wang. In addition, the notes were revised and polished by Stephen Simpson.

Contents

Chapter 1

Unsolvability of Hilbert's Tenth Problem

1.1 Hilbert's Tenth Problem

Definition 1.1.1 (Hilbert's Tenth Problem). Given a polynomial p with integer coefficients, to decide whether there exist integers $w_1, \ldots, w_n$ such that $p(w_1, \ldots, w_n) = 0$.

Definition 1.1.2. A *Diophantine equation* is an equation of the form

$$p(w_1, \ldots, w_n) = 0$$

where $p(w_1, \ldots, w_n)$ is a polynomial with *integer coefficients*, i.e., coefficients from $\mathbb{Z}$. Hilbert's Tenth Problem is: to find an algorithm for deciding whether a given Diophantine equation has an *integer solution*, i.e., $w_1, \ldots, w_n \in \mathbb{Z}$.

Hilbert proposed this problem in 1900. There was no progress until the 1950s, when M. Davis conjectured that Hilbert's Tenth Problem is unsolvable, i.e., no such algorithm exists. Davis, Putnam, and J. Robinson made further progress toward this result, and Matiyasevich completed the proof in 1969.

A typical method for showing that a problem P is unsolvable is to reduce the Halting Problem to P. Thus, a solution for P would give a solution to the Halting Problem, and as the Halting Problem is known to be unsolvable, P must then also be unsolvable. This is the method used here. We shall show that the Halting Problem is reducible to Hilbert's Tenth Problem.

The starting point for our presentation is the undecidability of true first-order arithmetic, T_1. Let the language L_1 consist of $\{+, \times, 0, 1, =\}$, where $+$ and $\times$ are binary operations, 0 and 1 are constants, and $=$ is a binary relation. The *terms* of L_1 are variables $x, y, z, \ldots$, the constants 0 and 1, and $t_1 + t_2$, $t_1 \times t_2$ where t_1, t_2 are terms. The *formulas* of L_1 are atomic formulas $t_1 = t_2$ where t_1, t_2 are terms, and $\neg A$, $A \vee B$, $A \wedge B$, $A \Rightarrow B$, $A \Leftrightarrow B$, $\exists x\, A$, $\forall x\, A$, where A, B are formulas and x is a variable. As usual, a sentence is a formula with no free variables.

Let $\mathbb{N} = \{0, 1, 2, \ldots\}$, the set of natural numbers. We also use $\mathbb{N}$ to denote the structure

$$(\mathbb{N}, +, \times, 0, 1, =),$$

i.e., the intended model of first-order arithmetic. Formulas of L_1 may be interpreted as usual in $\mathbb{N}$, and each sentence of L_1 is either true or false in $\mathbb{N}$. A theorem of Tarski says there is no algorithm to determine the truth value of an L_1-sentence in $\mathbb{N}$. T_1 is the complete theory consisting of all sentences of L_1 which are true in $\mathbb{N}$. Thus Tarski's result is that the theory T_1 is undecidable. Actually, Tarski shows that the Halting Problem H and many other noncomputable sets and functions are definable over $\mathbb{N}$, i.e., definable over T_1.

When interpreted in $\mathbb{N}$, terms of L_1 are equivalent to polynomials with positive integer coefficients. For example, the term $(x+y)\times((1+1)\times z+y)$ is equivalent over $\mathbb{N}$ to $2xz+xy+2yz+y^2$, which is a polynomial in $\mathbb{N}[x, y, z]$. Atomic formulas of L_1 are similarly equivalent to Diophantine equations: $p(x_1, \ldots, x_n) = q(x_1, \ldots, x_n)$ is equivalent to

$$p(x_1, \ldots, x_n) - q(x_1, \ldots, x_n) = 0,$$

and this is a typical Diophantine equation. Thus the existential sentence

$$\exists x_1 \cdots \exists x_n \; p(x_1, \ldots, x_n) = q(x_1, \ldots, x_n)$$

holds in $\mathbb{N}$ if and only if the Diophantine equation $p(x_1, \ldots, x_n)-q(x_1, \ldots, x_n) = 0$ has at least one solution in $\mathbb{N}$.

Accordingly, we consider a modified form of Hilbert's Tenth Problem.

Definition 1.1.3 (Modified Hilbert's Tenth Problem). Given a polynomial $p(x_1, \ldots, x_n)$ with coefficients from $\mathbb{Z}$, to decide whether there exist $x_1, \ldots, x_n \in \mathbb{N}$ such that $p(x_1, \ldots, x_n) = 0$.

Remark 1.1.4. The Modified Hilbert's Tenth Problem is equivalent to the original problem. Suppose first that the Modified Hilbert's Tenth Problem were solvable. Then the Diophantine equation $p(w_1, \ldots, w_n) = 0$ has integer solutions if and only if $\exists x_1 \cdots \exists x_n \in \mathbb{N}$ such that $p(\pm x_1, \ldots, \pm x_n) = 0$, so Hilbert's Tenth Problem would be solvable. Conversely, if Hilbert's Tenth Problem were solvable, then $p(x_1, \ldots, x_n) = 0$ has natural number solutions if and only if $p(t_1^2+u_1^2+v_1^2+w_1^2, \ldots, t_n^2+u_n^2+v_n^2+w_n^2) = 0$ has integer solutions, so the Modified Hilbert's Tenth Problem would also be solvable. This relies on Lagrange's Theorem: every natural number is the sum of four squares.

Note that Tarski's Theorem and the Modified Hilbert's Tenth Problem both deal with different kinds of definability over $\mathbb{N}$. We use the proof of Tarski's Theorem (see our Math 558 notes [14]) as the starting point for our proof of unsolvability of the Modified Hilbert's Tenth Problem.

1.2 Σ_1 Relations and Functions

To warm up, we consider yet another kind of definability over $\mathbb{N}$.

Definition 1.2.1 (Δ_0 formulas). The Δ_0 formulas of L_1 are the smallest class of formulas closed under propositional connectives ($\wedge$, $\vee$, $\neg$, $\Rightarrow$, $\Leftrightarrow$) and bounded quantification ($\forall x < t$, $\exists x < t$, where t is a term not mentioning x).

Definition 1.2.2 (Δ_0 relations and functions). A relation $R \subseteq \mathbb{N}^k$ is Δ_0 if it is definable by a Δ_0 formula. A partial function ψ from $\mathbb{N}^k$ to $\mathbb{N}$ is Δ_0 if graph(ψ) is Δ_0.

Example 1.2.3. The "less than" relation $x < y$ is definable by the Δ_0 formula $\exists z < y\,(x + z + 1 = y)$.

Remark 1.2.4. The Δ_0 relations are only a small subclass of the primitive recursive relations. Nevertheless, many interesting relations are Δ_0. E.g., a result of Bennett shows that the 3-place exponential relation $x^y = z$ is Δ_0. We omit the proof.

Definition 1.2.5 (Σ_1 formulas). A formula G is Σ_1 it is of the form $\exists x\, F$ where F is Δ_0.

Definition 1.2.6 (Σ_1 relations and functions). A relation $R \subseteq \mathbb{N}^k$ is Σ_1 if it is definable over $\mathbb{N}$ by a Σ_1 formula. A partial function ψ from $\mathbb{N}^k$ to $\mathbb{N}$ is Σ_1 if graph(ψ) is Σ_1.

We shall prove the following theorem.

Theorem 1.2.7. *R is Σ_1 if and only if R is recursively enumerable, i.e., Σ^0_1. ψ is Σ_1 if and only if ψ is partial recursive.*

The forward direction of the theorem is obvious, as Σ_1 relations are clearly Σ^0_1, and Σ_1 partial functions are clearly partial recursive. (See my Math 558 notes [14].) We must show the converse direction. In particular, we must show that all primitive recursive functions are Σ_1.

Lemma 1.2.8. *The class of Σ_1 relations is closed under unbounded existential quantification, logical and, logical or, and bounded quantification.*

Proof. Suppose G is Σ_1. Then $\exists x\, G$ is equivalent to $\exists x\, \exists y\, F$, where F is Δ_0. This is then equivalent to $\exists z\, \exists x < z\, \exists y < z\, F$ which is Σ_1.

If $\exists x\, F$ and $\exists x\, G$ are both Σ_1, then $\exists x\, F \wedge \exists x\, G \equiv \exists x\, \exists y\, (F \wedge G)$. $F \wedge G$ is Δ_0 and so the formula is Σ_1. The case for disjunction is similar.

Also, $\exists x < t\, \exists y\, F \equiv \exists y\, \exists x < t\, F$ and so the class of Σ_1 relations is closed under bounded existential quantification.

We have $\forall x < t\, \exists y\, F \equiv \exists z\, \forall x < t\, \exists y < z\, F$. The formula $\forall x < t\, \exists y < z\, F$ is Δ_0 and thus the whole formula is Σ_1 as required. Thus the class of Σ_1 relations is closed under bounded universal quantification. □

To finish the proof of Theorem 1.2.7, we now briefly review Gödel's β function. The β function is a method of coding arbitrarily long finite sequences of integers in an arithmetically effective way.

Lemma 1.2.9. For all k there exist infinitely many a such that

$$a+1,\, 2a+1,\, \ldots,\, ka+1$$

are pairwise relatively prime.

Proof. Let a be any muliple of $k!$. If $ia+1$ and $ja+1$ are not relatively prime, $1 \leq i < j \leq k$, let p be a prime dividing both $ia+1$ and $ja+1$. In particular p does not divide a. Thus $p > k$ by our choice of a. On the other hand, p divides $(ja+1)-(ia+1) = (j-i)a$, so p divides $j-i$. This contradicts $p > k$. □

The following is a well known result in number theory. We omit its proof. See the Math 558 notes [14].

Lemma 1.2.10 (Chinese Remainder Theorem). Let $m_1, \ldots, m_k$ be pairwise relatively prime. Given $r_1, \ldots, r_k$ such that $0 \leq r_i < m_i$ for $i = 1, \ldots, k$, we can find r such that $r \equiv r_i \bmod m_i$ for all $i = 1, \ldots, k$.

Definition 1.2.11 (the β function). We define

$$\beta(a, r, i) = \mathrm{Rem}(r, a \cdot (i+1) + 1)$$

where $\mathrm{Rem}(y, x)$ is the remainder of y on division by x.

Corollary 1.2.12. Given $r_0, \ldots, r_k \geq 0$, we can find $a, r \geq 0$ such that $\beta(a, r, i) = r_i$ for all $i = 0, \ldots, k$.

Proof. By Lemma 1.2.9 above, let a be such that $a+1, 2a+1, \ldots, (k+1)a+1$ are pairwise relatively prime, and $a > \max(r_0, \ldots, r_n)$. By the Chinese Remainder Theorem, we can find r such that $r \equiv r_i \bmod a(i+1)+1$ for $i = 0, \ldots, k$. Thus $\beta(a, r, i) = r_i$ for $i = 0, \ldots, k$. □

Lemma 1.2.13. The β function is Σ_1.

Proof. It suffices to show that Rem is Σ_1. We have

$$\mathrm{Rem}(y, x) = r \iff r < x \wedge \exists q < y \, (y = qx + r) \, .$$

Thus Rem and the β function are Δ_0, hence Σ_1. □

Lemma 1.2.14. All primitive recursive functions are Σ_1.

Proof. $Z(x) = 0$ is Σ_1 via $y = 0$.
$S(x) = x+1$ is Σ_1 via $y = x+1$.
$P_{ki}(x_1, \ldots, x_n) = x_i$ is Σ_1 via $y = x_i$.

Given $f(x_1,\ldots,x_n) = h(g_1(x_1,\ldots,x_n),\ldots,g_m(x_1,\ldots,x_n)))$, where $h, g_1,\ldots,g_m$ are Σ_1, we have that f is Σ_1, because

$$y = f(x_1,\ldots,x_n) \iff \exists z_1 \cdots \exists z_m \left(y = h(z_1,\ldots,z_m)) \wedge \bigwedge_{i=1}^{m} z_i = g_i(x_1,\ldots,x_n) \right).$$

Thus the class of Σ_1 functions is closed under composition.

Given $f(x_1,\ldots,x_n)$ defined by

$$\begin{aligned} f(0,x_1,\ldots,x_n) &= g(x_1,\ldots,x_n) \\ f((x+1,x_1,\ldots,x_n) &= h(x,f(x,x_1,\ldots,x_n),x_1,\ldots,x_n) \end{aligned}$$

where g, h are Σ_1, f is Σ_1 because

$$\begin{aligned} y = f(x,x_1,\ldots,x_n) &\iff \exists\langle y_0,y_1,\ldots,y_x\rangle\,(y_0 = g(x_1,\ldots,x_n) \wedge \\ &\qquad (\forall i < x)\, y_{i+1} = h(i,y_i,x_1,\ldots,x_n)) \\ &\iff \exists a\,\exists r\,(\beta(a,r,0) = g(x_1,\ldots,x_n) \wedge \beta(a,r,x) = y \wedge \\ &\qquad (\forall i < x)\,\beta(a,r,i+1) = h(i,\beta(a,r,i),x_1,\ldots,x_n))\,. \end{aligned}$$

Thus the class of Σ_1 functions is closed under primitive recursion.

It now follows that all primitive recursive functions are Σ_1. □

We can now prove:

Theorem 1.2.15. If $\psi : \mathbb{N}^k \xrightarrow{P} \mathbb{N}$ is partial recursive, then ψ is Σ_1.

Proof. Let e be an index of ψ, i.e., the Gödel number of a program which computes ψ. Then $\psi = \varphi_e^{(k)}$, i.e., $\psi(x_1,\ldots,x_k) \simeq y \iff \varphi_e^{(k)}(x_1,\ldots,x_k) \simeq y \iff \exists n\,(\mathrm{State}(e,x_1,\ldots,x_k,n))_0 = 0 \wedge (\mathrm{State}(e,x_1,\ldots,x_k,n))_{k+1} = y)$, where $(\mathrm{State}(e,x_1,\ldots,x_k,n))_0$ and $(\mathrm{State}(e,x_1,\ldots,x_k,n))_{k+1}$ are primitive recursive functions (see Math 558 notes [14]). Thus ψ is Σ_1. □

The proof of Theorem 1.2.7 is now complete.

Corollary 1.2.16. The Halting Problem H is Σ_1.

1.3 Diophantine Relations and Functions

Definition 1.3.1. A relation $R \subseteq \mathbb{N}^k$ is said to be *Diophantine* if there exists a polynomial $p(x_1,\ldots,x_k,y_1,\ldots,y_n)$ with coefficients from $\mathbb{Z}$, such that

$$R = \{\langle x_1,\ldots,x_k\rangle \in \mathbb{N}^k \mid \exists y_1 \cdots \exists y_n\, p(x_1,\ldots,x_k,y_1,\ldots,y_n) = 0\}\,.$$

Here $y_1,\ldots,y_n$ range over $\mathbb{N}$. A partial function ψ is said to be *Diophantine* if $\mathrm{graph}(\psi)$ is Diophantine.

The following theorem is due to Matiyasevich 1969. It is known as Matiyasevich's Theorem, or as the MDRP Theorem (standing for Matiyasevich, Davis, Robinson, Putnam).

Theorem 1.3.2 (MDRP Theorem). R is Diophantine $\iff$ R is Σ_1. ψ is Diophantine $\iff$ ψ is partial recursive.

Corollary 1.3.3. The Halting Problem $H = \{e \mid \varphi_e^{(1)}(0)\downarrow\} \subseteq \mathbb{N}$ is Diophantine.

Corollary 1.3.4. Hilbert's Tenth Problem is unsolvable.

So, our goal now is to prove the MDRP Theorem.

Note that the forward direction of the MDRP Theorem is obvious, as ψ Diophantine implies ψ Σ_1, which implies ψ partial recursive. For the converse, we must show that all partial recursive functions are Diophantine.

By Theorem 1.2.7, it suffices to show that all Σ_1 functions are Diophantine. We begin with the following easy lemma.

Lemma 1.3.5. The binary relation $<$ is Diophantine. The class of Diophantine relations is closed under unbounded existential quantification, logical and, logical or, and bounded existential quantification.

Proof. Clearly $<$ is Diophantine, since $x < y \iff \exists z\,(x + z + 1 = y)$.

If $R(x_1, \ldots, x_k, y) \equiv \exists \overline{z}\, p(x_1, \ldots, x_k, y, \overline{z}) = 0$ is Diophantine, then so is $\exists y\, R(x_1, \ldots, x_k, y) \equiv \exists y\, \exists \overline{z}\, p(x_1, \ldots, x_k, y, \overline{z}) = 0$, so trivially the class of Diophantine relations is closed under unbounded existential quantification.

Suppose $R_1 = \{\langle x_1, \ldots, x_k\rangle \in \mathbb{N}^k \mid \exists \overline{y}\, p(x_1, \ldots, x_k, \overline{y}) = 0\}$ and $R_2 = \{\langle x_1, \ldots, x_k\rangle \in \mathbb{N}^k \mid \exists \overline{z}\, q(x_1, \ldots, x_k, \overline{z}) = 0\}$ are both Diophantine. We then have

$$\begin{aligned} \exists \overline{y}\, p(x_1, \ldots, x_k, \overline{y}) = 0 \wedge \exists \overline{z}\, q(x_1, \ldots, x_k, \overline{z}) = 0 \quad &\iff \\ \exists \overline{y}\, \exists \overline{z}\, (p(x_1, \ldots, x_k, \overline{y}) = 0 \wedge q(x_1, \ldots, x_k, \overline{z}) = 0) \quad &\iff \\ \exists \overline{y}\, \exists \overline{z}\, (p(x_1, \ldots, x_k, \overline{y})^2 + q(x_1, \ldots, x_k, \overline{z})^2 = 0) \quad & \end{aligned}$$

so $R_1 \wedge R_2$ is Diophantine. Thus the class of Diophantine relations is closed under logical and.

Similarly, for logical or, we have

$$\begin{aligned} \exists \overline{y}\, p(x_1, \ldots, x_k, \overline{y}) = 0 \vee \exists \overline{z}\, q(x_1, \ldots, x_k, \overline{z}) = 0 \quad &\iff \\ \exists \overline{y}\, \exists \overline{z}\, (p(x_1, \ldots, x_k, \overline{y}) = 0 \vee q(x_1, \ldots, x_k, \overline{z}) = 0) \quad &\iff \\ \exists \overline{y}\, \exists \overline{z}\, p(x_1, \ldots, x_k, \overline{y}) \cdot q(x_1, \ldots, x_k, \overline{z}) = 0 \quad & \end{aligned}$$

so $R_1 \vee R_2$ is Diophantine. Thus the class of Diophantine relations is closed under logical or.

We also have $(\exists x < t)\, \exists \overline{y}\, p(x, x_1, \ldots, x_n, \overline{y}) = 0$ if and only if $\exists x\, (x < t \wedge \exists \overline{y}\, p(x, x_1, \ldots, x_n, \overline{y}) = 0)$. Thus the class of Diophantine relations is closed under bounded existential quantification. $\square$

In addition, we have the following easy lemma.

Lemma 1.3.6. Addition, multiplication, and the functions Quot and Rem given by

$$y = qx + r, \quad r < x, \quad \mathrm{Quot}(y, x) = q, \quad \mathrm{Rem}(y, x) = r$$

as well as the Gödel β function are Diophantine. The class of Diophantine functions is closed under composition.

Proof. Trivially $+$ and $\cdot$ are Diophantine. We have $\mathrm{Quot}(y, x) = q \iff \exists r\,(r < x \wedge y = qx + r)$, so Quot is Diophantine, and similarly for Rem. Closure under composition is easy, as in the proof of Lemma 1.2.14. It now follows that β is Diophantine. $\square$

By Lemma 1.3.5, to prove the MDRP Theorem, it remains only to show that the class of Diophantine relations is closed under bounded universal quantification. This is the hard part of the proof. Note that bounded universal quantification was crucial in the proof of Lemma 1.2.14.

We shall follow the exposition of Davis [5]. Most of the work is contained in the following lemma.

Lemma 1.3.7 (Main Lemma). The following functions are Diophantine.

1. $(n, k) \mapsto n^k$
2. $(n, k) \mapsto \binom{n}{k}$
3. $n \mapsto n!$
4. $(a, b, k) \mapsto \prod_{i=0}^{k}(a + bi)$

The proof of the Main Lemma is difficult, and we postpone it to Section 1.6 below.

1.4 Bounded Universal Quantification

Our goal is to show that if R is Σ_1 then R is Diophantine. As we have already seen, it suffices to prove that the class of Diophantine relations is closed under bounded universal quantification. Here is a flawed attempt at a proof of this.

Flawed Proof. We attempt to imitate the proof of Lemma 1.2.14 using the idea of coding via Gödel's β function. Assume that

$$(\forall i)_{1 \leq i \leq k}\ \exists y_1 \cdots \exists y_n\ p(k, i, \ldots, y_1, \ldots, y_n) = 0.$$

For each $1 \leq i \leq k$ pick witnesses $y_1^{(i)}, \ldots, y_n^{(i)}$ such that $p(k, i, \ldots, y_1^{(i)}, \ldots, y_n^{(i)}) = 0$. Let u be an upper bound for k and $y_j^{(i)}$, $1 \leq i \leq k$, $1 \leq j \leq n$. Let t be any multiple of $u!$. By the proof of Lemma 1.2.9, the moduli $t + 1, \ldots, kt + 1$ are pairwise relatively prime. By the Chinese Remainder Theorem 1.2.10, we can find $r_1, \ldots, r_n$ such that $r_j \equiv y_j^{(i)} \bmod it+1$ for all $1 \le i \le k$, $1 \le j \le n$. Hence for $1 \le i \le k$ we have

$$p(k, i, \ldots, r_1, \ldots, r_n) \equiv 0 \mod it+1.$$

Form the product $\prod_{i=1}^{k}(it+1) = ct+1$. We have $0 \equiv it+1 \equiv ct+1 \bmod it+1$. Multiplying by c and i respectively, we have $0 \equiv cit+c \equiv cit+i \bmod it+1$, which implies $c \equiv i \bmod it+1$. It follows that $p(k, c, \ldots, r_1, \ldots, r_n) \equiv 0 \bmod it+1$ for all i, $1 \le i \le k$. Since the $it+1$, $1 \le i \le k$ are pairwise relatively prime, we have

$$\begin{aligned} p(k, c, \ldots, r_1, \ldots, r_n) &\equiv 0 \mod \prod_{i=1}^{k}(it+1) \\ &\equiv 0 \mod ct+1. \end{aligned}$$

The upshot is that we have "packaged" all of our equations for $1 \le i \le k$ into one equation. But our problem is that it is only modulo $ct+1$.

Conversely, assume t is a multiple of $u!$, $u \ge k$, $ct+1 = \prod_{i=1}^{k}(it+1)$ and $\exists r_1 \cdots \exists r_n \, p(k, c, \ldots, r_1, \ldots, r_n) \equiv 0 \bmod ct+1$. As before we have $c \equiv i \bmod it+1$ for each $1 \le i \le k$. Let $y_j^{(i)} = \mathrm{Rem}(r_j, it+1)$. Then $r_j \equiv y_j^{(i)} \bmod it+1$, hence $p(k, i, \ldots, y_1^{(i)}, \ldots, y_n^{(i)}) \equiv 0 \bmod it+1$. If we knew that

$$|p(k, i, \ldots, y_1^{(i)}, \ldots, y_n^{(i)})| < it+1,$$

we could conclude

$$p(k, i, \ldots, y_1^{(i)}, \ldots, y_n^{(i)}) = 0$$

and we would be finished. □

In order to repair this flawed argument, we first present a simple lemma, Lemma 1.4.1. After that, the proof of closure under bounded universal quantification is given by Lemma 1.4.2.

Lemma 1.4.1. Given a polynomial $p(k, i, \ldots, y_1, \ldots, y_n)$ we can find a polynomial $q(k, \ldots, u)$ such that

1. $q(k, \ldots, u) \ge u$
2. $q(k, \ldots, u) \ge k$
3. $q(k, \ldots, u) \ge |p(k, i, \ldots, y_1, \ldots, y_n)|$ for all $i \le k$ and $y_1, \ldots, y_n \le u$.

Proof. Let $q(k, \ldots, u) = |p|(k, k, \ldots, u, \ldots, u) + u + k$ where $|p|$ is just p with all coefficients replaced by their absolute values. □

Lemma 1.4.2. $(\forall i)_{1 \le i \le k} \exists y_1 \cdots \exists y_n \, p(k, i, \ldots, y_1, \ldots, y_n) = 0$ if and only if there exist $u, t, c, r_1, \ldots, r_n$ such that:

1. $t = q(k, \ldots, u)!$,

2. $ct+1 = \prod_{i=1}^{k}(it+1)$ divides each of $\prod_{y=0}^{u}(r_j - y)$, $1 \le j \le n$,

3. $p(k, c, \dots, r_1, \dots, r_n) \equiv 0 \mod ct+1$.

The point of this lemma is that, by 1.3.7, the right-hand side is Diophantine. Thus we see that the class of Diophantine relations is closed under bounded universal quantification.

Proof. $\Rightarrow$: As before, we can find $u, t, c, r_1, \dots, r_n$ such that

$$p(k, c, \dots, r_1, \dots, r_n) \equiv 0 \mod ct+1$$

and $r_j \equiv y_j^{(i)} \mod it+1$. Thus $it+1$ divides $r_j - y_j^{(i)}$. Since $y_j^{(i)} \le u$, $it+1$ divides $\prod_{y=0}^{u}(r_j - y)$. Since $it+1$, $1 \le i \le k$ are pairwise relatively prime, it follows that $ct+1$ divides $\prod_{y=0}^{u}(r_j - y)$ for $1 \le j \le n$, as required.

$\Leftarrow$: For each $1 \le i \le k$ pick a prime divisor p_i of $it+1$. Since $t = q(k, \dots, u)!$, we have $p_i > q(k, \dots, u)$. Let $y_j^{(i)} = \mathrm{Rem}(r_j, p_i)$. Note that $y_j^{(i)} < p_i$. We claim $y_j^{(i)} \le u$. To see this, note that p_i divides $it+1$ which divides $ct+1$ which divides $\prod_{y=0}^{u}(r_j - y)$, hence p_i divides $r_j - y$ for some $y \le u$. Then $y \equiv r_j \equiv y_j^{(i)} \mod p_i$. Noting also that $y \le u \le q(k, \dots, u) < p_i$, we see that $y = y_j^{(i)}$. Therefore $y_j^{(i)} \le u$.

Next we claim that $p(k, i, \dots, y_1^{(i)}, \dots, y_n^{(i)}) = 0$ for $1 \le i \le k$. By assumption we have $p(k, c, \dots, r_1, \dots, r_n) \equiv 0 \mod ct+1$. Recall that $ct+1 = \prod_{i=1}^{k} it+1$ and $c \equiv i \mod it+1$. Therefore $c \equiv i \mod p_i$. Moreover $r_i \equiv y_j^{(i)} \mod p_i$, so

$$\begin{aligned} p(k, i, \dots, y_1^{(i)}, \dots, y_n^{(i)}) &\equiv 0 \mod ct+1 \\ &\equiv 0 \mod it+1 \\ &\equiv 0 \mod p_i. \end{aligned}$$

Since $y_1^{(i)}, \dots, y_n^{(i)} \le u$, we have $|p(k, i, \dots, y_1^{(i)}, \dots, y_n^{(i)})| \le q(k, \dots, u) < p_i$. Hence $p(k, i, \dots, y_1^{(i)}, \dots, y_n^{(i)}) = 0$ and our lemma is proved. $\square$

Lemma 1.4.2 shows that the class of Diophantine relations is closed under bounded universal quantification. This completes the proof of the MDRP Theorem 1.3.2, except that it remains to prove the Main Lemma.

1.5 The Pell Equation

The Main Lemma 1.3.7 asserts that the exponential function $(n, k) \mapsto n^k$ and similar functions are Diophantine. In order to prove this, we need a Diophantine function which is of exponential growth. It turns out that the solutions of a particular Diophantine equation known as Pell's equation not only grow exponentially but also are convenient in other ways. Following Davis ([5], reprinted in [6, Appendix]), we give a self-contained, elementary presentation of all of the number theory which we shall use.

1.5.1 Basic Properties

We begin with basic properties of the Pell equation.

Definition 1.5.1 (the Pell equation). A *Pell equation* is an equation of the form $x^2 - dy^2 = 1$ where $d = a^2 - 1$, $a \geq 2$, $a \in \mathbb{N}$.

Examples 1.5.2.

1. $a = 2$, $x^2 - 3y^2 = 1$.
2. $a = 3$, $x^2 - 8y^2 = 1$.
3. $a = 4$, $x^2 - 15y^2 = 1$.

Remark 1.5.3. If (x, y) is any integer solution of the Pell equation, then clearly

$$(x + y\sqrt{d})(x - y\sqrt{d}) = x^2 - dy^2 = 1.$$

Furthermore, (x, y) is a solution if and only if $(|x|, |y|)$ is a solution, so we may focus on solutions with $x, y \geq 0$. In this case we have $x+y\sqrt{d} \geq 1$, with equality only if $(x, y) = (1, 0)$.

Remark 1.5.4. There are two obvious solutions of the Pell equation, $(1, 0)$ and $(a, 1)$. Moreover, there is an easy way of generating more solutions, as follows.

Lemma 1.5.5. If (x, y) and (x', y') are integer solutions of the Pell equation, then so is (x'', y'') given by

$$x'' + y''\sqrt{d} = (x + y\sqrt{d})(x' + y'\sqrt{d}).$$

Proof. Taking conjugates, we have

$$x'' - y''\sqrt{d} = (x - y\sqrt{d})(x' - y'\sqrt{d}).$$

Multiplying the two equations, we get

$$x''^2 - dy''^2 = (x^2 - dy^2)(x'^2 - dy'^2) = 1$$

and our lemma is proved. □

We shall now show that all solutions are generated in this way.

Definition 1.5.6. For $n \geq 0$ we define $x_n(a)$ and $y_n(a)$ by

$$x_n(a) + y_n(a)\sqrt{d} \;=\; (a + \sqrt{d})^n\,.$$

By Lemma 1.5.5, $(x_n(a), y_n(a))$ is a solution of Pell's equation. When a is fixed, we write $x_n = x_n(a)$ and $y_n = y_n(a)$.

Theorem 1.5.7. All natural number solutions of Pell's equation are of the form (x_n, y_n) for some n.

Proof. Otherwise there would be a solution (x, y) with

$$x_n + y_n\sqrt{d} < x + y\sqrt{d} < x_{n+1} + y_{n+1}\sqrt{d}\,.$$

By the above definition, this becomes

$$(a + \sqrt{d})^n < x + y\sqrt{d} < (a + \sqrt{d})^{n+1}.$$

Dividing gives

$$1 < \frac{x + y\sqrt{d}}{x_n + y_n\sqrt{d}} < a + \sqrt{d}$$

which simplifies to

$$1 < (x + y\sqrt{d})(x_n - y_n\sqrt{d}) < a + \sqrt{d}\,.$$

Multiplying the solutions as in Lemma 1.5.5 gives

$$1 < x' + y'\sqrt{d} < a + \sqrt{d}.$$

Taking negative reciprocals, we get

$$-1 < -x' + y'\sqrt{d} < -a + \sqrt{d}.$$

Adding, we get $0 < 2y'\sqrt{d} < 2\sqrt{d}$, which implies $0 < y' < 1$, a contradiction. $\square$

We now obtain recurrences and explicit formulas for x_n and y_n.

Lemma 1.5.8. We have

$$\begin{aligned} x_{n\pm m} &= x_n x_m \pm d y_n y_m, \\ y_{n\pm m} &= x_m y_n \pm x_n y_m. \end{aligned}$$

Proof. Note that

$$\begin{aligned} x_{n\pm m} + y_{n\pm m}\sqrt{d} &= (a + \sqrt{d})^{n+m} \\ &= (x_m \pm y_m\sqrt{d})(x_n \pm y_n\sqrt{d}) \\ &= (x_n x_m \pm d y_n y_m) + (x_n y_m \pm x_m y_n)\sqrt{d}. \end{aligned}$$

$\square$

Remark 1.5.9. In the special case $m = 1$, the previous lemma says

$$\begin{aligned} x_{n\pm1} &= a x_n \pm d y_n, \\ y_{n\pm1} &= a y_n \pm x_n. \end{aligned}$$

Adding these expressions for $x_{n\pm1}$ and $y_{n\pm1}$ respectively, we get recurrences

$$\begin{aligned} x_{n+1} &= 2a x_n - x_{n-1}, \\ y_{n+1} &= 2a y_n - y_{n-1}. \end{aligned}$$

Theorem 1.5.10. We have the following explicit formulas:

$$x_n = \left\lceil \frac{1}{2}(a+\sqrt{d})^n \right\rceil,$$

$$y_n = \left\lfloor \frac{1}{2\sqrt{d}}(a+\sqrt{d})^n \right\rfloor.$$

Proof. To get an explicit formula for x_n, we solve the recurrence $x_{n+1} = 2ax_n - x_{n-1}$. Setting $x_n = z^n$ we get $z^{n+1} = 2az^n - z^{n-1}$. Dividing by z^{n-1} we get the quadratic equation $z^2 = 2az - 1$ which has solutions $z = a \pm \sqrt{d}$. Thus $x_n = A(a+\sqrt{d})^n + B(a-\sqrt{d})^n$. Using our initial conditions $x_0 = 1 = A + B$ and $x_1 = a = A(a+\sqrt{d}) + B(a-\sqrt{d})$, we get $A = B = 1/2$. Thus $x_n = (1/2)((a+\sqrt{d})^n + (a-\sqrt{d})^n) = \lceil (1/2)(a+\sqrt{d})^n \rceil$.

Similarly, to get an explicit formula for y_n, we have $y_n = A(a+\sqrt{d})^n + B(a-\sqrt{d})^n$, but this time our initial conditions are $y_0 = 0 = A + B$ and $y_1 = 1 = A(a+\sqrt{d}) + B(a-\sqrt{d})$. These equations yield $A = 1/2\sqrt{d}$ and $B = -1/2\sqrt{d}$. Thus $y_n = (1/2\sqrt{d})((a+\sqrt{d})^n + (a-\sqrt{d})^n) = \lfloor (1/2\sqrt{d})(a+\sqrt{d})^n \rfloor$. □

1.5.2 Divisibility Properties of y_n

We now obtain some divisibility properties of y_n.

Theorem 1.5.11. $\mathrm{GCD}(x_n, y_n) = 1$.

Proof. Let p be a prime dividing x_n and y_n. Then p divides $x_n^2 - dy_n^2 = 1$, a contradiction. □

Lemma 1.5.12. $y_n \mid y_t$ if and only if $n \mid t$.

Proof. Assume $n \mid t$ and let $t = nk$. We prove $y_n \mid y_{nk}$ by induction on k. For $k = 0$ we have $y_n \mid 0 = y_0$, and for $k = 1$ we have $y_n \mid y_n$. Now $y_{n(k+1)} = y_{nk+n} = x_n y_{nk} + x_{nk} y_n$, and by induction hypothesis $y_n \mid y_{nk}$, hence $y_n \mid y_{n(k+1)}$.

Conversely, assume $y_n \mid y_t$. Let $t = qn + r$ with $0 \le r < n$. We then have $y_t = y_{qn+r} = x_r y_{qn} + x_{qn} y_r$. Since y_n divides y_{qn}, it follows that y_n divides $x_{qn} y_r$. But since $\mathrm{GCD}(y_{qn}, x_{qn}) = 1$, we have $\mathrm{GCD}(y_n, x_{qn}) = 1$. Thus y_n divides y_r, but since $r < n$ we have $y_r < y_n$. Hence $r = 0$, so $n \mid t$. □

Theorem 1.5.13. $y_n^2 \mid y_t$ if and only if $ny_n \mid t$.

Proof. Note that

$$x_{nk} + y_{nk}\sqrt{d} = (a+\sqrt{d})^{nk} = (x_n + y_n\sqrt{d})^k = \sum_{i=0}^{k} \binom{k}{i} x_n^{k-i} y_n^i d^{i/2}.$$

Comparing coefficients of $\sqrt{d}$, we see that

$$y_{nk} = \sum_{\substack{0 \le i \le k \\ i \text{ odd}}} \binom{k}{i} x_n^{k-i} y_n^i d^{(i-1)/2} \equiv k x_n^{k-1} y_n \mod y_n^3.$$

Setting $k = y_n$, we see that $y_n^2 \mid y_{nk}$, i.e., $y_n^2 \mid y_{ny_n}$. It follows by Lemma 1.5.12 that $y_n^2 \mid y_t$ for all t divisible by ny_n. Conversely, suppose $y_n^2 \mid y_t$. By Lemma 1.5.12 again, we have $n \mid t$, say $t = nk$, so $y_n^2 \mid y_{nk}$. Moreover, we have already seen that $y_{nk} \equiv kx_n^{k-1}y_n \bmod y_n^3$. It follows that $y_n^2 \mid kx_n^{k-1}y_n$, hence $y_n \mid kx_n^{k-1}$. Since $\mathrm{GCD}(x_n, y_n) = 1$, it follows that $y_n \mid k$, hence $ny_n \mid nk = t$. □

Recall that $x_{n+1} = 2ax_n - x_{n-1}$ and $y_{n+1} = 2ay_n - y_{n-1}$. We use these recurrences to establish some easy properties of x_n and y_n, by induction on n.

Theorem 1.5.14. If $a \equiv b \bmod c$, then $x_n(a) \equiv x_n(b)$, $y_n(a) \equiv y_n(b) \bmod c$.

Proof. For $n = 0$ we have $x_0(a) = 1 = x_0(b)$ and $y_0(a) = 0 = y_0(b)$. For $n = 1$ we have $x_1(a) = a \equiv b = x_1(b) \bmod c$, and $y_1(a) = 1 = y_1(b)$. Inductively we have $x_{n+1}(a) = 2ax_n(a) - x_{n-1}(a) \equiv 2ax_n(b) - x_{n-1}(b) \equiv x_{n+1}(b) \bmod c$, and similarly $y_{n+1}(a) \equiv y_{n+1}(b) \bmod c$. □

Theorem 1.5.15. $y_n \equiv n \bmod a - 1$.

Proof. For $n = 0, 1$ we have $y_0 = 0$, $y_1 = 1$. Inductively we have $y_{n+1} = 2ay_n - y_{n-1} \equiv 2an - (n-1) = 2(a-1)n + n + 1 \equiv n + 1 \bmod a - 1$. □

Theorem 1.5.16. If n is even, y_n is even. If n is odd, y_n is odd.

Proof. The initial values $y_0 = 0$ and $y_1 = 1$ are known. We have $y_{n+1} = 2ay_n - y_{n-1} \equiv -y_{n-1} \equiv y_{n-1} \bmod 2$, so our result is obvious by induction. □

1.5.3 Congruence Properties of x_n

We now prove a theorem telling for which i, j, n are $x_i \equiv x_j \bmod x_n$.

Lemma 1.5.17. $x_{2n\pm j} \equiv -x_j \bmod x_n$.

Proof. By Lemma 1.5.8 we have $x_{2n\pm j} = x_{n+(n\pm j)} = x_nx_{n\pm j} + dy_ny_{n\pm j} = x_nx_{n\pm j} + dy_n(y_nx_j \pm x_ny_j)$. Continuing modulo x_n, we have $x_{2n\pm j} \equiv dy_n^2x_j = (x_n^2 - 1)x_j \equiv -x_j$. □

Lemma 1.5.18. $x_{4n\pm j} \equiv x_j \bmod x_n$.

Proof. By the previous lemma we have $x_{4n\pm j} = x_{2n+(2n\pm j)} \equiv -x_{2n\pm j} \equiv -(-x_j) = x_j \bmod x_n$. □

Lemma 1.5.19. For all $0 \leq i < j \leq 2n$ we have $x_i \not\equiv x_j \bmod x_n$. The only exception is when $a = 2$, $n = 1$, $x_0 \equiv x_2 \bmod x_1$.

Proof. If x_n is odd, put $q = (x_n - 1)/2$. Since $2q < x_n$, the numbers

$$-q, -q+1 \ldots, -1, 0, 1, \ldots, q-1, q$$

are all pairwise $\not\equiv \bmod x_n$. Recalling $x_n = ax_{n-1} + dy_{n-1}$, we have $x_n \geq ax_{n-1} \geq 2x_{n-1}$, hence $x_{n-1} \leq x_n/2$, hence $x_{n-1} \leq q$, since x_{n-1} is an integer. It now follows that

$$-q \leq -x_{n-1} < \cdots < -x_1 < -x_0 = -1 < 0 < 1 = x_0 < x_1 < \cdots < x_{n-1} \leq q$$

are pairwise $\not\equiv$ mod x_n. Moreover, Lemma 1.5.17 tells us that $x_{n+1} \equiv -x_{n-1}$, $\ldots$, $x_{2n-1} \equiv -x_1$, $x_{2n} \equiv -x_0$, all mod x_n, and trivially $x_n \equiv 0$ mod x_n. It is now clear that all of $x_0, x_1, \ldots, x_{n-1}, x_n, x_{n+1}, \ldots, x_{2n}$ are pairwise $\not\equiv$ mod x_n.

If x_n is even, put $q = x_n/2$. Since $2q \leq x_n$, the numbers

$$-q+1, -q+2, \ldots, -1, 0, 1, \ldots, q-1, q$$

are all pairwise $\not\equiv$ mod x_n. As before we have $x_{n-1} \leq q$, so our result follows as before, unless $x_{n-1} = q = x_n/2$. In this exceptional situation we have $2x_{n-1} = x_n = ax_{n-1} + dy_{n-1}$, hence $a = 2$ and $y_{n-1} = 0$, hence $n = 1$, $x_1 = a = 2$, $x_2 = 2ax_1 - x_0 = 8 - 1 = 7 \equiv 1 = x_0$ mod x_1. □

Lemma 1.5.20. If $x_i \equiv x_j$ mod x_n, where $n \geq 1$, $0 < i \leq n$, and $0 \leq j < 4n$, then either $j = i$ or $j = 4n - i$.

Proof. Case 1: $j \leq 2n$. By Lemma 1.5.19, $i = j$ unless the exceptional case occurs. But this implies $\{i, j\} = \{0, 2\}$ and $n = 1$, contradicting our assumptions.

Case 2: $2n < j < 4n$. Set $j' = 4n - j$. Then $0 < j' < 2n$. By Lemma 1.5.18, $x_{j'} \equiv x_j$ mod x_n, hence $x_i = x_{j'}$ mod x_n. By Lemma 1.5.19, $i = j'$ unless the exceptional case occurs. This cannot happen, because both i and j' are > 0. □

Theorem 1.5.21. If $0 < i \leq n$ and $x_i \equiv x_j$ mod x_n, then $i = \pm j$ mod $4n$.

Proof. Put $j = 4nq + r$, where $0 \leq r < 4n$. By Lemma 1.5.18, $x_i \equiv x_j \equiv x_r$ mod x_n. By Lemma 1.5.20, $i = r$ or $i = 4n - r$. Thus $j \equiv r \equiv \pm i$ mod $4n$. □

1.5.4 Diophantine Definability of x_n and y_n

Theorem 1.5.22. The functions $(a, k) \mapsto x_k(a)$ and $(a, k) \mapsto y_k(a)$ are Diophantine.

Proof. We show that $x = x_k(a)$ and $y = y_k(a)$ if and only if there exist b, u, v, s, t and w_i, $1 \leq i \leq 6$, satisfying the following system of equations:

$$x^2 + (a^2 - 1)y^2 = 1 \tag{1.1}$$
$$u^2 - (a^2 - 1)v^2 = 1 \tag{1.2}$$
$$s^2 - (b^2 - 1)t^2 = 1 \tag{1.3}$$
$$v = w_1 y^2 \tag{1.4}$$
$$b = 1 + 4w_2 y \tag{1.5}$$
$$b = a + w_3 u \tag{1.6}$$
$$s = x + w_4 u \tag{1.7}$$
$$t = k + 4w_5 y \tag{1.8}$$
$$y = k + w_6 \tag{1.9}$$

Note that (1.4)–(1.9) amount to saying y^2 divides v, $b \equiv 1$ mod $4y$, $b \equiv a$ mod u, $s \equiv x$ mod u, $t \equiv k$ mod $4y$, $y \geq k$.

⇒: Assume $x = x_k(a)$ and $y = y_k(a)$. Set $n = 2ky$, $u = x_n(a)$, $v = y_n(a)$. Clearly (1.1) and (1.2) hold. Since $y_k(a) \geq k$, (1.9) holds. By Theorem 1.5.13, y_k^2 divides $y_{2ky_k} = y_n$, i.e., y^2 divides v, so (1.4) is satisfied. By Theorem 1.5.11 we have $\mathrm{GCD}(u, v) = 1$ hence $\mathrm{GCD}(u, y) = 1$. Because n is even, $y_n(a) = v$ is even (Theorem 1.5.16), hence $u = x_n(a)$ is odd, hence $\mathrm{GCD}(u, 4y) = 1$. By the Chinese Remainder Theorem, we can find b such that $b \equiv 1 \bmod 4y$ and $b \equiv a \bmod u$, so (1.5) and (1.6) are satisfied. Set $s = x_k(b)$ and $t = y_k(b)$, so (1.3) is satisfied. Since $b \equiv a \bmod u$, we have $x_k(b) \equiv x_k(a) \bmod u$, so (1.7) is satisfied. By Theorem 1.5.15, $t = y_k(b) \equiv k \bmod b - 1$. Since $4y \mid b - 1$, we have $t \equiv k \bmod 4y$, so (1.8) is satisfied. Thus all of (1.1)–(1.9) are satisfied.

⇐: Assume (1.1)–(1.9). We want to prove $x = x_k(a)$ and $y = y_k(a)$. By (1.1)–(1.3) there exist i, j, n such that $x = x_i(a)$, $y = y_i(a)$, $u = x_n(a)$, $v = y_n(a)$, $s = x_j(b)$, $t = y_j(b)$. It remains only to show that $i = k$.

By (1.6) we have $a \equiv b \bmod x_n(a)$, hence $x_j(a) \equiv x_j(b) \bmod x_n(a)$. By (1.7) we have $x_i(a) \equiv x_j(b) \bmod x_n(a)$. Hence $x_i(a) \equiv x_j(a) \bmod x_n(a)$. By (1.4) we have $y_i(a)^2 \mid y_n(a)$, hence $y_i(a) \leq y_n(a)$, hence $i \leq n$. By Theorem 1.5.21 it follows that $i = \pm j \bmod 4n$. By (1.4) we have $y_i(a)^2 \mid y_n(a)$, hence by Theorem 1.5.13 $y_i(a) \mid n$. Thus $i \equiv \pm j \bmod 4y_i(a)$. By (1.5) we have $b \equiv 1 \bmod 4y_i(a)$, i.e., $4y_i(a) \mid b - 1$. By Theorem 1.5.15, $y_j(b) \equiv j \bmod b - 1$, hence $y_j(b) \equiv j \bmod 4y_i(a)$. But by (1.8) we also have $y_j(b) \equiv k \bmod 4y_i(a)$. Thus $i \equiv \pm j \equiv \pm k \bmod 4y_i(a)$. By (1.9) we have $k \leq y_i(a)$, and obviously $i \leq y_i(a)$, hence $i = k$ and we are done. □

1.6 Proof of the Main Lemma

In this section we use properties of Pell's equation to prove the Main Lemma 1.3.7. We begin by proving that the exponential function is Diophantine.

Lemma 1.6.1. For $n, k \geq 1$ and $a \geq 2$ we have

$$n^k \equiv x_k(a) - (a - n)y_k(a) \mod 2an - n^2 - 1.$$

Proof. The proof is by induction on k. For the base cases $k = 0$ and $k = 1$, we have $x_0 - (a - n)y_0 = 1 = n^0$ and $x_1 - (a - n)y_1 = a - (a - n) = n = n^1$. Assuming our congruence for $k - 1$ and for k, we derive it for $k + 1$ using the recurrences $x_{k+1} = 2ax_k - x_{k-1}$ and $y_{k+1} = 2ay_k - y_{k-1}$. Namely,

$$\begin{aligned} x_{k+1} - (a - n)y_{k+1} &= 2a(x_k - (a - n)y_k) - (x_{k-1} - (a - n)y_{k-1}) \\ &\equiv 2an^k - n^{k-1} = n^{k-1}(2an - 1) \mod 2an - n^2 - 1 \\ &\equiv n^{k-1}n^2 = n^{k+1} \mod 2an - n^2 - 1. \end{aligned}$$

Our congruence is now proved for all k. □

Lemma 1.6.2. If $n^k < a$, then $n^k < 2an - n^2 - 1$.

Proof. Set $g(z) = 2az - z^2 - 1$ where z is a real variable. We have $g(1) = 2a - 2 \geq a$. Moreover, for $1 \leq z < a$ we have $g'(z) = 2a - 2z > 0$, hence $g(z) \geq a$. In particular, for $1 \leq n \leq n^k < a$ we have $g(n) \geq a > n^k$. □

Theorem 1.6.3. The function $(n,k) \mapsto n^k$ is Diophantine.

Proof. Set $a = x_{k+1}(n+1)$. By Theorem 1.5.22 this is a Diophantine function of n and k. By Theorem 1.5.10 we have $n^k < a$. Hence by Lemma 1.6.2 we have $n^k < 2an - n^2 - 1$. By Lemma 1.6.1 we have $n^k \equiv x_k(a) - (a-n)y_k(a) \bmod 2an - n^2 - 1$ for any a. But then, for this particular a, it follows that $n^k =$ the remainder of $x_k(a) - (a-n)y_k(a)$ on division by $2an - n^2 - 1$. It is now clear that $(n,k) \mapsto n^k$ is Diophantine, since $(a,k) \mapsto x_k(a), y_k(a)$ are Diophantine. $\square$

Having shown that the exponential function is Diophantine, we now show that the other functions mentioned in the Main Lemma are Diophantine.

Theorem 1.6.4. The function $(n,k) \mapsto \binom{n}{k}$ is Diophantine.

Proof. Given n and k, choose $M > 2^n$. We have

$$\frac{(M+1)^n}{M^k} = \sum_{i=0}^{n} \binom{n}{i} M^{i-k} = q + \epsilon$$

where $q = \sum_{i=k}^{n} \binom{n}{i} M^{i-k}$ is an integer, and

$$\epsilon = \sum_{i=0}^{k-1} \binom{n}{i} M^{i-k} \leq \frac{1}{M} \sum_{i=0}^{n} \binom{n}{i} = \frac{1}{M} 2^n < 1.$$

Moreover, $q \equiv \binom{n}{k} \bmod M$, and $\binom{n}{k} < 2^n < M$. It is now clear that $z = \binom{n}{k}$ if and only if

$$\exists M\, \exists q\, \left[\, M > 2^n \;\wedge\; q = \mathrm{Quot}((M+1)^n, M^k) \;\wedge\; z = \mathrm{Rem}(q, M) \,\right].$$

Thus $(n,k) \mapsto \binom{n}{k}$ is Diophantine. $\square$

Lemma 1.6.5. For any $M > (2n)^{n+1}$ we have

$$n! = \mathrm{Quot}\left(M^n, \binom{M}{n}\right) = \left\lfloor \frac{M^n}{\binom{M}{n}} \right\rfloor.$$

Proof. We have

$$\begin{aligned}
\frac{M^n}{\binom{M}{n}} &= \frac{M^n n!}{M(M-1)\cdots(M-n+1)} \\
&= \frac{n!}{(1-\frac{1}{M})\cdots(1-\frac{n-1}{M})} \\
&< \frac{n!}{(1-\frac{n}{M})^n} \\
&= n! \left(\frac{1}{1-\alpha}\right)^n
\end{aligned}$$

where $\alpha = n/M$. Moreover

$$\frac{1}{1-\alpha} = 1 + \frac{\alpha}{1-\alpha} < 1 + 2\alpha,$$

hence

$$\begin{aligned}
\left(\frac{1}{1-\alpha}\right)^n &< (1+2\alpha)^n \\
&= 1 + \sum_{i=1}^{n} \binom{n}{i} (2\alpha)^i \\
&= 1 + 2\alpha \sum_{i=1}^{n} \binom{n}{i} (2\alpha)^{i-1} \\
&< 1 + 2\alpha \sum_{i=0}^{n} \binom{n}{i} \\
&= 1 + (2\alpha)(2^n) = 1 + 2^{n+1}\alpha.
\end{aligned}$$

Thus

$$n! \leq \frac{M^n}{\binom{M}{n}} < n! + 1$$

provided $n!\,2^{n+1}\alpha < 1$, and this follows from $n(n!)2^{n+1} < (2n)^{n+1} < M$. □

Theorem 1.6.6. The function $n \mapsto n!$ is Diophantine.

Proof. By the previous lemma we have

$$z = n! \iff \exists M \left[M > (2n)^{n+1} \wedge z = \mathrm{Quot}\left(M^n, \binom{M}{n}\right)\right].$$

This is Diophantine in view of Theorems 1.6.3 and 1.6.4. □

Theorem 1.6.7. The function

$$(a, b, n) \mapsto h(a, b, n) = \prod_{i=0}^{n} (a + bi)$$

is Diophantine.

Proof. Given a, b, n, choose $M > (a + bn)^{n+1} \geq h(a, b, n)$ such that M is relatively prime to b. Then b is invertible mod M, i.e., $\exists c\,(bc \equiv 1 \bmod M)$, hence

$abc \equiv a \bmod M$. It follows that

$$\begin{aligned} h(a,b,n) &= \prod_{i=0}^{n}(a+bi) \\ &\equiv \prod_{i=0}^{n}(abc+bi) \mod M \\ &\equiv b^{n+1}\prod_{i=0}^{n}(ac+i) \mod M \\ &= b^{n+1}\binom{ac+n}{n+1}(n+1)!\ , \end{aligned}$$

hence $h(a,b,n) =$ the remainder of $b^{n+1}\binom{ac+n}{n+1}(n+1)!$ on division by M. It is now clear that $z = h(a,b,n)$ if and only if

$$\exists M\,\exists c\,\Big[\,M > (a+bn)^{n+1}\ \wedge\ bc \equiv 1 \bmod M\ \wedge\ z = \mathrm{Rem}\Big(b^{n+1}\tbinom{ac+n}{n+1}(n+1)!, M\Big)\,\Big]$$

and this is Diophantine in view of Theorems 1.6.3, 1.6.4, 1.6.6. $\square$

This completes the proof of the Main Lemma 1.3.7. Therefore, we have now proved the MDRP Theorem 1.3.2 and with it the unsolvability of Hilbert's Tenth Problem.

Chapter 2

Unsolvability of the Word Problem for Groups

This chapter consists mainly of a proof that the word problem for groups is unsolvable. This result is due to P. Novikov 1955 and Boone 1959. Boone's proof was simplified by Britton 1963. We follow the exposition of Rotman [12, Chapter 12]. Note also that a more streamlined proof has been given by Aanderaa/Cohen [2].

At the end of the chapter we present some related results, including unsolvability of the triviality problem for groups.

2.1 Finitely Presented Semigroups

We shall first prove that the word problem for semigroups is unsolvable. This result is due to Post 1947 and Markov 1947 and is much easier than unsolvability of the word problem for groups.

Definition 2.1.1. A *semigroup* is a set S together with an associative binary operation $\cdot : S \times S \to S$. We consider only semigroups with an *identity element*, i.e., $1 \in S$ such that $s \cdot 1 = 1 \cdot s = s$ for all $s \in S$.

Example 2.1.2. Let $a_1, \ldots, a_n$ be a finite alphabet. Let S_n be the set of words on $a_1, \ldots, a_n$. A *word* is a finite sequence of letters of the alphabet, $W = a_{i_1} \cdots a_{i_k}$, where $1 \le i_j \le n$ for $1 \le j \le k$. Here k is the *length* of W. If $k = 0$, W is the empty word. Note that S_n is a semigroup under concatenation. For example, If $U = abaac$, $V = baba$, then $UV = abaacbaba$. This semigroup

$$S_n = \langle a_1, \ldots, a_n \rangle$$

is called the *free semigroup* on $a_1, \ldots, a_n$.

Definition 2.1.3. Let R be a subset of $S_n \times S_n$. We define an equivalence relation $\approx_R$ on S_n. For $W, W' \in S_n$, define $W \approx_R W'$ if and only if there exists a finite sequence of words $W = W_0, W_1, \ldots, W_t = W'$ such that, for each $i < t$, $W_i \sim_R W_{i+1}$, i.e., $W_i = UXV$ and $W_{i+1} = UYV$ for some (X,Y) or $(Y,X) \in R$. For $W \in S_n$ we write $[W]_R = \{W' \in S_n \mid W \approx_R W'\}$ = the equivalence class of W modulo $\approx_R$. We put $S = S_n/{\approx_R}$ = the set of such equivalence classes. This is a semigroup, with the operation $\cdot$ being given by $[U]_R \cdot [V]_R = [UV]_R$. The identity element is $1 = [\varepsilon]_R$ where ε is the empty word. We frequently write W instead of $[W]_R$. Our semigroup S is written as

$$S = \langle a_1, \ldots a_n \mid R \rangle.$$

Each $(X,Y) \in R$ is viewed as a *relation* $X = Y$ which holds in S.

Example 2.1.4. Let S be the semigroup $\langle a, b \mid a^3 = 1, ab = ba \rangle$. We refer to S as the semigroup with generators a, b and relations $a^3 = 1$, $ab = ba$. Elements of S are words on the alphabet a, b except we can reduce equivalent words, e.g., $aababa = aaabba = aaabab = aaaabb = abb$. In fact, each word is equivalent to a unique one of the form $a^i b^j$, where $0 \le i \le 2$, $j \ge 0$. Thus each element of S has a normal form. The multiplication of normal forms is given by $a^i b^j a^s b^t = a^k b^{j+t}$, where $k = \mathrm{Rem}(i+s, 3)$.

Definition 2.1.5. A *finitely presented semigroup* is a semigroup of the form $\langle a_1, \ldots, a_n \mid R \rangle$, where $a_1, \ldots, a_n$ is a finite set of generators and R is a finite set of relations.

Definition 2.1.6. Let $S = \langle a_1, \ldots, a_n \mid R \rangle$ be a finitely presented semigroup. The *word problem* for S is the problem, given two words $W, W' \in S_n$, to decide whether $W = W'$ in S, i.e., whether $W \approx_R W'$.

Example 2.1.7. The word problem for $\langle a, b \mid a^3 = 1, ab = ba \rangle$ is solvable, because $a^i b^j = a^s b^t$ in S if and only if $i \equiv s \bmod 3$, and $j = t$. In fact, each word on a, b is equivalent to a unique normal form $a^i b^j$, $0 \le i \le 2$, $j \ge 0$, and two normal forms are equivalent if and only if they are equal.

Remark 2.1.8. In general, the word problem for a finitely presented semigroup $S = \langle A \mid R \rangle$ is a recursively enumerable or Σ^0_1 problem. This is because $W = W'$ in S if and only if $\exists t$ $\exists$ finite sequence of words $W_0, W_1, \ldots, W_t$ such that

$$W \equiv W_0 \sim_R W_1 \sim_R \cdots \sim_R W_t \equiv W' .$$

Theorem 2.1.9 (Post, Markov). We can construct a finitely presented semigroup S such that the word problem for S is unsolvable.

In order to prove this theorem, we shall encode the Halting Problem into the word problem for a particular finitely presented semigroup.

Recall that a k-place partial function ψ is partial recursive if and only if it is computable by some register machine program $\mathcal{P}$. Please refer to the Math 558 notes [14] for the definition of register machine programs.

Let $\mathcal{P}$ be a register machine program. Recall that $\mathcal{P}(x_1, \ldots, x_k)$ is the run of $\mathcal{P}$ started with $x_1, \ldots, x_k$ in registers $R_1, \ldots, R_k$ and all other registers empty.

Lemma 2.1.10. We can find a program $\mathcal{P}$ such that, given $x \in \mathbb{N}$, it is undecidable whether $\mathcal{P}(x)$ halts.

Proof. By the Enumeration Theorem, let $\mathcal{P}$ be a program computing the partial recursive function $e \mapsto \varphi_e^{(1)}(0)$: that is, $\mathcal{P}$ takes a number e, constructs the program with that Gödel number, and then runs the program with input 0. Thus $\mathcal{P}(e)$ halts if and only if $e \in H$, where H is the Halting Set. By Turing's work, H is undecidable, so the Halting Problem for $\mathcal{P}$ is undecidable. $\square$

Notation 2.1.11. We write

$$p_0,\, p_1,\, \ldots,\, p_i,\, \ldots$$

for the prime numbers $2, 3, 5, 7, 11, \ldots$. Thus p_i is the ith prime, where we start indexing with 0.

Lemma 2.1.12. Given a k-place partial recursive function $\psi(x_1, \ldots, x_k)$, we can find a 1-place partial recursive function $\psi^*(z)$ such that

$$\psi^*(p_1^{x_1} \cdots p_k^{x_k}) \;\simeq\; p_{k+1}^{\psi(x_1,\ldots,x_k)}$$

for all $x_1, \ldots, x_k$, and $\psi^*(z)$ is computable by a register machine program using only two registers, R_1 and R_2.

Proof. Let $\mathcal{P}$ be a register machine program which computes ψ. Let

$$P_1, \ldots, P_k, P_{k+1}, \ldots, P_s$$

be the registers used in $\mathcal{P}$. We may safely assume that, whenever $\mathcal{P}(x_1, \ldots, x_k)$ halts, it leaves all registers except possibly P_{k+1} empty. Our new program $\mathcal{P}^*$ for ψ^* will be constructed so as to simulate $\mathcal{P}$ using only two registers, R_1 and R_2. In the new program, R_1 is used to hold a number z which encodes the contents of $P_1, \ldots, P_s$ via prime power coding, i.e.,

$$z = \prod_{i=1}^{s} p_i^{z_i}$$

where z_i is the content of P_i. Then R_2 is used for scratch work. Each P_i^+ instruction is replaced by a program for $z \mapsto z \cdot p_i$. Each P_i^- instruction is replaced by a program for

$$z \mapsto \begin{cases} z/p_i \text{ if } p_i \text{ divides } z, \\ z \text{ otherwise.} \end{cases}$$

We shall see that this simulation can be performed using only R_1 and R_2. It is then clear that $\psi^*(p_1^{x_1} \cdots p_k^{x_k}) \simeq p_{k+1}^{\psi(x_1,\ldots,x_k)}$ for all $x_1, \ldots, x_k$.

The details of the simulation are as follows.

We replace $\longrightarrow P_i^+ \longrightarrow$ in $\mathcal{P}$ by Figure 2.1 in $\mathcal{P}^*$.

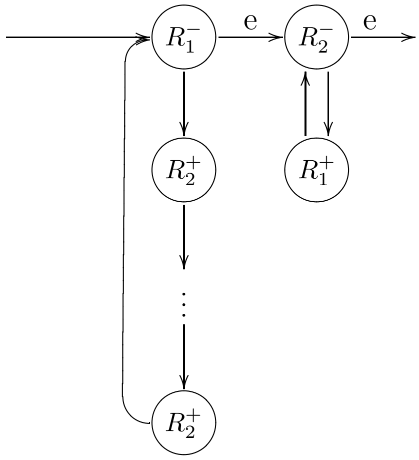
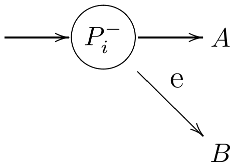
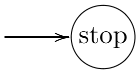
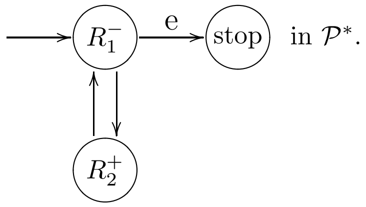


Figure 2.1: Incrementing P_i. The number of R_2^+ instructions is p_i.

We replace 

in $\mathcal{P}$ by Figure 2.2 in $\mathcal{P}^*$.

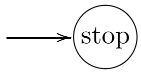
We replace 

in $\mathcal{P}$ by 

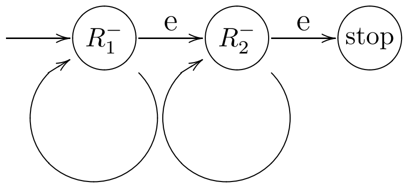
in $\mathcal{P}^*$.

This completes the proof of Lemma 2.1.12. □

Theorem 2.1.13. We can find a program $\mathcal{P}$ using only two registers, R_1 and R_2, such that, given $x \in \mathbb{N}$, it is undecidable whether $\mathcal{P}(x)$ halts. Furthermore, when it halts, R_1 and R_2 are empty.

Proof. We begin with the program of Lemma 2.1.10. Using Lemma 2.1.12 we convert it to a program using only R_1 and R_2. We then replace 

by



to clear R_1 and R_2 before halting. □

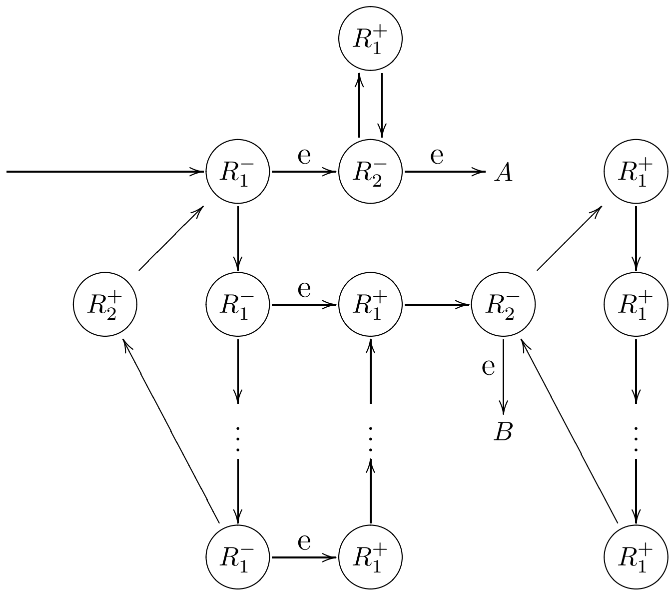


Figure 2.2: Decrementing P_i. The number of R_1^- instructions is p_i.

We now construct a semigroup S with unsolvable word problem.

Definition 2.1.14. Let $\mathcal{P}$ be a program using only two registers R_1, R_2 as in Theorem 2.1.13. Let $I_1, \ldots, I_l$ be the instructions of $\mathcal{P}$. As usual, I_1 is the first instruction executed, and I_0 is the halt instruction. Our semigroup S will have $l+3$ generators $a, b, q_0, q_1, \ldots, q_l$. If R_1 and R_2 contain x and y respectively, and if I_m is about to be executed, then we represent this state as a word $ba^x q_m a^y b$. Thus a serves as a counting token, and b serves as an end-of-count marker. For each $m = 1, \ldots, l$, if I_m says "increment R_1 and go to I_{n_0}," we represent this as a production $q_m \to aq_{n_0}$ or as a relation $q_m = aq_{n_0}$. If I_m says "increment R_2 and go to I_{n_0}," we represent this as a production $q_m \to q_{n_0}a$ or as a relation $q_m = q_{n_0}a$. If I_m says "if R_1 is empty go to I_{n_0} otherwise decrement R_1 and go to I_{n_1}," we represent this as a pair of productions $bq_m \to bq_{n_0}$, $aq_m \to q_{n_1}$, or as a pair of relations $bq_m = bq_{n_0}$, $aq_m = q_{n_1}$. If I_m says "if R_2 is empty go to I_{n_0} otherwise decrement R_2 and go to I_{n_1}," we represent this as a pair of productions $q_m b \to q_{n_0} b$, $q_m a \to q_{n_1}$, or as a pair of relations $q_m b = q_{n_0} b$, $q_m a = q_{n_1}$. Thus the total number of productions or relations is $l^+ + 2l^-$, where $l = l^+ + l^-$ and l^+ is the number of increment instructions and l^- is the number of decrement instructions. Let S be the semigroup described by these generators and relations.

Theorem 2.1.15. $\mathcal{P}(x)$ halts if and only if $ba^x q_1 b = bq_0 b$ in S.

Proof. The "if" part is clear. For the "only if" part, assume that $ba^x q_1 b = bq_0 b$ in S. This implies that there is a sequence of words $ba^x q_1 b = W_0 = \cdots =$

$W_n = bq_0b$ where each W_{i+1} is obtained from W_i by a forward or backward production. We claim that the backward productions can be eliminated. In other words, if there are any backward productions, we can replace the sequence $W_0, \ldots, W_n$ by a shorter sequence. This is actually obvious, because if there is a backward production then there must be one which is immediately followed by a forward production, and these two must be inverses of each other, because $\mathcal{P}$ is deterministic. Thus we see that $ba^xq_1b = bq_0b$ via a sequence of forward productions. This implies that $\mathcal{P}(x)$ halts. Our claim is proved. $\square$

From the previous theorem, it follows that our semigroup S has unsolvable word problem. This proves Theorem 2.1.9.

2.2 The Boone Group

Definition 2.2.1. A *group* is a semigroup G such that $\forall g \in G\, \exists g^{-1} \in G$ such that $gg^{-1} = g^{-1}g = 1$.

Notation 2.2.2. Let A be an alphabet. We introduce new letters a^{-1}, $a \in A$, and we write $A^{-1} = \{a^{-1} \mid a \in A\}$. We also write $(a^{-1})^{-1} = a$. A word on $A \cup A^{-1}$ is said to *involve* a if it contains an occurrence of a or a^{-1}.

Definition 2.2.3. A *group presentation* is a semigroup presentation

$$G = \langle A \cup A^{-1} \mid R \rangle$$

where R includes semigroup relations

$$aa^{-1} = a^{-1}a = 1,$$

i.e., $aa^{-1} = a^{-1}a = \varepsilon$, for all $a \in A$, where ε is the empty word. We abbreviate this as

$$G = \langle A \mid R \rangle.$$

Note that G is a group, because for any word $a_{i_1}^{e_1} \cdots a_{i_k}^{e_k}$ on $A \cup A^{-1}$ we have

$$(a_{i_1}^{e_1} \cdots a_{i_k}^{e_k})^{-1} = a_{i_k}^{-e_k} \cdots a_{i_1}^{-e_1}.$$

Definition 2.2.4. A *finitely presented group* is a group with a finite presentation, i.e., $G = \langle A \mid R \rangle$ where A and R are finite.

Definition 2.2.5. Let G be a finitely generated group, and let A be a finite generating set. The *word problem for G* is the problem, given a word W on $A \cup A^{-1}$, to decide whether $W = 1$ in G.

Remark 2.2.6. If G is a finitely generated group, the degree of unsolvability of the word problem of G is independent of finite set of generators chosen. The same holds for semigroups.

We now exhibit a finitely presented group with unsolvable word problem. To do this, we build upon our construction of a finitely presented semigroup with unsolvable word problem. We use some special features of the earlier construction.

Remark 2.2.7. In Section 2.1 we constructed a finitely presented semigroup $S = \langle A \cup Q \mid R \rangle$ with unsolvable word problem, where

$$A = \{a, b\}, \qquad Q = \{q_0, \ldots, q_l\}.$$

Recall that the relations of S were of the form

$$R = \{X_i q_{m_i} Y_i = U_i q_{n_i} V_i \mid i \in I\}$$

where X_i, Y_i, U_i, V_i are words on A. We showed that, given words X, Y on A, it is undecidable whether $X q_m Y = b q_0 b$ in S.

We now introduce a new generator $q = q_{l+1}$ into Q, and we introduce a new relation $b q_0 b = q$ into R. With this trivially modified presentation of the semigroup S, we now have $Q = \{q, q_0, \ldots, q_l\}$. Moreover, given words X, Y on A, it is undecidable whether $X q_m Y = q$ in S.

Notation 2.2.8. If $X = a_{i_1} \cdots a_{i_k}$ is a word on A, we write

$$\overline{X} = a_{i_1}^{-1} \cdots a_{i_k}^{-1}.$$

Note that $\overline{X} \neq X^{-1}$. If X and Y are words on A, we write $(X q_m Y)^* = \overline{X} q_m Y$.

We now construct a group with unsolvable word problem.

Definition 2.2.9 (the Boone group). Let

$$S = \langle A \cup Q \mid X_i q_{m_i} Y_i = U_i q_{n_i} V_i, i \in I \rangle$$

be a finitely presented semigroup as in Remark 2.2.7 above. Let G be the group with generators

$$A \cup Q \cup \{r_i \mid i \in I\} \cup \{x, t, k\}$$

and relations

$$xa = ax^2$$

$$r_i a = a x r_i x$$

$$r_i^{-1} \overline{X_i} q_{m_i} Y_i r_i = \overline{U_i} q_{n_i} V_i$$

$$tx = xt, \quad tr_i = r_i t$$

$$kx = xk, \quad kr_i = r_i k$$

$$k(q^{-1} t q) = (q^{-1} t q) k$$

for all $a \in A$ and $i \in I$. Note that G is a finitely presented group. This particular group is due to Boone.

Theorem 2.2.10 (Boone). Let X and Y be words on A. Put

$$\Sigma = (Xq_mY)^* = \overline{X}q_mY.$$

The following are pairwise equivalent.

1. $Xq_mY = q$ in S.
2. $\Sigma = LqR$ in G, where L, R are some words on $x, x^{-1}, r_i, r_i^{-1}, i \in I$.
3. $k(\Sigma^{-1}t\Sigma) = (\Sigma^{-1}t\Sigma)k$ in G.

Corollary 2.2.11. The word problem for the Boone group G is unsolvable.

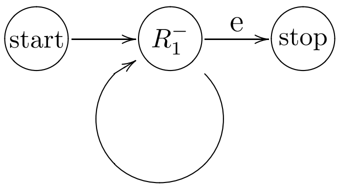
Proof. This is immediate from $1 \Leftrightarrow 3$ in Boone's Theorem 2.2.10, plus the known undecidability of $Xq_mY = q$ in S. $\square$

Theorem 2.2.12 (P. Novikov, Boone). The word problem for groups is unsolvable.

Proof. This is immediate from Corollary 2.2.11. $\square$

Before starting the proof of Boone's Theorem 2.2.10, we give an example.

Example 2.2.13. Consider the register machine program $\mathcal{P}$ which empties R_1 and halts:



.

The Post semigroup relations for $\mathcal{P}$ are:

$$\begin{aligned} aq_1 &= q_1, \\ bq_1 &= bq_0, \\ bq_0b &= q. \end{aligned}$$

The corresponding Boone group relations are:

$$\begin{aligned} r_1^{-1}a^{-1}q_1r_1 &= q_1, \\ r_2^{-1}b^{-1}q_1r_2 &= b^{-1}q_0, \\ r_3^{-1}b^{-1}q_0br_3 &= q. \end{aligned}$$

In addition, the Boone group has a generator x and relations $xa = ax^2$, $xb = bx^2$, $r_ia = axr_ix$, $r_ib = bxr_ix$, $i = 1, 2, 3$. This gives a subgroup G_3 of the Boone group (see also Lemma 2.3.6 below). The full Boone group is obtained by introducing additional generators t, k and their associated relations.

Consider $\mathcal{P}(1)$, the run of $\mathcal{P}$ starting with 1 in R_1, 0 in R_2. In the semigroup we have

$$baq_1b = bq_1b = bq_0b = q.$$

Hence in the group we have

$$\begin{aligned} q &= r_3^{-1}b^{-1}q_0br_3 \\ &= r_3^{-1}r_2^{-1}b^{-1}q_1r_2br_3 \\ &= r_3^{-1}r_2^{-1}b^{-1}q_1bxr_2xr_3 \\ &= r_3^{-1}r_2^{-1}b^{-1}r_1^{-1}a^{-1}q_1r_1bxr_2xr_3 \\ &= \underbrace{r_3^{-1}r_2^{-1}x^{-1}r_1^{-1}x^{-1}}_{L}b^{-1}a^{-1}q_1b\underbrace{xr_1x^2r_2xr_3}_{R}. \end{aligned}$$

Setting $\Sigma = (baq_1b)^* = b^{-1}a^{-1}q_1b$, we have $q = L\Sigma R$, where L is a word on $x^{-1}, r_i^{-1}, i \in I$, and R is a word on $x, r_i, i \in I$. This is an instance of statement 2 of Boone's Theorem.

We now begin the proof of Boone's Theorem.

Proof of $1 \Rightarrow 2$ *and* $2 \Rightarrow 3$. Assume $Xq_mY = q$ in S. Say

$$Xq_mY = W_0 = \cdots = W_n = q$$

where for each $\nu = 1, \ldots, n$ there exists $i \in I$ such that $W_{\nu-1}$ and W_ν are of the form $PX_iq_{m_i}Y_iQ$ and $PU_iq_{n_i}V_iQ$, where P, Q are words on A.

Note that for any word P on A, we have $r_iP = PR$ and $\overline{P}r_i^{-1} = L\overline{P}$, where R, L are words on x, x^{-1}, r_i, r_i^{-1}. Hence in G we have

$$\begin{aligned} \overline{PU_i}q_{n_i}V_iQ &= \overline{P}r_i^{-1}\overline{X_i}q_{m_i}Y_ir_iQ \\ &= L\overline{PX_i}q_{m_i}Y_iQR, \end{aligned}$$

hence $W_{\nu-1}^* = L_\nu W_\nu^* R_\nu$ for each $\nu = 1, \cdots, n$. Hence $W_0^* = LW_n^*R$ where

$$L = L_1 \cdots L_n, \quad R = R_n \cdots R_1$$

are words on $x, x^{-1}, r_i, r_i^{-1}, i \in I$. But $W_0^* = (Xq_mY)^* = \Sigma$, and $W_n^* = q^* = q$. Thus we have $\Sigma = LqR$ in G. Now, by the relations of G, we have

$$\begin{aligned} k(\Sigma^{-1}t\Sigma) &= kR^{-1}q^{-1}L^{-1}tLqR \\ &= kR^{-1}q^{-1}tqR \\ &= R^{-1}kq^{-1}tqR \\ &= R^{-1}q^{-1}tqkR \\ &= R^{-1}q^{-1}tqRk \\ &= R^{-1}q^{-1}L^{-1}tLqRk \\ &= (\Sigma^{-1}t\Sigma)k\,. \end{aligned}$$

Thus we have proved $1 \Rightarrow 2$ and $2 \Rightarrow 3$ in Boone's Theorem. □

It remains to prove $3 \Rightarrow 2$ and $2 \Rightarrow 1$.

2.3 HNN Extensions and Britton's Lemma

In order to finish the proof of Boone's Theorem, we first study HNN extensions.

Remark 2.3.1. Given a group G, and given $p \in G$, the map $G \to G$ given by $g \mapsto p^{-1}gp$ is an automorphism of G. Such automorphisms are called *inner automorphisms*. We shall see that all of the relations used to define the Boone group G describe properties of inner automorphisms.

Theorem 2.3.2 (Higman/Neumann/Neumann). Let G be a group. Let H, K be subgroups of G which are isomorphic to each other. Let $\phi : H \cong K$ be a particular isomorphism of H onto K. Then there exists a group $G^* \supseteq G$ and a group element $p \in G^*$ such that $p^{-1}hp = \phi(h)$ for all $h \in H$.

Definition 2.3.3 (HNN extensions). Let $G = \langle A \mid R\rangle$ be a group presentation. Let $\phi : H \cong K$ be an isomorphism of a subgroup of G onto another subgroup of G. Consider the group presentation $G^* = \langle A^* \mid R^*\rangle$ where $A^* = A \cup \{p\}$, and $R^* = R \cup \{p^{-1}Xp = \phi(X)\}_X$ where X ranges over a set of words on $A \cup A^{-1}$ which generate H. We sometimes write this as

$$G^* = \langle G, p \mid p^{-1}Xp = \phi(X)\rangle_X.$$

By the HNN Theorem 2.3.2, the identity map $a \mapsto a$, $a \in A$, gives an embedding of G into G^*. Then G^* is called an *HNN extension* of G, with *stable letter p*.

An important special case of an HNN extension is when ϕ is the identity map and $H = K$, as follows.

Definition 2.3.4 (commuting HNN extensions). Let $G = \langle A \mid R\rangle$ be a group. Let H be any subgroup of G. Consider $G' = \langle A' \mid R'\rangle$ where $A' = A \cup \{p\}$, and $R' = R \cup \{p^{-1}Xp = X\}_X$ where X ranges over a set of generators of H. Then G' is called a *commuting HNN extension* of G, with stable letter p. Thus we have

$$G' = \langle G, p \mid p^{-1}Xp = X\rangle_X.$$

Note also that $p^{-1}Xp = X$ can be written as $pX = Xp$.

Remark 2.3.5. The Boone group is nothing but a finite sequence of HNN extensions. More precisely, each of the letters in our presentation of the Boone group was introduced as a stable letter for an HNN extension. In particular, the letters t and k in Definition 2.2.9 are stable letters for commuting HNN extensions. We spell all this out in the proof of the following lemma.

Lemma 2.3.6. The Boone group G (see Definition 2.2.9) is obtained as an iterated HNN extension.

Proof. We start with the infinite cyclic group $G_0 = \langle x\rangle$. Clearly

$$G_1 = \langle x, a, a \in A \mid a^{-1}xa = x^2, a \in A\rangle$$

is a multiple HNN extension (see Definition 2.6.2 below) of G_0 with stable letters a, $a \in A$.

Consider the free product

$$G_2 = G_1 * \langle q, q_0, \dots, q_l \rangle$$

where $\langle q, q_0, \dots, q_l \rangle$ is the free group on $q, q_0, \dots, q_l$. We claim that

$$G_3 = \langle G_2, r_i, i \in I \mid r_i^{-1} a x r_i = a x^{-1}, r_i^{-1} \overline{X_i} q_{m_i} Y_i r_i = \overline{U_i} q_{n_i} V_i, a \in A, i \in I \rangle$$

is a multiple HNN extension of G_2 with stable letters r_i, $i \in I$. To see this, consider the subgroups H_i and K_i of G_2 generated by $\overline{X_i} q_{m_i} Y_i$, ax, $a \in A$, and $\overline{U_i} q_{n_i} V_i$, ax^{-1}, $a \in A$, respectively. It is not hard to see that H_i and K_i are free on these generators. Hence there are isomorphisms $\phi_i : H_i \cong K_i$ given by $\phi_i(\overline{X_i} q_{m_i} Y_i) = \overline{U_i} q_{n_i} V_i$, $\phi_i(ax) = ax^{-1}$, $a \in A$. Thus G_3 is a multiple HNN extension of G_2 as claimed.

Next we have

$$G_4 = \langle G_3, t \mid tx = xt, tr_i = r_i t, i \in I \rangle$$

which is a commuting HNN extension of G_3 with stable letter t. Finally, the Boone group is

$$G = G_5 = \langle G_4, k \mid kx = xk, kr_i = r_i k, k(q^{-1} t q) = (q^{-1} t q) k, i \in I \rangle$$

which is a commuting HNN extension of G_4 with stable letter k. □

Our proof of Boone's Theorem will be based on a detailed understanding of HNN extensions. A key property is given by Britton's Lemma, below.

Definition 2.3.7. In an HNN extension, a *pinch* is a word of the form $p^{-1} X p$ or $p X p^{-1}$ where X is a word on $A \cup A^{-1}$ lying in H or K respectively. A word containing no pinches is said to be *reduced*.

Remark 2.3.8. In an HNN extension, any word is equivalent to a reduced word. This is because the relations of G^* allow us to replace pinches by words not involving p or p^{-1}. Namely, if X is a word on $A \cup A^{-1}$ lying in H, then $p^{-1} X p = \phi(X)$ is equivalent to a word on $A \cup A^{-1}$ lying in K. Likewise, if X is a word on $A \cup A^{-1}$ lying in K, then $p X p^{-1} = \phi^{-1}(X)$ is equivalent to a word on $A \cup A^{-1}$ lying in H.

Lemma 2.3.9 (Britton's Lemma). Let W be a word involving p or p^{-1}. If $W = 1$ in G^*, then W contains a pinch.

The proofs of the HNN Theorem 2.3.2 and Britton's Lemma 2.3.9 are spread out over Sections 2.4, 2.5, 2.6 below.

2.4 Free Products With Amalgamation

In order to prove the HNN Theorem and Britton's Lemma, we first introduce free products with amalgamation. The proof of the HNN Theorem is at the end of this section.

Definition 2.4.1 (free product). Let G_1, G_2 be groups, which we assume to be disjoint except for the identity element, 1. The *free product* $G_1 * G_2$ is the group consisting of all formal products $g_1 \cdots g_n$ where $n \geq 0$, $g_i \neq 1$, and adjacent g_i belong to distinct G_j. Note that distinct n-tuples $g_1, \ldots, g_n$ as above give rise to distinct elements of $G_1 * G_2$.

Remark 2.4.2. Intuitively, the free product $G_1 * G_2$ is the "largest" group generated by $G_1 \cup G_2$. One way to see this is in terms of generators and relations: if $G_1 = \langle A_1 \mid R_1 \rangle$ and $G_2 = \langle A_1 \mid R_2 \rangle$, then $G_1 * G_2 = \langle A_1 \cup A_2 \mid R_1 \cup R_2 \rangle$. Another way to see it is in terms of a universal mapping property:

$$\begin{tikzcd}
& K & \\
G_1 \arrow[ur] \arrow[r] & G_1 * G_2 \arrow[u, dashed] & G_2 \arrow[ul] \arrow[l]
\end{tikzcd}$$

This means that, given maps from G_1 and G_2 to K, a unique map from $G_1 * G_2$ to K is determined.

Example 2.4.3. The free group on n generators may be viewed as a free product

$$F_n = \langle a_1, \ldots, a_n \rangle = \langle a_1 \rangle * \cdots * \langle a_n \rangle$$

where $\langle a_1 \rangle$, ..., $\langle a_n \rangle$ are infinite cyclic groups.

Corollary 2.4.4. *G_1 and G_2 are subgroups of $G_1 * G_2$. Moreover, in $G_1 * G_2$ we have $G_1 \cap G_2 = 1$.*

Corollary 2.4.5. *For all $g_1, \ldots, g_n$, $n \geq 1$, $g_i \neq 1$, with adjacent g_i from distinct G_j, we have $g_1 \cdots g_n \neq 1$ in $G_1 * G_2$.*

Definition 2.4.6 (free product with amalgamation). Let H be a subgroup embedded in both G_1 and G_2 via $\iota_1 : H \hookrightarrow G_1$ and $\iota_2 : H \hookrightarrow G_2$. Define $G_1 *_H G_2 = G_1 * G_2 / N$ where N is the normal subgroup of $G_1 * G_2$ generated by $\iota_1(h)\iota_2(h)^{-1}$, $h \in H$. The group $G_1 *_H G_2$ is called a *free product with amalgamation.* In terms of generators and relations, if $G_1 = \langle A_1 \mid R_1 \rangle$, $G_2 = \langle A_2 \mid R_2 \rangle$, and $A_1 \cap A_2 = \emptyset$, then

$$G_1 *_H G_2 = \langle A_1, A_2 \mid R_1, R_2, \iota_1(X) = \iota_2(X) \rangle_X$$

where X ranges over a set of generators of H.

Remark 2.4.7. There is a universal mapping property given by the following diagram:

$$\begin{tikzcd}
& K & \\
G_1 \arrow[ur] \arrow[r] & G_1 *_H G_2 \arrow[u, dashed] & G_2 \arrow[ul] \arrow[l] \\
& H \arrow[ul] \arrow[ur] &
\end{tikzcd}$$

This means that, given maps from G_1 and G_2 to K which induce the same map from H to K, a unique map from $G_1 *_H G_2$ to K is determined.

Remark 2.4.8. To obtain a concrete description of the elements of $G_1 *_H G_2$, assume $H \subseteq G_1$ and $H \subseteq G_2$. For each $g \in G_j \setminus H$, $j = 1, 2$ let $\overline{g} \in G_j \setminus H$ be a representative of the coset gH. Note that we have uniquely $g = \overline{g}h$ for some $h \in H$. Then $G_1 *_H G_2$ is concretely the set of formal products $\overline{g_1} \cdots \overline{g_n} h$, $n \geq 0$, where adjacent g_i come from distinct $G_j \setminus H$, and $h \in H$.

Corollary 2.4.9. *G_1 and G_2 are subgroups of $G_1 *_H G_2$, and $G_1 \cap G_2 = H$.*

Proof. The first statement is immediate from Remark 2.4.8. Suppose $g_1 \in G_1 \setminus H$, $g_2 \in G_2 \setminus H$. We have $g_1 = \overline{g_1}h_1$, $g_2 = \overline{g_2}h_2$, and $\overline{g_1} \neq \overline{g_2}$, hence $g_1 \neq g_2$. □

Corollary 2.4.10. *Let $n \geq 1$ and let $g_1, \ldots, g_n \in G_j \setminus H$ where adjacent g_i come from distinct G_j. Then in $G_1 *_H G_2$ we have $g_1 \cdots g_n \notin H$, hence $g_1 \cdots g_n \neq 1$.*

Proof. For $1 \leq i \leq n$ we have $g_i \notin H$, hence $g_i = \overline{g_i}h_i$, $h_i \in H$. We then have

$$\begin{aligned}
g_1 \cdots g_n &= \overline{g_1}h_1\overline{g_2}h_2 \cdots \overline{g_n}h_n \\
&= \overline{g_1}\overline{h_1\overline{g_2}}h_2' \cdots \overline{g_n}h_n \\
&\cdots \\
&= \overline{g_1}\overline{g_2'} \cdots \overline{g_n'}h'
\end{aligned}$$

which is clearly $\notin H$. □

We now use a free product with amalgamation to prove the HNN Theorem.

Proof of the HNN Theorem 2.3.2. Let $\phi : H \cong K$ with $H, K \subseteq G$. Let $M = G * \langle u \rangle$ where u is a new letter. Let P be the subgroup of M generated by $G \cup u^{-1}Hu$. Note that $P = G * u^{-1}Hu$ within M, because there can be no equation $g_0(u^{-1}h_1u)g_1(u^{-1}h_2u) \cdots g_{n-1}(u^{-1}h_nu)g_n = 1$ with $g_i \in G$, $h_i \in H$, $h_1 \neq 1$, $g_1 \neq 1$, $h_2 \neq 1$, $\ldots$, $g_{n-1} \neq 1$, $h_n \neq 1$, $n \geq 1$.

Similarly, let $N = G * \langle v \rangle$ where v is a new letter, and let $Q = G * v^{-1}Kv$ be the subgroup of N generated by $G \cup v^{-1}Kv$. Clearly $P \cong Q$ via θ defined by $\theta(g) = g$, $\theta(u^{-1}hu) = v^{-1}\phi(h)v$ for all $g \in G$, $h \in H$.

Consider the free product with amalgamation $G' = M *_\theta N$. Note that $G \hookrightarrow G'$ via $g \mapsto g$. For all $h \in H$ we have $u^{-1}hu = v^{-1}\phi(h)v$, hence $p^{-1}hp = \phi(h)$ where $p = uv^{-1}$. This proves the HNN Theorem. □

We still need to prove Britton's Lemma.

2.5 Proof of $3 \Rightarrow 2$

In this section we prove Britton's Lemma (Lemma 2.3.9) in the special case of commuting HNN extensions (see Definition 2.3.4). We then use this special case to obtain the implication $3 \Rightarrow 2$ in Boone's Theorem 2.2.10.

Notation 2.5.1. In the proof of Britton's Lemma and Boone's Theorem, we shall frequently write $W \equiv W'$ for words W and W', meaning that W and W' are *identical* as words. This is in contrast to $W = W'$ which means merely that W and W' are *equal* as elements of some group.

Let $G = \langle A \mid R \rangle$ be a group. Let H be a subgroup of G generated by words X_i, $i \in I$, on $A \cup A^{-1}$. Let t be a new letter, and consider the commuting HNN extension

$$G' = \langle A, t \mid R, t^{-1}X_i t = X_i, i \in I \rangle .$$

Lemma 2.5.2. $G' \cong G *_H (H \times \langle t \rangle)$ via the canonical map $a \mapsto a$, $t \mapsto t$, $a \in A$.

Proof. Let $H = \langle x_i, i \in I \mid S \rangle$ be a presentation of H on generators x_i corresponding to X_i, $i \in I$. Then $G *_H (H \times \langle t \rangle)$ has the presentation

$$\langle A, x_i, i \in I, t \mid R, S, t^{-1}x_i t = x_i, X_i = x_i, i \in I \rangle.$$

In this presentation, the relations S are superfluous, so we have

$$\langle A, x_i, i \in I, t \mid R, t^{-1}x_i t = x_i, X_i = x_i, i \in I \rangle.$$

Now the generators x_i, $i \in I$ are superfluous, so we have simply

$$\langle A, t \mid R, t^{-1}X_i t = X_i, i \in I \rangle$$

which is G'. □

Lemma 2.5.3. Let W be a word involving t or t^{-1}. If $W = 1$ in G', then W contains a pinch, i.e., a subword of the form $t^{-1}Xt$ or tXt^{-1} where X is a word on $A \cup A^{-1}$ lying in H.

Proof. If W contains a subword of the form $t^{-1}t$ or tt^{-1}, we are done. Hence we may safely assume

$$W \equiv W_0 t^{e_1} W_1 t^{e_2} W_2 \cdots W_{n-1} t^{e_n} W_n = 1$$

where $n \geq 1$, $e_i \neq 0$, W_i is a word on $A \cup A^{-1}$, and $W_1, \ldots, W_{n-1}$ are nonempty.

We proceed by induction on n. If $n = 1$ we have $W \equiv W_0 t^{e_1} W_1 = 1$ in G', hence $t^{e_1} = W_0^{-1} W_1^{-1}$ in G'. However, by Lemma 2.5.2 $G' \cong G *_H (H \times \langle t \rangle)$ is a free product with amalgamation, hence by Corollary 2.4.9 $t^{e_1} \in G \cap (H \times \langle t \rangle) = H$, which is clearly impossible.

Assume now that $n > 1$. Apply Corollary 2.4.10 to the factorization

$$W \equiv W_0(t^{e_1})W_1 \cdots W_{n-1}(t^{e_n})W_n = 1.$$

Clearly $t^{e_1}, \ldots, t^{e_n} \notin H$, hence at least one of $W_0, W_1, \ldots, W_n$ lies in H. If $W_0 \in H$, replace $W_0(t^{e_1})$ by $(W_0 t^{e_1}) \in H \times \langle t \rangle \setminus H$. If $W_n \in H$, replace $(t^{e_n})W_n$ by $(t^{e_n}W_n) \in H \times \langle t \rangle \setminus H$. Applying Corollary 2.4.10 to the resulting factorization, we see that at least one of $W_1, \ldots, W_{n-1}$ lies in H. Thus

$$W \equiv \cdots t^{e_i} W_i t^{e_{i+1}} \cdots = 1$$

where $W_i \in H$, $1 \leq i \leq n-1$. If e_i and e_{i+1} are of opposite sign, then we have our pinch, so we are done. If e_i and e_{i+1} are of the same sign, consider the equivalent word

$$W' \equiv \cdots t^{e_i+e_{i+1}} W_i W_{i+1} \cdots = 1.$$

Since W' contains one less power of t, it follows by induction that W' contains a pinch. But then W contains a pinch. $\square$

We have now proved the special case of Britton's Lemma for commuting HNN extensions (Lemma 2.5.3). The reader who is impatient to see the proof of the full Britton's Lemma may skip to the next section. We now use the special case to prove the implication $3 \Rightarrow 2$ in Boone's Theorem.

Proof of $3 \Rightarrow 2$. Recall from the proof of Lemma 2.3.6 that the Boone group $G = G_5$ is a commuting HNN extension of G_4 with stable letter k, namely

$$G = \langle G_4, k \mid kx = xk, kr_i = r_ik, k(q^{-1}tq) = (q^{-1}tq)k, i \in I \rangle,$$

where G_4 is the subgroup of G generated by the generators other than k. Moreover, G_4 is a commuting HNN extension of G_3 with stable letter t, namely

$$G_4 = \langle G_3, t \mid tx = xt, tr_i = r_it, i \in I \rangle,$$

where G_3 is the subgroup of G_4 generated by its generators other than t.

Assume 3, i.e., $k(\Sigma^{-1}t\Sigma) = (\Sigma^{-1}t\Sigma)k$. By Britton's Lemma with stable letter k, $\Sigma^{-1}t\Sigma$ belongs to the subgroup generated by $x, r_i, q^{-1}tq, i \in I$. Thus there is an equation

$$W \equiv \Sigma^{-1}t\Sigma R_0(q^{-1}t^{e_1}q)R_1 \cdots R_{n-1}(q^{-1}t^{e_n}q)R_n = 1$$

where the R_j are (possibly empty) words on $x, x^{-1}, r_i, r_i^{-1}, i \in I$, and $e_j = \pm 1$. Choose this equation so that n is as small as possible. By Britton's Lemma with stable letter t, W contains a pinch t^eXt^{-e} where $e = \pm 1$ and $X = R$ for some word R on $x, x^{-1}, r_i, r_i^{-1}, i \in I$. There are two cases.

Case 1: t^e is the first occurrence of t in W. Thus $t^eXt^{-e} \equiv t\Sigma R_0 q^{-1} t^{e_1}$. Hence $e = 1$, $e_1 = -1$, and $X \equiv \Sigma R_0 q^{-1}$. Since $X = R$, we have $\Sigma = RqR_0^{-1}$, which gives us 2 in Boone's Theorem.

Case 2: $t^e X t^{-e} \equiv t^{e_j} q R_j q^{-1} t^{e_{j+1}}$ for some j, $1 \le j \le n-1$. Hence $e_j = e$, $e_{j+1} = -e$ and $X \equiv q R_j q^{-1}$. We then have

$$\begin{aligned} q^{-1} t^{e_j} q R_j q^{-1} t^{e_{j+1}} q &= q^{-1} t^{e_j} C t^{e_{j+1}} q \\ &= q^{-1} t^{e_j} R t^{e_{j+1}} q \\ &= q^{-1} R q \\ &= q^{-1} X q \\ &= q^{-1} q R_j q^{-1} q = R_j \end{aligned}$$

so in W we may replace $R_{j-1} q^{-1} t^{e_j} q R_j q^{-1} t^{e_{j+1}} q R_{j+1}$ by $R_{j-1} R_j R_{j+1}$ contradicting minimality of n. This completes the proof of $3 \Rightarrow 2$. □

2.6 Proof of Britton's Lemma

Having proved a special case of Britton's Lemma, we now use it to prove the full lemma.

Lemma 2.6.1 (Britton's Lemma). Let

$$G^* = \langle G, p \mid p^{-1} X p = \phi(X) \rangle_X$$

be an HNN extension of G with stable letter p (see Definition 2.3.3). If W is a word involving p or p^{-1}, and if $W = 1$ in G^*, then W contains a pinch.

Proof. If W has a subword of the form $p^{-1}p$ or pp^{-1}, we are done. Assume this is not the case, i.e.,

$$W \equiv W_0 p^{e_1} W_1 \cdots W_{n-1} p^{e_n} W_n = 1$$

where $n \ge 1$, $e_1, \ldots, e_n \ne 0$, $W_0, \ldots, W_n$ are words on $A \cup A^{-1}$, and $W_1, \ldots, W_{n-1}$ are nonempty.

Introduce a new letter t, and form

$$G^{*\prime} = \langle A, p, t \mid R, p^{-1} X p = \phi(X), t^{-1} X t = X \rangle_X$$

which is a commuting HNN extension of G^* with stable letter t. In $G^{*\prime}$ we have $(tp)^{-1} X (tp) = p^{-1} t^{-1} X t p = p^{-1} X p = \phi(X)$, so there is a homomorphism $\psi : G^* \to G^{*\prime}$ given by $a \mapsto a$, $p \mapsto tp$, $a \in A$. Thus in $G^{*\prime}$ we have

$$W' = W_0 (tp)^{e_1} W_1 \cdots W_{n-1} (tp)^{e_n} W_n = \psi(W) = 1.$$

Applying Lemma 2.5.3 to W', we see that W' contains a "special pinch," i.e., a subword of the form $t^{-1} Y t$ or $t Y t^{-1}$ where Y is a word on $A \cup A^{-1} \cup \{p\}$ lying in H.

If our special pinch is $t^{-1} Y t$, then we have $e_i < 0$, $e_{i+1} > 0$, and $Y \equiv W_i$ for some i, $1 \le i \le n-1$. Since Y lies in H, W_i lies in H. Going back to G^*, we see that W has a subword $p^{-1} W_i p$ and this is a pinch.

If our special pinch is tYt^{-1}, then we have $e_i > 0$, $e_{i+1} < 0$, and $Y \equiv pW_ip^{-1}$ for some i, $1 \leq i \leq n-1$. Since Y lies in H, $W_i = p^{-1}Yp$ lies in K. Going back to G^*, we see that W has a subword pW_ip^{-1} and this is a pinch.

This completes the proof of Britton's Lemma. $\square$

We shall also need the following generalization.

Definition 2.6.2 (multiple HNN extension). Let $G = \langle A \mid R \rangle$ be a group presentation. Assume that we have isomorphisms $\phi_i : H_i \cong K_i$, $i \in I$, where H_i and K_i are subgroups of G. Consider the group presentation

$$G^* = \langle A, p_i, i \in I \mid R, p_i^{-1}X_ip_i = \phi_i(X_i) \rangle$$

where X_i, $\phi_i(X_i)$ range over generators of H_i, K_i respectively. We call this a *multiple HNN extension* with stable letters p_i, $i \in I$.

Lemma 2.6.3 (multiple Britton Lemma). Let G^* be a multiple HNN extension of G as above. If $W = 1$ in G^* and W involves at least one stable letter, then W contains a *pinch*, i.e., a subword $p_i^{-1}Xp_i$ or $p_iXp_i^{-1}$ where X is a word on $A \cup A^{-1}$ lying in H_i or K_i respectively.

Proof. Let $p_1, \ldots, p_n$ be the stable letters occurring in W. We proceed by induction on n. We may assume that $G = G_0 \subseteq G_1 \subseteq \cdots \subseteq G_n = G^*$ where, for each $i = 0, \ldots, n-1$, $G_{i+1} = \langle G_i, p_i \mid \ldots \rangle$ is an HNN extension of G_i with stable letter p_i. By Britton's Lemma with stable letter p_n, W contains a subword $p_n^{-1}Xp_n$ or $p_nXp_n^{-1}$ where X is a word on $A, A^{-1}, p_1, p_1^{-1}, \ldots, p_{n-1}, p_{n-1}^{-1}$ and X lies in H_n or K_n respectively. If X does not involve $p_1, \ldots, p_{n-1}$, then we have our pinch. Otherwise, let Z be a word on $A \cup A^{-1}$ such that $X = Z$ in G_{n-1}. Then $XZ^{-1} = 1$ in G_{n-1}, so by inductive hypothesis XZ^{-1} contains a pinch. But Z^{-1} is a word on $A \cup A^{-1}$ only, so X contains a pinch. $\square$

2.7 Proof of $2 \Rightarrow 1$

We now complete the proof of Boone's Theorem 2.2.10.

Proof of $2 \Rightarrow 1$. Assume 2, i.e.,

$$L\overline{X}q_mYR = q\,,$$

where X and Y are words on A, and L and R are words on $x, x^{-1}, r_i, r_i^{-1}, i \in I$. Note that our equation takes place in G_3, the subgroup of the Boone group generated by $A, q, q_0, \ldots, q_l, x, r_i, i \in I$. Recall also from the proof of Lemma 2.3.6 that G_3 is a multiple HNN extension of G_2 with stable letters $r_i, i \in I$. Here G_2 is the subgroup generated by $A, q, q_0, \ldots, q_l, x$.

We may safely assume that L, R are *freely reduced*, i.e., they do not contain subwords of the form $x^{-1}x$, xx^{-1}, $r_i^{-1}r_i$, $r_ir_i^{-1}$, $i \in I$. Using this assumption, we have:

Lemma 2.7.1. L and R are $\{r_i \mid i \in I\}$-reduced.

Proof. Otherwise, L or R contains a pinch of the form $r_i^{-1}x^e r_i$ or $r_i x^e r_i^{-1}$ where $e \neq 0$. Thus, it suffices to show that x^e, $e \neq 0$, cannot belong to the subgroup H_i generated by $\overline{X_i} q_{m_i} Y_i, ax, a \in A$, or to the subgroup K_i generated by $\overline{U_i} q_{n_i} V_i, ax^{-1}, a \in A$.

In the first case, suppose

$$W \equiv W_0(\overline{X_i} q_{m_i} Y_i)^{e_1} W_1 \cdots W_{n-1}(\overline{X_i} q_{m_i} Y_i)^{e_{n-1}} W_n = x^e$$

where $e_j = \pm 1$ and each W_j is a (possibly empty) word on $ax, (ax)^{-1}, a \in A$. This takes place in the free product $G_2 = G_1 * \langle q, q_0, \ldots, q_l \rangle$ where

$$G_1 = \langle x, a, a \in A \mid a^{-1}xa = x^2, a \in A \rangle$$

(see the proof of Lemma 2.3.6). Therefore, by Corollary 2.4.5, W cannot involve q_{m_i}. Hence $n = 0$ and $W \equiv W_0$. Thus we have

$$x^{-e}W_0 \equiv x^{-e}(a_1 x)^{f_1} \cdots (a_k x)^{f_k} = 1$$

where $a_j \in A$ and $f_j = \pm 1$. By the multiple Britton Lemma 2.6.3 with stable letters $a, a \in A$, we see that $x^{-e}W_0$ contains a pinch of the form $a^{-1}x^n a$ or $ax^{2n}a^{-1}$ for some $a \in A$. By inspection of $x^{-e}W_0$, both forms are impossible. Hence W_0 is empty, hence $x^e = 1$, hence $e = 0$, a contradiction.

The second case is similar. This proves the lemma. □

We continue with the proof of $2 \Rightarrow 1$. We are assuming $L\overline{X} q_m Y R = q$ in the Boone group, and we wish to obtain $X q_m Y = q$ in the Post semigroup.

Let N be the number of occurences of $r_i, r_i^{-1}, i \in I$ in L and R. We proceed by induction on N. If $N = 0$, we have $L = x^s$, $R = x^t$, and our assumption becomes

$$x^s \overline{X} q_m Y x^t = q.$$

Since no r_i appears, this holds in the free product $G_1 * \langle q, q_0, \ldots, q_l \rangle$, and clearly $x^s\overline{X}$ and Yx^t belong to G_1. By Corollary 2.4.5, it follows that $q_m = q$ and $x^s\overline{X} = Yx^t = 1$. Hence $s = t = 0$ and X and Y are empty, so our conclusion $X q_m Y = q$ in S trivially holds.

Assume now that $N > 0$. Hence, by the multiple Britton Lemma 2.6.3 with stable letters $r_i, i \in I$, there is a pinch in $L\overline{X} q_m Y R$. By Lemma 2.7.1, L and R are $\{r_i \mid i \in I\}$-reduced, i.e., they individually do not contain a pinch. Hence there must be a pinch which spans L and R. It follows that

$$L\overline{X} q_m Y R \ \equiv \ L' r_i^e x^s \overline{X} q_m Y x^t r_i^{-e} R'$$

where $e = \pm 1$, $L \equiv L' r_i^e x^s$, $R \equiv x^t r_i^{-e} R'$, and $r_i^e x^s \overline{X} q_m Y x^t r_i^{-e}$ is a pinch.

If $e = -1$, then $x^s \overline{X} q_m Y x^t$ lies in the subgroup H_i generated by $\overline{X_i} q_{m_i} Y_i$, ax, $a \in A$. If $e = 1$, then $x^s \overline{X} q_m Y x^t$ lies in the subgroup K_i generated by $\overline{U_i} q_{n_i} V_i$, ax^{-1}, $a \in A$. Since we are in the free product $G_1 * \langle q, q_0, \ldots, q_l \rangle$, it is clear that $m = m_i$, hence $q_m = q_{m_i}$. We consider only the case $e = -1$, the other case being similar.

Since $x^s\overline{X}q_mYx^t$ lies in H_i, there exists an equation

$$W \equiv x^s\overline{X}q_mYx^tW_0(\overline{X_i}q_mY_i)^{e_1}W_1\cdots W_{n-1}(\overline{X_i}q_mY_i)^{e_n}W_n = 1$$

where $e_j = \pm 1$ and W_j is a possibly empty, freely reduced word on $ax, (ax)^{-1}, a \in A$. Choose this equation so that n is as small as possible. Since our equation $W = 1$ holds in the free product $G_1 * \langle q_m \rangle$, it follows that $1 + e_1 + \cdots + e_n = 0$ and, by Corollary 2.4.5, each of the words between two consecutive occurrences of q_m or q_m^{-1} are $= 1$ in G_1. In particular, if $e_j = 1$ and $e_{j+1} = -1$, then $Y_iW_jY_i^{-1} = 1$, and if $e_j = -1$ and $e_{j+1} = 1$, then $\overline{X_i}^{-1}W_j\overline{X_i} = 1$. Either way, $W_j = 1$, hence $(\overline{X_i}q_mY_i)^{e_j}W_j(\overline{X_i}q_mY_i)^{e_{j+1}} = 1$ contradicting minimality of n. Therefore, we must have $e_1 = \cdots = e_n$. Since $1 + e_1 + \cdots + e_n = 0$, it follows that $n = 1$ and $e_1 = -1$. We now have

$$\begin{aligned} W &\equiv x^s\overline{X}q_mYx^tW_0(\overline{X_i}q_mY_i)^{-1}W_1 \\ &\equiv x^s\overline{X}q_mYx^tW_0Y_i^{-1}q_m^{-1}\overline{X_i}^{-1}W_1 = 1 \end{aligned}$$

in the free product $G_1 * \langle q_m \rangle$. It follows by Corollary 2.4.5 that $Yx^tW_0Y_i^{-1} = 1$, hence $x^s\overline{X}\,\overline{X_i}^{-1}W_1 = 1$.

Lemma 2.7.2. Y_i is an initial segment of Y, and X_i is a final segment of X.

Proof. We first show that Y_i is an initial segment of Y. Let $Y' \equiv Y_i^{-1}Y$ after cancelling subwords of the form $a^{-1}a$ for all $a \in A$. It suffices to show that the first letter of Y' is *positive* (i.e., an element of A, not of A^{-1}). If not, let $b^{-1} \in A^{-1}$ be the first letter of Y', and consider $x^tW_0Y' = x^tW_0Y_i^{-1}Y = 1$. Applying the multiple Britton Lemma with stable letters $a, a \in A$, we see that x^tW_0Y' contains a pinch a^eZa^{-e}, where $e = \pm 1$ and Z lies in $\langle x \rangle$. Since W_0 is a freely reduced word on $ax, (ax)^{-1}, a \in A$, our pinch is not contained in x^tW_0. Hence a^{-e} must be the first letter of Y'. Our pinch is then $a^eZa^{-e} \equiv bxb^{-1}$, hence x belongs to the subgroup of $\langle x \rangle$ generated by x^2, a contradiction.

We have now proved that Y_i is an initial segment of Y. The proof that X_i is a final segment of X is similar. □

By the previous lemma, let $Y = Y_iY'$ and $X = X'X_i$, where X' and Y' are words on A. We then have

$$W_0Y'x^t = W_0YY_i^{-1}x^t = 1\,, \qquad x^s\overline{X'}W_1 = x^s\overline{X}\,\overline{X_i}^{-1}W_1 = 1.$$

Consider the automorphism ψ of G_1 given by $\psi(a) = a$ for $a \in A$, and $\psi(x) = x^{-1}$. In particular, for all $a \in A$ we have $\psi(ax) = ax^{-1} = r_i^{-1}axr_i$. Since W_0 and W_1 are words on $ax, (ax)^{-1}, a \in A$, we have

$$\psi(W_0) = r_i^{-1}W_0r_i\,, \qquad \psi(W_1) = r_i^{-1}W_1r_i.$$

Moreover,

$$\psi(W_0 Y' x^t) = \psi(W_0) Y' x^{-t} = 1\,, \qquad \psi(x^s \overline{X'} W_1) = x^{-s} \overline{X'} \psi(W_1) = 1.$$

We now have:

$$\begin{aligned}
q &= L\overline{X} q_m Y R \\
&= L' r_i^{-1} x^s \overline{X} q_{m_i} Y x^t r_i R' \\
&= L' r_i^{-1} x^s \overline{X'}\, \overline{X_i} q_{m_i} Y_i Y' x^t r_i R' \\
&= L' r_i^{-1} W_1^{-1} \overline{X_i} q_{m_i} Y_i W_0^{-1} r_i R' \\
&= L' \psi(W_1)^{-1} r_i^{-1} \overline{X_i} q_{m_i} Y_i r_i \psi(W_0)^{-1} R' \\
&= L' \psi(W_1)^{-1} \overline{U_i} q_{n_i} V_i \psi(W_0)^{-1} R' \\
&= L' x^{-s} \overline{X'}\, \overline{U_i} q_{n_i} V_i Y' x^{-t} R'.
\end{aligned}$$

Note that $L'x^{-s}$ and $x^{-t}R'$ are words on $x, x^{-1}, r_i, r_i^{-1}, i \in I$ with $N-2$ occurrences of $r_i, r_i^{-1}, i \in I$. Hence, by induction hypothesis, $X'U_i q_{n_i} V_i Y' = q$ in the semigroup S. Thus $Xq_mY = X'X_i q_{m_i} Y_i Y' = X'U_i q_{n_i} V_i Y' = q$ in S, and we have proved 1. □

This completes our proof of Boone's Theorem 2.2.10. Thus we have proved that the word problem for groups is unsolvable.

2.8 Some Refinements

In this section we state without proof some refinements of Theorem 2.2.12 concerning unsolvability of the word problem for groups.

The following result is due to Higman. For a proof, see Aanderaa/Cohen [2] or Rotman [12, Chapter 12] or Shoenfield [13, Appendix].

Theorem 2.8.1 (Higman's Theorem). Let $G = \langle A \mid R \rangle$ be a recursively presented group, i.e., A and R are recursive. Then G is recursively embeddable in a finitely presented group.

This following result is due to C. Miller [8, Corollary 3.9]. The proof uses Higman's Theorem.

Theorem 2.8.2 (C. Miller). We can construct a finitely presented group G such that G and all nontrivial quotient groups of G have unsolvable word problem.

In another direction, let $G = \langle A \mid R \rangle$ be a finitely presented group, and consider the following sets of words on $A \cup A^{-1}$.

1. $S_1 = \{W \mid W = 1 \text{ in } G\}$.

2. $S_2 = \{W \mid W \neq 1 \text{ in some finite homomorphic image of } G\}$.

Remark 2.8.3. It is easy to see that S_1 and S_2 are disjoint and recursively enumerable. Therefore, if S_1 and S_2 are complementary (i.e., if G is *residually finite*), then S_1 and S_2 are recursive. To say that S_1 is recursive means exactly that the word problem for G is solvable.

The following result is due to Slobodskoi [17]. See also Kharlampovich [9].

Theorem 2.8.4 (Slobodskoi). We can construct a finitely presented group G such that both S_1 and S_2 are nonrecursive.

The following stronger result has been announced by Aanderaa [1].

Theorem 2.8.5 (Aanderaa). We can construct a finitely presented group G such that S_1 and S_2 are recursively inseparable.

2.9 Unsolvability of the Triviality Problem

In this section we consider group-theoretic problems of another kind, concerning not just a single group, but rather a family of groups.

Definition 2.9.1 (triviality problem). The *triviality problem for groups* is as follows.

> Given a finitely presented group $G = \langle A \mid R \rangle$, to decide whether G is the trivial group, i.e., $G = 1$.

Note that this is a Σ^0_1 problem, because A is finite, and $G = 1 \iff \forall a \in A\, \exists n\, \exists$ finite sequence of words such that $a \equiv W_0 \sim_R W_1 \sim_R \cdots \sim_R W_n \equiv 1$.

We shall show that the triviality problem for groups is unsolvable. This and similar results (see Corollary 2.9.10 below) are due to Adian 1955 and Rabin 1958. It turns out that these results follow fairly easily from the unsolvability of the word problem for groups.

Let G be a fixed, finitely presented group. We reduce the word problem for G to the triviality problem for finite presented groups. The reduction is given by the following definition and lemma.

Definition 2.9.2. Let $G = \langle A \mid R \rangle$ be a fixed, finitely presented group. Given a word W on $A \cup A^{-1}$, let $G'_W = \langle A' \mid R' \rangle$, where $A' = A \cup \{x, y, z\}$, and R' consists of R plus the relations

$$\begin{array}{lccll}
(1) & x^{-1}(W^{-1}y^{-1}Wy)x & = & z^{-1}yz & \\
(2) & x^{-2}(yxy)x^2 & = & z^{-2}yz^2 & \\
(3) & x^{-3}yx^3 & = & z^{-3}(yzy)z^3 & \\
(4) & x^{-3-i}(ya_iy)x^{3+i} & = & z^{-3-i}yz^{3+i}, & 1 \le i \le n,
\end{array}$$

where $A = \{a_1, \ldots, a_n\}$. Note that G'_W is a finitely presented group.

Lemma 2.9.3.

1. If $W \neq 1$ in G, then G embeds into G'_W.
2. If $W = 1$ in G, then G'_W is trivial.

Proof. Assume first that $W \neq 1$ in G. Within the free product $G * \langle x, y \rangle$, consider the subgroup H generated by y plus the left hand sides of equations (1)–(4). It is straightforward to check that H is free on these generators (use Corollary 2.4.5). Similarly, in the free group $\langle y, z \rangle$, consider the subgroup K generated by y plus the right hand sides of (1)–(4). Again, K is free on these generators. Thus, there is an obvious isomorphism $\theta : H \cong K$, and we have

$$G'_W \;\cong\; (G * \langle x, y \rangle) *_\theta \langle y, z \rangle,$$

i.e., G'_W is the free product of $G * \langle x, y \rangle$ and $\langle y, z \rangle$ with H and K amalgamated via θ. It follows that $G \hookrightarrow G * \langle x, y \rangle \hookrightarrow G'_W$.

Now assume $W = 1$ in G. Then $W^{-1}y^{-1}Wy = 1$ in G'_W, hence by (1) $y = 1$. Hence by (2) $x = 1$, by (3) $z = 1$, and by (4) $a_i = 1$, $1 \leq i \leq n$. We conclude that $G'_W = 1$. $\square$

Theorem 2.9.4 (unsolvability of the triviality problem). The triviality problem for finitely presented groups is unsolvable.

Proof. Let G be a finitely presented group such that word problem for G is unsolvable. Then $W = 1$ in G if and only if G'_W is trivial. Thus the word problem for G reduces to the triviality problem for finitely presented groups. Hence, the latter problem is unsolvable. $\square$

Using Theorem 2.9.4, S. Novikov has obtained the following undecidability result in geometry. We state this result without proof.

Theorem 2.9.5 (S. Novikov). Fix $n \geq 5$. If M is a finitely presented, compact, connected, n-dimensional manifold without boundary, then it is undecidable whether M is diffeomorphic to the n-sphere, S^n. Instead of diffeomorphic, we can say homeomorphic.

Remark 2.9.6. To each finitely presented, connected manifold M is associated a finitely presented group $\pi_1(M)$, the *fundamental group* of M, consisting of the homotopy classes of closed paths in M. It is well known that the fundamental group of the n-sphere, S^n, is trivial. Conversely, there is a theorem of Smale saying that, under certain circumstances, if the fundamental group of an n-dimensional manifold M is trivial, then $M \cong S^n$. Smale's result is used in the proof of S. Novikov's result. For an exposition of the proof of S. Novikov's result, see Nabutovsky [10, Appendix]. Nabutovsky has applied S. Novikov's result to draw some purely geometrical consequences.

One easily generalizes Theorem 2.9.4 as follows.

Definition 2.9.7. Let P be a property of groups which is invariant under isomorphism. We call P a *Markov property* if there exist finitely presented groups G_1, G_2 such that (1) G_1 has property P, (2) for any group $H \supseteq G_2$, H does not have property P.

Examples 2.9.8. Let $P =$ triviality, finiteness, Abelianness, solvability, nilpotence, etc. Each of these properties is a Markov property.

Theorem 2.9.9 (Adian, Rabin). Let P be a Markov property. Given a finitely presented group H, it is undecidable whether H has property P.

Proof. Fix a finitely presented group K with unsolvable word problem. Given a word W in K, form the finitely presented group

$$H_W = G_1 \times (K \times G_2)'_W \,.$$

If $W = 1$ in K, then $H_W = G_1$ has property P. If $W \neq 1$ in K, then $G_2 \hookrightarrow K \times G_2 \hookrightarrow (K \times G_2)'_W \hookrightarrow H_W$, so H_W does not have property P. Thus, the word problem for K is reducible to the problem of deciding whether a given finitely presented group has property P. Hence, the latter problem is unsolvable. $\square$

Corollary 2.9.10. Given a finite presented group H, it is undecidable whether H is trivial, finite, Abelian, solvable, nilpotent, etc.

Chapter 3

Recursively Enumerable Sets and Degrees

In this chapter we study the lattice of recursively enumerable sets of natural numbers, under inclusion. We also study the partial ordering of degrees of unsolvability of recursively enumerable sets of natural numbers, under Turing reducibility. A standard reference for these subjects is Soare [18]. A useful supplementary reference is Rogers [11].

3.1 The Lattice of R.E. Sets

The purpose of this section is to introduce the lattice of recursively enumerable sets. We begin by reviewing some basic properties of Σ^0_1 relations on $\mathbb{N}$, the set of natural numbers.

Definition 3.1.1. Recall that $R \subseteq \mathbb{N}^k$ is *recursive* if the characteristic function $\chi_R : \mathbb{N}^k \to \mathbb{N}$, defined by $\chi_R(x_1, \ldots, x_k) = 1$ if $R(x_1, \ldots, x_k)$ holds, 0 otherwise, is recursive.

Definition 3.1.2. Recall that $S \subseteq \mathbb{N}^k$ is Σ^0_1 if

$$S = \{\langle x_1, \ldots, x_k\rangle \in \mathbb{N}^k \mid \exists y\, R(x_1, \ldots, x_k, y)\}$$

where $R \subseteq \mathbb{N}^{k+1}$ is recursive.

Remark 3.1.3. In our definition of S being Σ^0_1, instead of saying that R is recursive, we could say that R is primitive recursive. Also, by Theorem 1.2.7, this is equivalent to S being Σ_1, i.e., we can say that R is Δ_0. Moreover, by Matiyasevich's Theorem 1.3.2, this is equivalent to S being Diophantine. However, we shall not make use of these results.

Proposition 3.1.4. S is recursive $\iff$ S is Δ^0_1, i.e., S and $\neg\, S$ are Σ^0_1.

Proof. The $\Longrightarrow$ direction is trivial. For the $\Longleftarrow$ direction, assume that S is Δ^0_1, say

$$S(x_1, \ldots, x_k) \equiv \exists y\, R_1(x_1, \ldots, x_k, y), \quad \neg\, S(x_1, \ldots, x_k) \equiv \exists y\, R_2(x_1, \ldots, x_k, y).$$

Let $f(x_1, \ldots, x_k) =$ the least y such that $R_1(x_1, \ldots, x_k, y) \vee R_2(x_1, \ldots, x_k, y)$. Then f is a recursive function, and $S(x_1, \ldots, x_k) \equiv R_1(x_1, \ldots, x_k, f(x_1, \ldots, x_k))$, hence S is recursive. $\square$

Proposition 3.1.5. If $S_1, S_2 \subseteq \mathbb{N}^k$ are Σ^0_1, then so are $S_1 \cup S_2$ and $S_1 \cap S_2$.

Proof. Let $S_i(x_1, \ldots, x_k) \equiv \exists y\, R_i(x_1, \ldots, x_k, y)$, $i = 1, 2$, where R_1, R_2 are recursive. We have

$$(S_1 \cup S_2)(x_1, \ldots, x_k) \equiv \exists y\, (R_1(x_1, \ldots, x_k, y) \vee R_2(x_1, \ldots, x_k, y))$$

and

$$(S_1 \cap S_2)(x_1, \ldots, x_k) \equiv \exists y\, (R_1(x_1, \ldots, x_k, (y)_1) \wedge R_2(x_1, \ldots, x_k, (y)_2))$$

so $S_1 \cup S_2$ and $S_1 \cap S_2$ are Σ^0_1. $\square$

The next proposition is known as the Σ^0_1 Uniformization Principle.

Proposition 3.1.6. Let $S \subseteq \mathbb{N}^{k+1}$ be Σ^0_1. Then there is a partial recursive function $\psi : \mathbb{N}^k \xrightarrow{P} \mathbb{N}$ such that

1. $\psi(x_1, \ldots, x_k)\downarrow \iff \exists y\, S(x_1, \ldots, x_k, y)$,
2. $\psi(x_1, \ldots, x_k)\downarrow \Longrightarrow S(x_1, \ldots, x_k, \psi(x_1, \ldots, x_k))$.

Proof. Let $S(x_1, \ldots, x_k, y) \equiv \exists z\, R(x_1, \ldots, x_k, y, z)$ where R is recursive. Put $\theta(x_1, \ldots, x_k) \simeq$ the least w such that $R(x_1, \ldots, x_k, (w)_0, (w)_1)$. Note that θ is a partial recursive function. Put $\psi(x_1, \ldots, x_k) \simeq (\theta(x_1, \ldots, x_k))_0$. $\square$

Proposition 3.1.7. $\psi : \mathbb{N}^k \xrightarrow{P} \mathbb{N}$ is partial recursive $\iff$ graph(ψ) is Σ^0_1.

Proof. $\Longleftarrow$: If the graph of ψ is Σ^0_1, let $S = \text{graph}(\psi)$, and apply the previous lemma to conclude that ψ is partial recursive.

$\Longrightarrow$: If ψ is partial recursive, let $\mathcal{P}$ be a program which computes ψ. Then $\psi(x_1, \ldots, x_k) \simeq \varphi^{(k)}_e(x_1, \ldots, x_k)$ where $e = \#(\mathcal{P})$. Thus $\psi(x_1, \ldots, x_k) \simeq y$ if and only if

$$\exists n\, ((\text{State}(e, x_1, \ldots, x_k, n))_0 = 0 \wedge (\text{State}(e, x_1, \ldots, x_k, n))_{k+1} = y)$$

where the State function is primitive recursive. (See the Math 558 notes [14].) Thus graph(ψ) is Σ^0_1. $\square$

Proposition 3.1.8. $S \subseteq \mathbb{N}^k$ is Σ^0_1 if and only if $S = \text{domain}(\psi)$ for some partial recursive function $\psi : \mathbb{N}^k \xrightarrow{P} \mathbb{N}$.

Proof. If S is Σ^0_1, say $S(x_1,\ldots,x_k) \equiv \exists y\, R(x_1,\ldots,x_k,y)$ where R is recursive, then we may take $\psi(x_1,\ldots,x_k) \simeq$ the least y such that $R(x_1,\ldots,x_k,y)$, and clearly this is partial recursive. Conversely, if ψ is partial recursive, say $\psi = \varphi^{(k)}_e$, then the domain of ψ is $\{\langle x_1,\ldots,x_k\rangle \in \mathbb{N}^k \mid \exists n\,(\mathrm{State}(e,x_1,\ldots,x_k,n))_0 = 0\}$ which is clearly Σ^0_1. □

The next proposition is known as the Σ^0_1 Reduction Principle.

Proposition 3.1.9. If $S_1, S_2 \subseteq \mathbb{N}^k$ are Σ^0_1, then we can find Σ^0_1 sets $S'_1, S'_2 \subseteq \mathbb{N}^k$ such that $S'_1 \subseteq S_1$, $S'_2 \subseteq S_2$, $S'_1 \cup S'_2 = S_1 \cup S_2$, and $S'_1 \cap S'_2 = \emptyset$.

Proof. Since S_1 and S_2 are Σ^0_1, we can express them as

$$S_1(x_1,\cdots,x_k) \equiv \exists y\, R_1(x_1,\ldots,x_k,y), \quad S_2(x_1,\ldots,x_k) \equiv \exists y\, R_2(x_1,\cdots,x_k,z)$$

where R_1, R_2 are recursive. Define S'_1 and S'_2 by

$$S'_1(x_1,\ldots,x_k) \equiv \exists y\,[\,R_1(x_1,\ldots,x_k,y) \wedge \neg\,\exists z < y\, R_2(x_1,\ldots,x_k,z)\,]\,,$$

$$S'_2(x_1,\ldots,x_k) \equiv \exists y\,[\,R_2(x_1,\cdots,x_k,z) \wedge \neg\,\exists z \leq y\, R_1(x_1,\ldots,x_k,z)\,]\,.$$

Clearly this works. Note the similarity to the proof of Rosser's Theorem. □

Corollary 3.1.10. If $P_1, P_2 \subseteq \mathbb{N}^k$ are Π^0_1, and if $P_1 \cap P_2 = \emptyset$, then there is a recursive $R \subseteq \mathbb{N}^k$ such that $P_1 \subseteq R$ and $P_2 \cap R = \emptyset$.

Proof. Let $S_1 = \mathbb{N}^k \setminus P_1$, $S_2 = \mathbb{N}^k \setminus P_2$, and apply the Reduction Principle 3.1.9. Then $S'_1 \cup S'_2 = S_1 \cup S_2 = \mathbb{N}^k$, $S'_1 \cap S'_2 = \emptyset$, hence by Proposition 3.1.4 S'_1, S'_2 are recursive. Set $R = S'_2$. □

Remark 3.1.11. The previous corollary is known as the Π^0_1 Separation Principle. On the other hand, there is no Σ^0_1 Separation Principle, as shown by the next proposition.

Definition 3.1.12. $S_1, S_2 \subseteq \mathbb{N}^k$ are said to be *recursively inseparable* if there is no recursive $R \subseteq \mathbb{N}^k$ such that $S_1 \subseteq R$ and $R \cap S_2 = \emptyset$.

Proposition 3.1.13. We can find Σ^0_1 sets $B_1, B_2 \subseteq \mathbb{N}$ such that $B_1 \cap B_2 = \emptyset$ and B_1, B_2 are recursively inseparable.

Proof. Put $B_i = \{e \mid \varphi^{(1)}_e(e) \simeq i\}$ for $i = 1, 2$. Clearly $B_1 \cap B_2 = \emptyset$ and B_1, B_2 are Σ^0_1. If B_1, B_2 were recursively separable, let $f : \mathbb{N} \to \{1, 2\}$ be recursive such that $f(e) = 2$ for all $e \in B_1$, and $f(e) = 1$ for all $e \in B_2$. Since f is recursive, $f = \varphi^{(1)}_e$ for some e. If $f(e) = 1$, then $\varphi^{(1)}_e(e) = 1$, which implies $e \in B_1$, which implies $f(e) = 2$, a contradiction. The contradiction is similar if we assume $f(e) = 2$. Thus B_1, B_2 are recursively inseperable. □

We now introduce the lattice of recursively enumerable sets.

Definition 3.1.14 (recursively enumerable sets). Let A be a subset of $\mathbb{N}$. We say that A is *recursively enumerable*, abbreviated *r.e.*, if it is either empty or the range of a recursive function $f : \mathbb{N} \to \mathbb{N}$.

Theorem 3.1.15. A is recursively enumerable $\iff$ A is Σ^0_1. Moreover, if A is recursively enumerable and infinite, then A is the range of a one-to-one recursive function.

Proof. Let $A = \text{range}(f)$ where $f : \mathbb{N} \to \mathbb{N}$ is recursive. Then $x \in A \iff \exists w\, f(w) = x$ and this is Σ^0_1. Now assume that A is infinite and Σ^0_1, say $A = \{x \mid \exists y\, R(x, y)\}$ where R recursive. Put

$$B = \{2^x 3^y \mid R(x, y) \wedge \neg\, \exists z < y\, R(x, z)\}.$$

Then B is infinite and recursive. Define $\pi_B : \mathbb{N} \to \mathbb{N}$ by $\pi_B(n) =$ the nth smallest element of B. This is known as the *principal function* of B. Clearly π_B is recursive, since we can obtain it by recursion as $\pi_B(0) =$ least element of B, $\pi_B(n+1) =$ least $w \in B$ such that $w > \pi_B(n)$. Now, let $f(n) = (\pi_B(n))_0$. Clearly f is one-to-one and recursive, and $\text{range}(f) = A$. □

The previous theorem says that r.e. sets are the same thing as Σ^0_1 sets. Thus we have the following properties of r.e. sets.

Theorem 3.1.16.

1. Let $A_1, A_2 \subseteq \mathbb{N}$ be recursively enumerable. Then we can find recursively enumerable sets $A'_1 \subseteq A_1$, $A'_2 \subseteq A_2$ such that $A'_1 \cup A'_2 = A_1 \cup A_2$ and $A'_1 \cap A'_2 = \emptyset$.

2. We can find recursively enumerable sets $B_1, B_2 \subseteq \mathbb{N}$ such that $B_1 \cap B_2 = \emptyset$ and B_1, B_2 are recursively inseparable.

Proof. Part 1 is a special case of the Reduction Principle 3.1.9. Part 2 is a restatement of Proposition 3.1.13. □

An algebraic context for results of this kind is lattice theory.

Definition 3.1.17 (lattices). A *lattice* is a partially ordered set $\mathcal{L} = (\mathcal{L}, \leq)$ in which any two elements have a least upper bound and a greatest lower bound.

Examples 3.1.18. We consider two familiar examples of lattices.

1. Consider the set of positive integers partially ordered by divisibility, i.e., $a \leq b \iff a$ divides b. This is a lattice. The l.u.b. and g.l.b. operations are just LCM and GCD.

2. Let X be a set. The powerset $P(X) = \{Y \mid Y \subseteq X\}$ is a lattice under inclusion, i.e., $Y \leq Z \iff Y \subseteq Z$. The l.u.b. and g.l.b. operations are given by $\cup$ and $\cap$.

Definition 3.1.19 (the lattice of r.e. sets). We write

$$\mathcal{E} = \{A \subseteq \mathbb{N} \mid A \text{ is recursively enumerable}\}.$$

By Proposition 3.1.5, $\mathcal{E}$ is a lattice under inclusion. The l.u.b. and g.l.b. operations are given by $\cup$ and $\cap$. We refer to $\mathcal{E}$ as *the lattice of r.e. sets*.

Definition 3.1.20 (lattice terminology). In an abstract lattice-theoretic context, the lattice operations l.u.b. and g.l.b. may be denoted $\vee$ and $\wedge$ respectively. If $\mathcal{L}$ is any lattice, we say that $\mathcal{L}$ is *distributive* if the laws $a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$ and $a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$ hold. All of the lattices considered in this chapter are distributive.

If a lattice $\mathcal{L}$ has a bottom element and a top element, they are denoted 0 and 1 respectively. For example, $P(X)$ and $\mathcal{E}$ are lattices with 0 and 1. The lattice of positive integers under divisibility has 0 but no 1.

Let $\mathcal{L}$ be a distributive lattice with 0 and 1. An element $a \in \mathcal{L}$ is said to be *complemented* within $\mathcal{L}$ if there exists $b \in \mathcal{L}$ (necessarily unique) such that $a \wedge b = 0$ and $a \vee b = 1$. The whole lattice $\mathcal{L}$ is said to be *complemented* if every element of $\mathcal{L}$ is complemented within $\mathcal{L}$. For example, the lattice $P(X)$ is complemented. A *Boolean algebra* is defined to be a complemented distributive lattice. Thus $P(X)$ is a Boolean algebra, but $\mathcal{E}$ is not.

Remark 3.1.21. Let $A \in \mathcal{E}$ be an r.e. set. By Proposition 3.1.4, A is complemented within $\mathcal{E}$ if and only if A is recursive. Since nonrecursive r.e. sets exist, it follows that the lattice $\mathcal{E}$ is noncomplemented. Theorem 3.1.16 above expresses further lattice-theoretic properties of $\mathcal{E}$.

Remark 3.1.22. Later in this chapter (Sections 3.6 and 3.7), we shall prove the following two theorems of Friedberg, which express yet more lattice-theoretic properties of $\mathcal{E}$. This is the beginning of a large subject.

1. If $A \subseteq \mathbb{N}$ is r.e. and not recursive, then we can find nonrecursive r.e. sets B_1, B_2 such that $A = B_1 \cup B_2$ and $B_1 \cap B_2 = \emptyset$. (Furthermore, we can demand that B_1, B_2 are recursively inseparable. According to Rogers [11, Exercise 12.21], this refinement is due to K. Ohashi.)

2. We can find a nonrecursive r.e. set $A \subseteq \mathbb{N}$ such for any r.e. set $B \supseteq A$, either $B \setminus A$ is finite or $\mathbb{N} \setminus B$ is finite. Such an r.e. set A is called a *maximal* r.e. set.

3.2 Many-One Completeness

A useful way to compare the recursion-theoretic complexity of subsets of $\mathbb{N}$, whether recursively enumerable or not, is via many-one reducibility.

Definition 3.2.1 (many-one reducibility). Let $A, B \subseteq \mathbb{N}$. We say that A is *many-one reducible* to B, abbreviated $A \leq_m B$, if there exists a recursive function $f : \mathbb{N} \to \mathbb{N}$ such that $\forall x\, (x \in A \iff f(x) \in B)$.

Definition 3.2.2 (m-completeness). An r.e. set C is said to be *many-one complete* if, for all r.e. sets A, $A \leq_m C$. We sometimes write *m-complete* as an abbreviation for many-one complete.

Example 3.2.3. The most straightforward example is as follows. Let

$$C = \{ 2^e 3^x \mid \varphi_e^{(1)}(x) \downarrow \} .$$

Clearly C is Σ_1^0, hence r.e. We claim that C is many-one complete. To see this, let A be an r.e. set. By Proposition 3.1.8, let $\psi(x)$ be a partial recursive function such that $A = \text{domain}(\psi)$. Let e be an index of ψ, i.e., the Gödel number of a program which computes ψ. (Using a notation to be introduced later, we can write $A = W_e$, i.e., e is an *index* of A.) For all $x \in \mathbb{N}$ we have $\psi(x) \simeq \varphi_e^{(1)}(x)$, hence $x \in A \iff \psi(x)\downarrow \iff \varphi_e^{(1)}(x)\downarrow \iff 2^e 3^x \in C$. Thus $A \leq_m C$ via the primitive recursive function $f(x) = 2^e 3^x$. We have now shown that C is m-complete.

In addition, we have the following examples.

Examples 3.2.4. Recall from Math 558 [14] the sets

$$H = \{e \in \mathbb{N} \mid \varphi_e^{(1)}(0)\downarrow\} = \text{the halting set}$$

and

$$K = \{e \in \mathbb{N} \mid \varphi_e^{(1)}(e)\downarrow\} = \text{the diagonal halting set}\,.$$

Clearly H and K are Σ_1^0, hence r.e.

Proposition 3.2.5. *H and K are many-one complete.*

Proof. Recall the Parametrization Theorem, which reads as follows. Given a partial recursive function $\theta(x, y)$, we can find a primitive recursive function $f(x)$ such that $\varphi_{f(x)}^{(1)}(y) \simeq \theta(x, y)$ for all x, y. (For a proof of the Paramatrization Theorem, see the Math 558 notes [14].) Given an r.e. set A, consider the partial recursive function $\theta(x, y) \simeq 1$ if $x \in A$, undefined otherwise. Apply the Parametrization Theorem to get a primitive recursive function $f(x)$ such that, for any y, $\varphi_{f(x)}^{(1)}(y)\downarrow \iff x \in A$. Setting $y = 0$, we see that $x \in A \iff f(x) \in H$. Setting $y = f(x)$, we see that $x \in A \iff f(x) \in K$. Thus $A \leq_m H$ and $A \leq_m K$ via f. $\square$

Examples 3.2.6. In Chapters 1 and 2 we considered several mathematical problems including Hilbert's Tenth Problem, the Word Problem for groups, and the Triviality Problem for groups. We pointed out that these each of these problems is Σ_1^0, i.e., recursively enumerable, and we proved that each of them is unsolvable, i.e., nonrecursive. More precisely, we showed how to many-one reduce the halting set H (or any other r.e. set) to each of them, via explicitly specified, primitive recursive functions. In particular, each of these problems is not only unsolvable but also many-one complete.

In a similar vein, one can show that many other well known unsolvable problems such as the Validity Problem for the predicate calculus, the Decision Problem for Z_1 (= first-order arithmetic), etc., are r.e. and many-one complete. It follows that each of these problems is many-one reducible to any of the others. In this sense, all of these problems are equivalent, i.e., they are all equally unsolvable.

Remark 3.2.7. Later in this chapter (see Sections 3.10–3.16), we shall study the general concept of *degrees of unsolvability*, due to Turing. From this point of view, the upshot of our examples above is that a great many unsolvable problems including the Halting Problem, Hilbert's Tenth Problem, the Word Problem for groups, the Validity Problem for predicate calculus, etc., are all of the same degree of unsolvability.

3.3 Creative Sets

In this section we define an interesting class of r.e. sets, the creative sets. We then prove a theorem due to Myhill 1955, which says that an r.e. set is creative if and only if it is many-one complete.

Notation 3.3.1. Let

$$W_e \;=\; \mathrm{domain}(\varphi_e^{(1)}) \;=\; \{x \in \mathbb{N} \mid \varphi_e^{(1)}(x)\downarrow\}\,.$$

By Proposition 3.1.8, the sequence W_e, $e = 0, 1, 2, \ldots$ is an enumeration of all the r.e. sets. We refer to this as the *standard enumeration* of the r.e. sets. Given an r.e. set A, an *index* or *r.e. index* of A is any $e \in \mathbb{N}$ such that $A = W_e$. Clearly any r.e. set has infinitely many indices.

Definition 3.3.2 (creative sets). An r.e. set C is said to be *creative* if there exists a partial recursive function $\psi(e)$ such that for all e, if $W_e \cap C = \emptyset$, then $\psi(e)\downarrow$ and $\psi(e) \notin W_e \cup C$. We call ψ a *creative function* for C.

Proposition 3.3.3. If C is creative, then C is not recursive.

Proof. If C were recursive, then $\mathbb{N} \setminus C$ would be recursively enumerable, say $\mathbb{N} \setminus C = W_e$. Then $W_e \cap C = \emptyset$, hence $\psi(x)\downarrow$ and $\psi(e) \notin W_e \cup C$, a contradiction since $W_e \cup C = \mathbb{N}$. □

Remark 3.3.4. We have just proved that creative sets are nonrecursive. In addition, we can say that a creative set C is "effectively nonrecursive." By this we mean that C is r.e. and nonrecursive and furthermore, the nonrecursiveness holds because of a computable function $\psi(e)$ which effectively provides a witness for the fact that W_e is not the complement of C, for all r.e. sets W_e.

Example 3.3.5. The diagonal halting set K of Example 3.2.4 is creative. Namely, a creative function for K is the identity function, $\psi(e) = e$ for all e. To see this, note that by definition $K = \{e \mid e \in W_e\}$. Hence, for all e, if $W_e \cap K = \emptyset$, then $e \notin W_e$ and $e \notin K$.

Exercise 3.3.6. Show that the r.e. sets C and H of Examples 3.2.3 and 3.2.4 are creative.

Exercise 3.3.7. Consider the set of Gödel numbers of sentences which are provable in the theory Z_1, first-order arithmetic, a.k.a., Peano Arithmetic. Show that this set is creative. Instead of Z_1, we could use any recursively axiomatizable theory to which Rosser's Theorem applies.

Theorem 3.3.8. Let A and B be r.e. sets. If A is creative and $A \leq_m B$, then B is creative.

Proof. Assume that A is creative via ψ, and assume that $A \leq_m B$ via f. By the Parametrization Theorem, let $h(x)$ be a primitive recursive function such that $\varphi^{(1)}_{h(x)}(y) \simeq \varphi^{(1)}_x(f(y))$ for all x, y. It follows that $W_{h(e)} = f^{-1}(W_e)$ for all e. Now, if $W_e \cap B = \emptyset$, then $W_{h(e)} \cap A = \emptyset$, hence $\psi(h(e)) \downarrow$ and $\psi(e) \notin W_{h(e)} \cup A$, hence $f(\psi(h(e))) \notin W_e \cup B$. Thus, a creative function for B is given by $e \mapsto f(\psi(h(e)))$. $\square$

Corollary 3.3.9. Let C be an r.e. set. If C is m-complete, then C is creative.

Proof. We have already seen that creative r.e. sets exist. For example, we have seen that K is creative. If C is m-complete, then $K \leq_m C$, hence by the previous theorem C is creative. $\square$

Our next goal is to prove the converse: if C is creative, then C is m-complete.

Lemma 3.3.10. If C is creative, then we can find a total recursive function $p(e)$ which is a creative function for C.

Proof. Let $\psi(e)$ be a creative function for A which is partial recursive. Consider the Σ^0_1 predicate

$$S(e, x) \;\equiv\; \psi(e) \simeq x \,\vee\, W_e \cap C \neq \emptyset\ .$$

Clearly $\forall e\, \exists x\, S(e, x)$. By Σ^0_1 uniformization (Proposition 3.1.6), we can find a total recursive function $p(e)$ such that $\forall e\, S(e, p(e))$ holds. Then $p(e)$ is a total creative function for C. $\square$

In order to prove that creative sets are m-complete, we need a mysterious and powerful theorem known as the Recursion Theorem.

Theorem 3.3.11 (Recursion Theorem). Let $\theta(w, x_1, \ldots, x_k)$ be a partial recursive function. Then we can find e such that

$$\varphi^{(k)}_e(x_1, \ldots, x_k) \simeq \theta(e, x_1, \ldots, x_k)$$

for all $x_1, \ldots, x_k$.

Example 3.3.12. We can find an e such that $\varphi^{(1)}_e(x) = e + x$ for all x. In particular, $\varphi^{(1)}_e(0) = e$, i.e., e is the Gödel number of a program which outputs e. Thus, there is a program which outputs its own Gödel number.

Actually, we need an even more powerful result, namely a uniform version of the Recursion Theorem. Here "uniform" means "parametrized."

Theorem 3.3.13 (Uniform Recursion Theorem). Let $\theta(w, y, x_1, \ldots, x_k)$ be a partial recursive function. Then we can find a primitive recursive function $h(y)$ such that

$$\varphi^{(k)}_{h(y)}(x_1, \ldots, x_k) \simeq \theta(h(y), y, x_1, \ldots, x_k)$$

for all $y, x_1, \ldots, x_k$.

Example 3.3.14. We can find a primitive recursive function $h(y)$ such that $\varphi^{(1)}_{h(y)}(x) \simeq h(y) + y + x$ for all y, x.

Remark 3.3.15. The Recursion Theorem follows easily from the Uniform Recursion Theorem, by treating the parameter y as a dummy variable.

Proof of the Uniform Recursion Theorem. We are given a partial recursive function $\theta(w, y, x_1, \ldots, x_k)$. Use the Parametrization Theorem to find a primitive recursive function $f(w, y)$ such that

$$\varphi^{(k)}_{f(w,y)}(x_1, \ldots, x_k) \simeq \theta(w, y, x_1, \ldots, x_k)$$

for all $y, x_1, \ldots, x_k$. In the same way, find a primitive recursive function $d(z)$ such that

$$\varphi^{(k)}_{d(z)}(x_1, \ldots, x_k) \simeq \varphi^{(k)}_{\varphi^{(1)}_z(z)}(x_1, \ldots, x_k)$$

for all $z, x_1, \ldots, x_k$. Here the expression on the right hand side is assumed to be undefined, if $\varphi^{(1)}_z(z)$ is undefined. Finally, let $g(y)$ be a primitive recursive function such that

$$\varphi^{(1)}_{g(y)}(z) \simeq f(d(z), y)$$

for all y, z. We then have

$$\begin{aligned}
\varphi^{(k)}_{d(g(y))}(x_1, \ldots, x_k) &\simeq \varphi^{(k)}_{\varphi^{(1)}_{g(y)}(g(y))}(x_1, \ldots, x_k) \\
&\simeq \varphi^{(k)}_{f(d(g(y)),y)}(x_1, \ldots, x_k) \\
&\simeq \theta(d(g(y)), y, x_1, \ldots, x_k)
\end{aligned}$$

so we may set $h(y) = d(g(y))$. ◻

We are now ready to prove the following result of Myhill.

Theorem 3.3.16. An r.e. set is creative if and only if it is many-one complete.

Proof. We have already seen in Corollary 3.3.9 that m-complete sets are creative. It remains to prove that creative sets are m-complete. Let C be a creative set. By Lemma 3.3.10, let $p(e)$ be a total recursive function which is a creative function for C. Let A be any r.e. set. We wish to show that $A \leq_m C$. Consider the partial recursive function $\theta(w, y, x) \simeq 1$ if $p(w) = x$ and $y \in A$, undefined otherwise. By the Uniform Recursion Theorem, let $h(y)$ be a primitive recursive function such that $\varphi^{(1)}_{h(y)}(x) \simeq \theta(h(y), y, x)$ for all y, x. Thus $W_{h(y)} = \{p(h(y))\}$ if $y \in A$, and $W_{h(y)} = \emptyset$ if $y \notin A$. For $y \in A$ we have $p(h(y)) \in W_{h(y)}$, hence $W_{h(y)} \cap C \neq \emptyset$, hence $p(h(y)) \in C$. For $y \notin A$ we have $W_{h(y)} = \emptyset$, hence $p(h(y)) \notin C$. Thus $A \leq_m C$ via the total recursive function $f(y) = p(h(y))$. This shows that C is m-complete. ◻

Myhill 1955 has also obtained the following result, which says that any two creative sets (or equivalently, m-complete sets) are *recursively isomorphic* to each other.

Theorem 3.3.17. If C_1 and C_2 are creative sets, then there is a total recursive function $\pi : \mathbb{N} \stackrel{\text{1-1 onto}}{\longrightarrow} \mathbb{N}$ such that $\pi(C_1) = C_2$. Note that π is a recursive permutation of the natural numbers.

Proof. We omit the proof. □

Remark 3.3.18. Using Myhill's results, we see that all of the specific, non-recursive r.e. sets mentioned in Section 3.2 are not only of the same degree of unsolvability, but also recursively isomorphic to each other.

3.4 Simple Sets

All of the nonrecursive r.e. sets which we have encountered so far are many-one complete, and hence creative. Nevertheless, there exist nonrecursive r.e. sets which are not creative. We now show one method for constructing such sets.

Definition 3.4.1 (simple sets). An r.e. set A is said to be *simple* if its complement

$$\overline{A} \;=\; \neg A \;=\; \mathbb{N} \setminus A$$

is infinite yet does not include an infinite r.e. set.

Clearly a simple set cannot be recursive, because by Proposition 3.1.4 the complement of a recursive set is r.e.

Theorem 3.4.2. There exists a simple set.

Proof. Consider the Σ^0_1 relation

$$S(e,x) \;\equiv\; x > 2e \text{ and } x \in W_e\,.$$

By Σ^0_1 uniformization, let $\psi(e)$ be a partial recursive function which uniformizes $S(e,x)$. In particular, if W_e is infinite, then $\psi(e)\downarrow$ and $\psi(e) > 2e$. Let A be the range of ψ. Thus A is an r.e. set which has nonempty intersection with every infinite r.e. set. To prove that A is simple, it remains to show that $\overline{A}$ is infinite. This is so because $|A \cap \{0,1,\ldots,2x\}| \leq x$ for all x, which follows from the fact that each element of $A \cap \{0,1,\ldots,2x\}$ is of the form $\psi(e)$ for some $e < x$. □

Theorem 3.4.3. A creative set is not simple.

Proof. Let C be a creative set, and let p be a creative function for C. By the Parametrization Theorem, let $f(e,x)$ be a primitive recursive function such that that $W_{f(e,x)} = W_e \cup \{x\}$ for all e, x. Let e_0 be an index of the empty set, i.e., $W_{e_0} = \emptyset$. Extend this to a recursive sequence of indices $e_0, e_1, e_2, \ldots$ by putting $e_{n+1} = f(e_n, p(e_n))$ for all n. By induction we have

$$W_{e_n} = \{p(e_0), p(e_1), \ldots, p(e_{n-1})\}$$

and $W_{e_n} \cap C = \emptyset$, hence $p(e_n)\downarrow$ and $p(e_n) \notin W_{e_n} \cup C$, for all n. Thus

$$\{p(e_0), p(e_1), \ldots, p(e_n), \ldots\}$$

is an infinite r.e. subset of $\overline{C}$. Hence C is not simple. □

Corollary 3.4.4. There exist nonrecursive r.e. sets which are not creative, hence not many-one complete.

3.5 Lattice-Theoretic Properties

Definition 3.5.1. A property of r.e. sets is said to be *lattice-theoretic* if it is definable over the $\mathcal{E}$, the lattice of r.e. sets.

Remark 3.5.2. In the previous definition, we could have used any of the languages $\{\cap, \cup, \subseteq\}$ or $\{\cap, \cup\}$ or $\{\subseteq\}$ or $\{\cap\}$ or $\{\cup\}$ for $\mathcal{E}$, without changing which properties of r.e. sets are definable over $\mathcal{E}$. This is because

$$A \subseteq B \;\equiv\; A \cup B = B \;\equiv\; A \cap B = A$$

and

$$A \cup B = \text{ the unique } C \text{ such that } \forall D\,(C \subseteq D \Leftrightarrow (A \subseteq D \wedge B \subseteq D))$$

and

$$A \cap B = \text{ the unique } C \text{ such that } \forall D\,(C \supseteq D \Leftrightarrow (A \supseteq D \wedge B \supseteq D))\,.$$

Moreover, the top and bottom elements $\mathbb{N}$ and $\emptyset$ and the equality relation $=$ for the lattice $\mathcal{E}$ are definable in any of these languages.

Examples 3.5.3. The following properties of r.e. sets are lattice-theoretic.

1. A is recursive $\iff$ A is complemented, i.e.,

$$\exists B\,(A \cup B = \mathbb{N} \text{ and } A \cap B = \emptyset).$$

2. A is finite $\iff$ $\forall B\,(A \cap B$ is recursive).
3. A is simple $\iff$ A nonrecursive and $\forall B\,(B$ infinite $\Rightarrow A \cap B \neq \emptyset)$.
4. A and B are recursively inseparable $\iff$

$$\neg\,\exists R\,(R \text{ recursive} \wedge A \subseteq R \wedge R \cap B = \emptyset).$$

Here of course the quantifiers range over $\mathcal{E}$.

The following surprising theorem is due to Harrington.

Theorem 3.5.4. An r.e. set A is creative if and only if

$$(\exists C \supseteq A)\,(\forall B \subseteq C)\,(\exists R)\,[\,R \text{ recursive, } R \cap C \text{ nonrecursive, } R \cap A = R \cap B\,]$$

where the quantifiers range over $\mathcal{E}$. Thus, the property of being creative is lattice-theoretic.

Proof. We omit the proof. □

3.6 The Friedberg Splitting Theorem

The purpose of this section is to prove the following splitting theorem, due essentially to Friedberg but with a small refinement due to Ohashi. We have previously stated this result as the first item in Remark 3.1.22.

Theorem 3.6.1. Let A be a nonrecursive r.e. set. Then we can find r.e. sets B_1, B_2 such that $A = B_1 \cup B_2$, $B_1 \cap B_2 = \emptyset$, and B_1, B_2 are recursively inseparable. It follows that B_1 and B_2 are nonrecursive.

Remark 3.6.2. Note that this statement is lattice-theoretic.

In order to present the proof, we first introduce some notation.

Notation 3.6.3 (finite approximation). Recall from the Math 558 notes [14] that $\varphi_e^{(1)}(x) \simeq y$ if and only if

$$\exists n\,[\,(\mathrm{State}(e,x,n))_0 = 0 \wedge (\mathrm{State}(e,x,n))_2 = y\,]\,.$$

We now introduce the finite approximation $\varphi_{e,s}^{(1)}(x) \simeq y$ if and only if

$$e,x,y < s \wedge \exists n < s\,[\,(\mathrm{State}(e,x,n))_0 = 0 \wedge (\mathrm{State}(e,x,n))_2 = y\,]\,.$$

Note that the 4-place relation $\varphi_{e,s}^{(1)}(x) \simeq y$ and the 3-place relation $\varphi_{e,s}^{(1)}(x)\downarrow$ are primitive recursive, and $\varphi_{e,s}^{(1)}(x) \simeq y$ implies $e,x,y < s$. Moreover,

$$\varphi_e^{(1)}(x) \simeq y \iff \exists s\ \varphi_{e,s}^{(1)}(x) \simeq y\,,$$

and

$$\varphi_e^{(1)}(x)\downarrow \iff \exists s\ \varphi_{e,s}^{(1)}(x)\downarrow\ .$$

In addition, there is a monotonicity property:

$$(\varphi_{e,s}^{(1)}(x) \simeq y \wedge s < t) \Rightarrow \varphi_{e,t}^{(1)}(x) \simeq y\,.$$

Recall also Notation 3.3.1, according to which $W_e = \mathrm{domain}(\varphi_e^{(1)})$. We now introduce the finite approximation

$$W_{e,s} = \mathrm{domain}(\varphi_{e,s}^{(1)})\,.$$

Again, the 3-place relation $x \in W_{e,s}$ is primitive recursive, and $x \in W_{e,s}$ implies $x, e < s$. Moreover $W_e = \bigcup_s W_{e,s}$. Also, $s < t$ implies $W_{e,s} \subseteq W_{e,t}$.

We now prove the Friedberg Splitting Theorem.

Proof of Theorem 3.6.1. We are given a nonrecursive r.e. set, A. Let $f : \mathbb{N} \to \mathbb{N}$ be a one-to-one, total recursive function such that $A = \mathrm{range}(f)$. Define $A^s = \{f(0), \ldots, f(s-1)\}$. We have $A = \bigcup_s A^s$. Also $s < t$ implies $A^s \subset A^t$.

For each $e \in \mathbb{N}$ and $i = 1, 2$ there will be a requirement

$$R_{2e+i} : B_i \cap W_e \neq \emptyset \text{ “if possible.”}$$

We order these requirements as

$$R_1, R_2, \ldots, R_{2e+1}, R_{2e+2}, \ldots$$

where lowered numbered requirements will receive higher priority.

The construction will consist of a definition of a recursive function

$$g : \mathbb{N} \to \{1, 2\}.$$

At stage $s + 1$ we shall define $g(s) = 1$ or $g(s) = 2$. We shall then define

$$B_1^{s+1} = \{f(n) \mid n \leq s, g(n) = 1\}, \qquad B_2^{s+1} = \{f(n) \mid n \leq s, g(n) = 2\},$$

beginning with $B_1^0 = B_2^0 = \emptyset$. At the end of the construction we shall define

$$B_1 = \bigcup_s B_1^s = \{f(n) \mid g(n) = 1\}, \qquad B_2 = \bigcup_s B_2^s = \{f(n) \mid g(n) = 2\}.$$

Obviously B_1, B_2 will be r.e. sets, because g is recursive. Moreover, this method of construction automatically guarantees that $A = B_1 \cup B_2$ and $B_1 \cap B_2 = \emptyset$.

The details of the construction are as follows.

Stage 0. $B_1^0 = B_2^0 = \emptyset$.

Stage s+1. At this stage we define $g(s)$, i.e., we decide whether $f(s)$ goes into B_1 or into B_2. Let e_s be the least e such that $f(s) \in W_{e,s}$ and either $B_1^s \cap W_{e,s} = \emptyset$ or $B_2^s \cap W_{e,s} = \emptyset$. If e_s is undefined, or if $B_1^s \cap W_{e,s} = \emptyset$, then define $g(s) = 1$, i.e., put $f(s)$ into B_1, i.e., $B_1^{s+1} = B_1^s \cup \{f(s)\}$, $B_2^{s+1} = B_2^s$. Otherwise, define $g(s) = 2$, i.e., put $f(s)$ into B_2, i.e., $B_2^{s+1} = B_2^s \cup \{f(s)\}$, $B_1^{s+1} = B_1^s$.

Note that, by construction, e_s takes on each possible value at most twice, so $\lim_s e_s = \infty$.

We claim that the construction gives r.e. sets B_1, B_2 which are recursively inseparable. To see this, assume for a contradiction that R is a recursive set separating B_1, B_2. We have $B_1 \subseteq R$ and $B_2 \cap R = \emptyset$. Let e and k be such that $W_e = R$ and $W_k = \mathbb{N} \setminus R$. Then $B_1 \cap W_k = \emptyset$ and $B_2 \cap W_e = \emptyset$. Hence for all s we have $B_1^s \cap W_{k,s} = \emptyset$ and $B_2^s \cap W_{e,s} = \emptyset$. For all sufficiently large s we have $e_s > e$ and $e_s > k$, hence by construction $f(s) \notin W_{e,s} \cup W_{k,s}$. Thus, there is a finite set F such that $(\forall x \in A \setminus F)\, \exists s\, (x \in A^{s+1} \setminus (W_{e,s} \cup W_{k,s}))$. On the other hand, it is obvious that $(\forall x \in \mathbb{N} \setminus A)\, \exists s\, (x \in (W_{e,s} \cup W_{k,s}) \setminus A^{s+1})$. Hence A is recursive, a contradiction. This completes the proof. $\square$

3.7 Maximal Sets

Definition 3.7.1 (maximal sets). An r.e. set $A \subseteq \mathbb{N}$ is said to be *maximal* if

1. $\overline{A} = \mathbb{N} \setminus A$ is infinite, yet
2. $\forall$ r.e. $B \supseteq A$, either $B \setminus A$ or $\mathbb{N} \setminus B$ is finite.

Remark 3.7.2. Maximality is a lattice-theoretic property. It is equivalent to A being a maximal element of the lattice $\mathcal{E}^* = \mathcal{E}/\{\text{finite sets}\}$.

Remark 3.7.3. If A is maximal, then A is simple. This is easily proved.

The following theorem is due to Friedberg. We have already stated this as item 2 in Remark 3.1.22.

Theorem 3.7.4. There exists a maximal set.

Remark 3.7.5 (movable markers). In our construction of a maximal set A, we shall have $A = \bigcup_s A^s$ where

$$A^0 \subseteq A^1 \subseteq \cdots \subseteq A^s \subseteq \cdots$$

and each A^s is finite. Here $A^0, A^1, \ldots, A^s, \ldots$ will be a recursive sequence of finite sets. We shall write $a_n^s =$ the nth element of $\overline{A^s}$, i.e.,

$$\overline{A^s} = \{a_0^s < a_1^s < \cdots < a_n^s < \cdots\}\,.$$

In addition, we shall have $a_n = \lim_s a_n^s =$ the nth element of $\overline{A}$, so that

$$\overline{A} = \{a_0 < a_1 < \cdots < a_n < \cdots\}\,.$$

A construction with these general features is known as a *movable marker* construction. We think of a_n^s as the position of the nth marker at stage s. Since $A^s \subseteq A^{s+1}$, we have $a_n^{s+1} = a_j^s$ for some $j \geq n$. This means that, whenever a marker is moved, it always lands on a position that was previously occupied by another marker. The fact that $a_n = \lim_s a_n^s < \infty$ means that the nth marker is moved only finitely many times, and its final position is a_n.

Definition 3.7.6 (e-states). We define $\sigma(e, x, s) =$ the *e-state* of x at stage s. This is defined by $\sigma(e, x, s) = \langle k_0, k_1, \ldots, k_e\rangle$, where $k_i = 1$ if $x \in W_{i,s}$, and $k_i = 0$ otherwise. In addition, we define $\sigma(e, x) = \lim_s \sigma(e, x, s) =$ the final e-state of x.

Remark 3.7.7. The e-states are a bookkeeping device. Our strategy will be to maximize the final e-state of a_e with respect to the lexicographic ordering of e-states. This ordering is defined by putting

$$\langle k_0, k_1, \ldots, k_e\rangle <_{\text{lex}} \langle l_0, l_1, \ldots, l_e\rangle$$

if and only if there exists $i \leq e$ such that $k_0 = l_0, \ldots, k_{i-1} = l_{i-1}, k_i < l_i$.

We now prove Theorem 3.7.4.

Proof. Our construction is as follows.

Stage 0. Put $A^0 = \emptyset$, and $a_e^0 = e$ for all e.

Stage s+1. We have $\overline{A^s} = \{a_0^s < a_1^s < \cdots < a_e^s < \cdots\}$. Choose the least e such that $\sigma(e, a_e^s, s) <_{\text{lex}} \sigma(e, a_j^s, s)$ for some $j > e$. For this e, choose the least such j. Put $A^{s+1} = A^s \cup \{a_e^s, a_{e+1}^s, \ldots, a_{j-1}^s\}$. If there is no such e, do nothing, i.e., $A^{s+1} = A^s$.

Remark 3.7.8. By construction, $a_i^s \leq a_i^{s+1}$ for all i. If no e is chosen at stage $s+1$, then $a_i^s = a_i^{s+1}$ for all i. If e is chosen at stage $s+1$, then $a_i^s = a_i^{s+1}$ for all $i < e$, and $a_j^s < a_j^{s+1}$ for all $j \geq e$, and $\sigma(e, a_e^s, s) <_{\text{lex}} \sigma(e, a_e^{s+1}, s+1)$.

Remark 3.7.9. Clearly the construction is primitive recursive, and $A = \bigcup_s A^s$ is an r.e. set. The idea behind the construction is that for each index i we have a requirement

$$R_i\colon \text{ for all } e \geq i,\ a_e \in W_i \text{ “if possible,”}$$

with priority i. Lower numbered requirements receive higher priority, embodied in the lexicographic ordering of e-states.

Lemma 3.7.10. For all e, $a_e = \lim_s a_e^s$ exists and is finite, i.e., the eth marker moves only finitely many times.

Proof. By induction on e, let s_1 be such that $\forall s > s_1\, \forall i < e\, (a_i^s = a_i^{s+1})$. It follows that $\forall s \geq s_1\, \forall i < e\, (i \text{ was not chosen at stage } s{+}1)$. Hence, for every $s \geq s_1$, if $a_e^s < a_e^{s+1}$ then e was chosen at stage $s{+}1$, and $\sigma(e, a_e^s, s) <_{\text{lex}} \sigma(e, a_e^{s+1}, s)$. Since there are only 2^{e+1} e-states, it follows that $\{s \geq s_1 \mid a_e^s < a_e^{s+1}\}$ is of cardinality $< 2^{e+1}$. This proves our lemma. □

Because of the previous lemma, we now know that

$$\overline{A} = \{a_0 < a_1 < \cdots < a_e < \cdots\}$$

is infinite.

Lemma 3.7.11. $\neg\, \exists e\, \exists j\, (e < j \wedge \sigma(e, a_e) <_{\text{lex}} \sigma(e, a_j))$.

Proof. Suppose not, i.e., $e < j$ and $\sigma(e, a_e) <_{\text{lex}} \sigma(e, a_j)$. By Lemma 3.7.10, for all sufficiently large s and all $i \leq j$ we have $a_i^s = a_i$ and $\sigma(e, a_i, s) = \sigma(e, a_i)$. In particular $a_e^s = a_e$ and $\sigma(e, a_e^s, s) = \sigma(e, a_e) <_{\text{lex}} \sigma(e, a_j) = \sigma(e, a_j^s, s)$. Hence some $i \leq e$ must have been chosen at stage $s+1$, hence $a_e^s < a_e^{s+1}$, and this is a contradiction. □

Lemma 3.7.12. $\forall e\, (W_e \cap \overline{A} \text{ or } \overline{W}_e \cap \overline{A} \text{ is finite })$.

Proof. By induction on e, we have $\forall k < e\, (W_k \cap \overline{A} \text{ or } \overline{W}_k \cap \overline{A} \text{ is finite})$, i.e., $\sigma(e{-}1, a_i) = \sigma(e{-}1, a_j)$ for all sufficiently large i and j. If $W_e \cap \overline{A}$ and $\overline{W}_e \cap \overline{A}$ are both infinite, there exist i and j such that $e \leq i < j$ and $\sigma(e{-}1, a_i) = \sigma(e{-}1, a_j)$ and $a_i \notin W_e$ and $a_j \in W_e$. It follows that $\sigma(i, a_i) <_{\text{lex}} \sigma(i, a_j)$, contradicting Lemma 3.7.11. □

This completes the proof of Theorem 3.7.4. □

3.8 The Owings Splitting Theorem and its Consequences

Remark 3.8.1. Recall from Definition 3.1.20 that a *Boolean algebra* is a complemented distributive lattice with 0 and 1. We state without proof the following well known algebraic facts.

1. Every distributive lattice with 0 and 1 is a sublattice of a Boolean algebra with the same 0 and 1.
2. Every Boolean algebra is isomorphic to a subalgebra of the Boolean algebra $(P(X), \cup, \cap, \emptyset, X)$, where X is a set.
3. Every finite Boolean algebra is isomorphic to $P(\{1, \ldots, n\})$ for some n, hence is of cardinality 2^n.

On the other hand, there are plenty of distributive lattices with 0 and 1 which are not complemented, i.e., not Boolean algebras. Examples are: (1) any linear ordering with a bottom element 0 and a top element 1 and at least one additional element; (2) the lattice of functions $[0,1]^X$ for any nonempty set X. Here $[0,1]$ is the unit interval in the real line.

Recall that $\mathcal{E}$, the lattice of r.e. sets, is a distributive lattice with 0 and 1. It is not a Boolean algebra, because there exist nonrecursive r.e. sets. The Friedberg Splitting Theorem says:

> If $A \in \mathcal{E}$ is nonrecursive, then there exist nonrecursive $B_1, B_2 \in \mathcal{E}$ such that $B_1 \cup B_2 = A$ and $B_1 \cap B_2 = \emptyset$.

We now consider some closely related lattices.

Definition 3.8.2. We define

$$\mathcal{E}^* = \mathcal{E}/\{\text{finite sets}\}\,.$$

Like $\mathcal{E}$, $\mathcal{E}^*$ is a distributive lattice with 0 and 1 and is not a Boolean algebra. Unlike $\mathcal{E}$, $\mathcal{E}^*$ is *atomless*, i.e.,

$$\forall a \in \mathcal{E}^* \,(a > 0 \Rightarrow \exists b \in \mathcal{E}^* \,(a > b > 0))\,.$$

Definition 3.8.3. Let C be a fixed r.e. set. We define

$$\mathcal{E}(C) = \{A \in \mathcal{E} \mid A \supseteq C\}\,,$$

the lattice of r.e. supersets of C. Again, $\mathcal{E}(C)$ is a distributive lattice, with $0 = C$ and $1 = \mathbb{N}$. Note that there is a lattice homomorphism $\mathcal{E} \to \mathcal{E}(C)$ given by $A \mapsto A \cup C$. This homomorphism is a retraction. We define

$$\mathcal{E}^*(C) = \mathcal{E}(C)/\{\text{finite sets}\}\,.$$

This is again a distributive lattice with 0 and 1.

Examples 3.8.4.

1. If C is cofinite, then $|\mathcal{E}^*(C)| = 1$. This is a degenerate case, which we shall ignore.

2. If C is a maximal set, then $\mathcal{E}(C) = \{C \cup F, C \cup (\mathbb{N} - F) \mid F \text{ is finite}\}$. Hence $\mathcal{E}^*(C) =$ the 2-element Boolean algebra $\{0, 1\}$. In fact, this property is equivalent to maximality of C.

3. Let $C = C_1 \cap C_2$ where C_1, C_2 are maximal sets such that $C_1 \cup C_2 = \mathbb{N}$, e.g., $\overline{C}_1 \subseteq \{\text{evens}\}$ and $\overline{C}_2 \subseteq \{\text{odds}\}$. Then $\mathcal{E}^*(C) =$ the 4-element Boolean algebra $\{0, c_1, c_2, 1\}$. Here c_1 and c_2 are the equivalence classes of C_1 and C_2 modulo finite sets.

4. Similarly let $C = C_1 \cap \cdots \cap C_n$ where C_i is maximal and $\overline{C}_i$, $i = 1, \cdots n$ are pairwise disjoint. Then $\mathcal{E}^*(C) =$ the 2^n-element Boolean algebra.

5. In the same vein, there exist r.e. sets C such that $\mathcal{E}^*(C)$ is an infinite Boolean algebra. This Boolean algebra can be atomless or non-atomless, depending on C.

6. If C is a coinfinite r.e. set which is not simple (e.g., $C = \emptyset$, or C creative), then the lattices $\mathcal{E}(C)$ and $\mathcal{E}^*(C)$ are not Boolean algebras.

Exercise 3.8.5. Show that $\mathcal{E}(C)$ is a Boolean algebra if and only if $\mathcal{E}^*(C)$ is a Boolean algebra.

Remark 3.8.6. As we have just seen, $\mathcal{E}(C)$ and $\mathcal{E}^*(C)$ can look quite different from $\mathcal{E}$ and $\mathcal{E}^*$. Nevertheless, the Friedberg Splitting Theorem generalizes to $\mathcal{E}(C)$. This is the content of the Owings Splitting Theorem, which we now state.

Theorem 3.8.7 (Owings). If $A \in \mathcal{E}(C)$ is noncomplemented, then there exist noncomplemented $B_1, B_2 \in \mathcal{E}(C)$ such that $A = B_1 \cup B_2$ and $B_1 \cap B_2 = C$.

Remark 3.8.8. Setting $C = \emptyset$, we recover the Friedberg Splitting Theorem.

Before proving the Owings Splitting Theorem, we examine its consequences concerning lattice-theoretic properties of $\mathcal{E}(C)$ and $\mathcal{E}^*(C)$.

Theorem 3.8.9. If $\mathcal{E}^*(C)$ is finite, then $\mathcal{E}^*(C)$ is a Boolean algebra. Hence $|\mathcal{E}^*(C)| = 2^n$ and C is the intersection of n maximal sets, for some n.

Proof. Suppose $\mathcal{E}^*(C)$ were not a Boolean algebra. It follows that $\mathcal{E}(C)$ is not a Boolean algebra. Let $A \in \mathcal{E}(C)$ be noncomplemented. By the Owings Splitting Theorem, let $A = B_1 \cup B_2$ where $B_1 \cap B_2 = C$ and B_1 and B_2 are noncomplemented. By the Owings Splitting Theorem again, let $B_2 = B_3 \cup B_4$, where $B_3 \cap B_4 = C$ and B_3 and B_4 are noncomplemented. Continuing in this fashion, we generate $B_1, B_3, B_5, \ldots$. This is an infinite sequence of noncomplemented elements of $\mathcal{E}(C)$, the intersection of any two of which is C. It follows that $\mathcal{E}^*(C)$ is infinite. This proves our theorem. $\square$

Corollary 3.8.10. $|\mathcal{E}^*(C)| \neq 3$.

Next we shall use the Owings Splitting Theorem to characterize the r.e. sets C for which $\mathcal{E}(C)$ is a Boolean algebra. Recall that this property implies that C is simple. We shall now define a subclass of the simple sets, called the hyperhypersimple sets. This class was first defined by Post.

Definition 3.8.11 (hyperhypersimple sets).

1. An *array* is a uniformly r.e. sequence of r.e. sets. Thus an array consists of a sequence of Σ^0_1 sets $B_i \subseteq \mathbb{N}$, $i = 0, 1, 2, \ldots$, such that in addition the 2-place relation $\{\langle x, i\rangle \mid x \in B_i\}$ is Σ^0_1. By the Parametrization Theorem, this is equivalent to saying that $B_i = W_{f(i)}$ for all i, where $f(i)$ is some primitive recursive function.

2. An r.e. set C is said to be *hyperhypersimple*, abbreviated *hhsimple*, if $\overline{C}$ is infinite and there does not exist an array of pairwise disjoint r.e. sets $W_{f(i)}$, $i = 0, 1, 2, \ldots$, such that $W_{f(i)} \cap \overline{C} \neq \emptyset$ for all i.

Theorem 3.8.12 (Lachlan). $\mathcal{E}(C)$ is a Boolean algebra if and only if C is hyperhypersimple.

Proof. In this proof we shall apply a uniform version of the Owings Splitting Theorem. The uniform version reads as follows.

> Given an r.e. index of a noncomplemented $A \in \mathcal{E}(C)$, we can recursively find r.e. indices of noncomplemented $B_1, B_2 \in \mathcal{E}(C)$ such that $B_1 \cup B_2 = A$ and $B_1 \cap B_2 = C$.

Assume now that $\mathcal{E}(C)$ is not a Boolean algebra. Let $A \in \mathcal{E}(C)$ be noncomplemented. Repeatedly apply the Owings Splitting Theorem as in the proof of Theorem 3.8.9 to generate an infinite sequence of noncomplemented sets

$$B_1, B_3, \ldots, B_{2i+1}, \ldots \in \mathcal{E}(C)\,,$$

the intersection of any two of which is C. By the uniformity, we may assume that $B_1, B_3, \ldots, B_{2i+1}$ are uniformly r.e., i.e., they form an array. This is almost what we want, except that these sets are not pairwise disjoint (unless $C = \emptyset$). To make them pairwise disjoint, let $\psi(x)$ be a partial recursive function which uniformizes the Σ^0_1 relation $S(x, i) \equiv x \in B_{2i+1}$. By the Parametrization Theorem, let $f(i)$ be a primitive recursive function such that $W_{f(i)} = \{x \mid \psi(x) \simeq i\}$ for all i. Then clearly $W_{f(i)}$, $i = 0, 1, 2, \ldots$, are pairwise disjoint, and $W_{f(i)} \cap \overline{C} = B_{2i+1} \cap \overline{C} \neq \emptyset$ for all i. Thus C is not hhsimple.

Conversely, assume that C is not hhsimple. Let $W_{f(i)}$, $i = 0, 1, 2, \ldots$ be an array of pairwise disjoint r.e. sets such that $W_{f(i)} \cap \overline{C} \neq \emptyset$ for all i. Set

$$A \,=\, C \cup \bigcup_{i=0}^{\infty} (W_i \cap W_{f(i)})\,.$$

Obviously $A \in \mathcal{E}(C)$. We claim that A is noncomplemented in $\mathcal{E}(C)$. Suppose $B \in \mathcal{E}(C)$ is the complement of A within $\mathcal{E}(C)$, i.e., $A \cap B = C$ and $A \cup B = \mathbb{N}$. Let e be such that $B = W_e$. Let $x \in W_{f(e)} \cap \overline{C}$. Since the $W_{f(i)}$, $i = 0, 1, 2, \ldots$, are pairwise disjoint, we have $x \in B \Leftrightarrow x \in W_e \Leftrightarrow x \in W_e \cap W_{f(e)} \Leftrightarrow x \in A$. This contradiction completes the proof. □

Exercises 3.8.13.

1. Show that C is hhsimple if and only if, for all r.e. sets A, $\overline{A} \cup C$ is r.e.
2. Show that if C_1 and C_2 are hhsimple then $C_1 \cap C_2$ is hhsimple.

3.9 Proof of the Owings Splitting Theorem

In this section we prove Theorem 3.8.7, the Owings Splitting Theorem.

Remark 3.9.1. The general framework for the proof will be the same as for the Friedberg Splitting Theorem. Let A and C be r.e. sets. Let f be a one-to-one recursive function which enumerates A. We write $A^s = \{f(0), \ldots, f(s-1)\}$. Similarly, let C^s be an enumeration of C. Just as in Section 3.6, we shall recursively decide at stage $s+1$ whether to put $f(s)$ into B_1 or into B_2. Thus we shall automatically have B_1, B_2 r.e. and $B_1 \cup B_2 = A$ and $B_1 \cap B_2 = \emptyset$.

In order to prove the Owings Splitting Theorem, we shall want to make sure that $B_1 \cup C$ and $B_2 \cup C$ are noncomplemented in $\mathcal{E}(C)$. Note that, for any r.e. sets B and C, $B \cup C$ is complemented in $\mathcal{E}(C)$ if and only if $\overline{B} \cup C$ is r.e. In other words, there exists e such that $W_e = \overline{B} \cup C$. In particular, $W_e \cap B \setminus C = \emptyset$. Therefore, our strategy in the construction will be to make $W_e \cap B_i \setminus C \neq \emptyset$ "if possible," for all e and for $i = 1, 2$. As in Section 3.6, these requirements will have a priority ordering given by $(e', i') < (e, i)$ if and only if $2e' + i' < 2e + i$.

As part of our construction, in order to mitigate the effect of C, we shall define an auxiliary recursive function $h(e, i, s)$ for all e and for $i = 1, 2$. This function will somehow control the process.

The details of our construction are as follows.

Stage 0. Let $B_1^0 = B_2^0 = \emptyset$. Let $h(e, i, 0) = 0$ for all e and for $i = 1, 2$.

Stage $s+1$. If $\exists x < h(e, i, s)$ such that $x \in W_{e,s} \cap B_i^s \setminus C^s$, let $h(e, i, s+1) = h(e, i, s)$. Otherwise, let $h(e, i, s+1) = h(e, i, s) + 1$. Set $y = f(s)$. Choose the least (e, i) such that $y \in W_{e,s}$ and $y < h(e, i, s)$. Put y into B_i, i.e., $B_i^{s+1} = B_i^s \cup \{y\}$. If no such (e, i) exists, put y into B_1.

By construction, B_1, B_2 are r.e. and $A = B_1 \cup B_2$ and $B_1 \cap B_2 = \emptyset$. Note also that $h(e, i, s) \leq h(e, i, s+1)$ for all s.

Definition 3.9.2. Say that (e, i) is *good* if $\lim_s h(e, i, s) < \infty$. Otherwise, say that (e, i) is *bad*.

Lemma 3.9.3. (e, i) is good $\iff$ $W_e \cap B_i \setminus C \neq \emptyset$.

Proof. $\Longrightarrow$: Assume (e,i) is good but $W_e \cap B_i \setminus C = \emptyset$, i.e., $W_e \cap B_i \subseteq C$. Since (e,i) is good, let s be so large that $h(e,i,t) = h(e,i,s)\ \forall t > s$. Let $t > s$ be so large that $\forall x < h(e,i,s)\,(x \in W_e \cap B_i \Rightarrow x \in C^t)$. Hence $\neg\,\exists x < h(e,i,t)\,(x \in W_{e,t} \cap B_i^t \setminus C^t)$. Therefore $h(e,i,t+1) = h(e,i,t)+1 > h(e,i,s)$, a contradiction.

$\Longleftarrow$: Assume (e,i) is bad and $W_e \cap B_i \setminus C \neq \emptyset$. Fix $x \in W_e \cap B_i \setminus C$. Since (e,i) is bad, $\lim_s h(e,i,s) = \infty$, hence for all sufficiently large s we have $x < h(e,i,s)$ and $x \in W_{e,s} \cap B_i^s \setminus C^s$, hence $h(e,i,s+1) = h(e,i,s)$, a contradiction. $\square$

Lemma 3.9.4. If (e,i) is good, then $\{s \mid (e,i) \text{ is chosen at stage } s+1\}$ is finite.

Proof. If (e,i) is chosen at stage $s+1$, we have $y = f(s) < h(e,i,s)$. Since $\lim_s h(e,i,s) < \infty$ and f is one-to-one, this can happen only finitely many times. $\square$

Lemma 3.9.5. Suppose (e,i) is bad. If (e,i) is chosen at stage $s+1$, then $y = f(s) \in C$.

Proof. Since (e,i) is chosen at stage $s+1$, we have $y = f(s) \in W_{e,s}$ and $y < h(e,i,s)$ and $y \in B_i^{s+1}$. So $y \in W_e \cap B_i$. Hence, by Lemma 3.9.3, $y \in C$. $\square$

Lemma 3.9.6. Assume that $A \cup C$ is noncomplemented in $\mathcal{E}(C)$. Then for $i = 1,2$, $B_i \cup C$ is noncomplemented in $\mathcal{E}(C)$.

Proof. Suppose that $B_i \cup C$ is complemented in $\mathcal{E}(C)$, say $W_e = \overline{B_i} \cup C$. In particular, $W_e \cap B_i \setminus C = \emptyset$, hence (e,i) is bad. By Lemma 3.9.4, let s_1 be so large that, for all $s \geq s_1$ and all good $(e',i') < (e,i)$, (e',i') is not chosen at stage $s+1$. Put

$$\widetilde{W}_e = \{x \mid \exists s \geq s_1\,(x \in W_{e,s} \wedge x \notin A^s \wedge x < h(e,i,s))\}\,.$$

Clearly $\widetilde{W}_e$ is r.e.

We claim that $\widetilde{W}_e \cup C$ is the complement of $A \cup C$ in $\mathcal{E}(C)$.

Suppose first that $x \notin A \cup C$. Since $B_i \subseteq A$, we have $\overline{A} \subseteq \overline{B_i}$, hence $x \in W_e$. Since (e,i) is bad, for all sufficiently large $s \geq s_1$ we have $x < h(e,i,s)$, and $x \in W_{e,s}$. Since $x \notin A$, $x \notin A^s$, hence $x \in \widetilde{W}_e$.

Suppose next that $y \in \widetilde{W}_e \cap A$. Since $y \in \widetilde{W}_e$, let $s \geq s_1$ be such that $y \in W_{e,s}$ and $x \notin A^s$ and $y < h(e,i,s)$. Since $y \in A$ and $y \notin A^s = \{f(0),\ldots,f(s)\}$, let $t > s$ be such that $y = f(t)$. Clearly $y \in W_{e,t}$ and $y < h(e,i,t)$. Hence, by construction, at stage $t+1$ some bad $(e',i') \leq (e,i)$ was chosen. Hence, by Lemma 3.9.5, $y \in C$.

We have now proved our claim. Thus $A \cup C$ is complemented in $\mathcal{E}(C)$. $\square$

The proof of the Owings Splitting Theorem 3.8.7 is now complete.

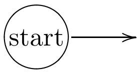
Remark 3.9.7. Our construction above is uniform. Therefore, given r.e. indices for A and C, we can use the Parametrization Theorem to primitive recursively find r.e. indices for B_1 and B_2. This extra uniformity was used in the proof of Lachlan's Theorem 3.8.12.

3.10 Oracle Computations

In this section we discuss *oracle computations*, i.e., computations where the computing device has the ability to consult an oracle. Intuitively, an *oracle* is a "black box" which, given a natural number as input, immediately produces a natural number as output. Our formal definitions are as follows.

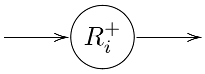
Definition 3.10.1 (oracles). An *oracle* is a total function $f : \mathbb{N} \to \mathbb{N}$. We write $\mathbb{N}^\mathbb{N}$ for the set of all oracles. Note that $\mathbb{N}^\mathbb{N}$ is also known as the *Baire space*.

Definition 3.10.2. Recall from Math 558 [14] that a register machine program consists of four kinds of instructions: a start instruction



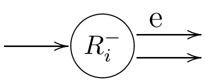
,

increment instructions



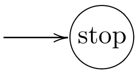
,

decrement instructions



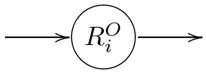
,

and stop instructions



.

We now introduce a fifth kind of instruction



called an *oracle instruction*. In the presence of an oracle f, the effect of R_i^O is to replace the content n of R_i by $f(n)$. In other words, if R_i contains n before executing R_i^O, then afterward R_i contains $f(n)$. We define an *oracle program* to be a register machine program as in Math 558 [14], except that oracle instructions are allowed.

Definition 3.10.3 (oracle computations). Let $\mathcal{P}$ be an oracle program, let f be an oracle, and let $x_1, \ldots, x_k \in \mathbb{N}$, where $k \geq 0$. We denote by $\mathcal{P}^f(x_1, \ldots, x_k)$ the unique run of $\mathcal{P}$ using oracle f starting with $x_1, \ldots, x_k$ in $R_1, \ldots, R_k$ and all other registers empty. As before, the *output* of $\mathcal{P}^f(x_1, \ldots, x_k)$ is the content of R_{k+1} if and when $\mathcal{P}^f(x_1, \ldots, x_k)$ halts. If $e = \#(\mathcal{P}) =$ the Gödel number of $\mathcal{P}$, we write

$$\varphi_e^{(k),f}(x_1, \ldots, x_k) \simeq \text{ the output of } \mathcal{P}^f(x_1, \ldots, x_k)\,.$$

We also introduce notations such as

$$W_e^f = \text{domain}(\varphi_e^{(1),f})\,.$$

Definition 3.10.4 (f-recursive functions, etc.). Let $f \in \mathbb{N}^\mathbb{N}$ be a fixed oracle. A partial function $\psi : \mathbb{N}^k \xrightarrow{P} \mathbb{N}$ is said to be *partial f-recursive* or *partial recursive in f* or *partial recursive relative to f*, if there exists $e \in \mathbb{N}$ such that $\psi(x_1, \ldots, x_k) \simeq \varphi_e^{(k),f}(x_1, \ldots, x_k)$ for all $x_1, \ldots, x_k \in \mathbb{N}$. Similarly, a set $A \subseteq \mathbb{N}$ is said to be *f-recursively enumerable* if and only if $A = W_e^f$ for some $e \in \mathbb{N}$, etc.

Remark 3.10.5. If the oracle f happens to be recursive, then clearly ψ is partial f-recursive $\iff$ ψ is partial recursive, A is f-r.e. $\iff$ A is r.e., etc. Thus we see that oracle computations are a generalization of ordinary, non-oracle computations.

Remark 3.10.6 (relativization to an oracle). Let $f \in \mathbb{N}^\mathbb{N}$ be a fixed oracle. A routine generalization of the Enumeration Theorem from Math 558 [14] asserts that for each $f \in \mathbb{N}^\mathbb{N}$ and each $k \geq 0$ the partial function

$$(e, x_1, \ldots, x_k) \mapsto \varphi_e^{(k),f}(x_1, \ldots, x_k)$$

is partial f-recursive. Similarly, all of our previous results about partial recursive functions, r.e. sets, the arithmetical hierarchy, etc., generalize routinely to partial f-recursive functions, f-r.e. sets, the f-arithmetical hierarchy, etc., where $f \in \mathbb{N}^\mathbb{N}$ is an arbitrary oracle. This process of routine generalization, replacing $\varphi_e^{(k)}$ by $\varphi_e^{(k),f}$, etc., is known as *relativization* to f.

Definition 3.10.7 (relativized arithmetical hierarchy). Let $f \in \mathbb{N}^\mathbb{N}$ be a fixed oracle. For $k, n \geq 1$, a relation $S \subseteq \mathbb{N}^k$ is said to be $\Sigma_n^{0,f}$ if there exists an f-recursive relation $R \subseteq \mathbb{N}^{k+n}$ such that

$$S(x_1, \ldots, x_k) \equiv \exists y_1 \, \forall y_2 \cdots y_n \, R(x_1, \ldots, x_k, y_1, y_2, \ldots, y_n)$$

where there are n alternating quantifiers, and the last quantifier is existential if n is odd, universal if n is even. A relation $P \subseteq \mathbb{N}^k$ is said to be $\Pi_n^{0,f}$ if $\neg P$ is $\Sigma_n^{0,f}$. A relation $D \subseteq \mathbb{N}^k$ is said to be $\Delta_n^{0,f}$ if it is both $\Sigma_n^{0,f}$ and $\Pi_n^{0,f}$. The classes $\Sigma_n^{0,f}$, $\Pi_n^{0,f}$, $\Delta_n^{0,f}$, $n \geq 1$ are known as the *f-arithmetical hierarchy.* All of the standard results about the arithmetical hierarchy (see Math 558 notes [14]) generalize routinely to the f-arithmetical hierarchy, for each $f \in \mathbb{N}^\mathbb{N}$.

Remark 3.10.8. Instead of viewing f as a fixed oracle, we may choose to view f as a variable ranging over the Baire space $\mathbb{N}^\mathbb{N}$. In this way, one develops a kind of recursion theory over $\mathbb{N}^\mathbb{N}$, including a version of the arithmetical hierarchy over $\mathbb{N}^\mathbb{N}$, etc. The starting point of this theory is the following definition.

Definition 3.10.9 (partial recursive functionals).

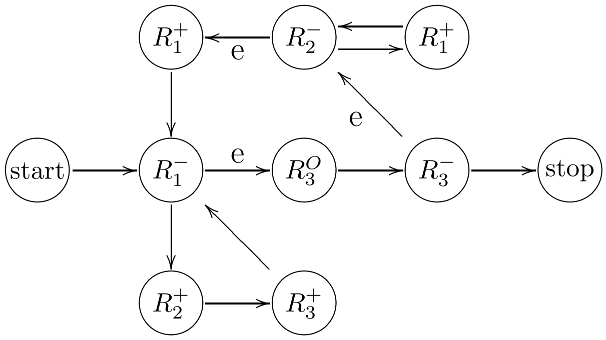
1. A *partial functional* is a partial function $\Psi : \mathbb{N}^\mathbb{N} \times \mathbb{N}^k \xrightarrow{P} \mathbb{N}$, where $k \geq 0$.

2. A partial functional $\Psi : \mathbb{N}^{\mathbb{N}} \times \mathbb{N}^k \xrightarrow{P} \mathbb{N}$ is said to be *partial recursive* if it is computable, i.e., if there exists $e \in \mathbb{N}$ such that

$$\Psi(f, x_1, \ldots, x_k) \simeq \varphi_e^{(k),f}(x_1, \ldots, x_k)$$

for all $f \in \mathbb{N}^{\mathbb{N}}$ and all $x_1, \ldots, x_k \in \mathbb{N}$.

Example 3.10.10. We exhibit an oracle program which computes the partial recursive functional $\Psi(f, x) \simeq$ the least $y \geq x$ such that $f(y) > 0$.



Letting e be the Gödel number of this program, we have $\varphi_e^{(1),f}(x) \simeq \Psi(f, x)$ for all $f \in \mathbb{N}^{\mathbb{N}}$ and all $x \in \mathbb{N}$.

The Enumeration, Parametrization, and Recursion Theorems from Math 558 [14] easily generalize to partial recursive functionals, as follows:

Theorem 3.10.11 (Enumeration Theorem). For each $k \geq 0$ we have a partial recursive functional

$$(f, e, x_1, \ldots, x_k) \mapsto \varphi_e^{(k),f}(x_1, \ldots, x_k).$$

Theorem 3.10.12 (Parametrization Theorem). Given a partial recursive functional $\Psi(f, x_0, x_1, \ldots, x_k)$, we can find a primitive recursive function $h(x_0)$ such that

$$\varphi_{h(x_0)}^{(k),f}(x_1, \ldots, x_k) \simeq \Psi(f, x_0, x_1, \ldots, x_k)$$

for all $f \in \mathbb{N}^{\mathbb{N}}$ and all $x_0, x_1, \ldots, x_k \in \mathbb{N}$.

Theorem 3.10.13 (Recursion Theorem). Given a partial recursive functional $\Psi(f, x_0, x_1, \ldots, x_k)$, we can find an index $e \in \mathbb{N}$ such that

$$\varphi_e^{(k),f}(x_1, \ldots, x_k) \simeq \Psi(f, e, x_1, \ldots, x_k)$$

for all $f \in \mathbb{N}^{\mathbb{N}}$ and all $x_1, \ldots, x_k \in \mathbb{N}$.

In this vein we obtain an alternative generalization of the arithmetical hierarchy from Math 558 [14], as follows.

Definition 3.10.14 (arithmetical hierarchy). For $k \geq 0$, a relation $R \subseteq \mathbb{N}^{\mathbb{N}} \times \mathbb{N}^k$ is said to be *recursive* if its characteristic function $\chi_R : \mathbb{N}^{\mathbb{N}} \times \mathbb{N}^k \to \mathbb{N}$ is recursive. For $k \geq 0$ and $n \geq 1$, a relation $S \subseteq \mathbb{N}^{\mathbb{N}} \times \mathbb{N}^k$ is said to be Σ^0_n if there exists a recursive relation $R \subseteq \mathbb{N}^{\mathbb{N}} \times \mathbb{N}^{k+n}$ such that

$$S(f, x_1, \ldots, x_k) \equiv \exists y_1 \, \forall y_2 \, \cdots \, y_n \, R(f, x_1, \ldots, x_k, y_1, y_2, \ldots, y_n)$$

where there are n alternating quantifiers, and the last quantifier is existential if n is odd, universal if n is even. A relation $P \subseteq \mathbb{N}^{\mathbb{N}} \times \mathbb{N}^k$ is said to be Π^0_n if $\neg P$ is Σ^0_n. A relation $D \subseteq \mathbb{N}^{\mathbb{N}} \times \mathbb{N}^k$ is said to be Δ^0_n if it is both Σ^0_n and Π^0_n.

3.11 Degrees of Unsolvability

We now introduce Turing degrees, a.k.a., degrees of unsolvability.

Definition 3.11.1 (Turing reducibility). Let f and g be total functions, i.e., $f, g \in \mathbb{N}^{\mathbb{N}}$. We say that f is *Turing reducible to* g, abbreviated $f \leq_T g$, if f is computable using g as an oracle, i.e., f is recursive relative to g, i.e., f is g-recursive, i.e.,

$$\exists e \, \forall x \, f(x) = \varphi_e^{(1),g}(x) \, .$$

Proposition 3.11.2. For $f, g, h \in \mathbb{N}^{\mathbb{N}}$, we have

1. $f \leq_T f$, and
2. if $f \leq_T g$ and $g \leq_T h$, then $f \leq_T h$.

Proof. Straightforward. Note that, if we write

$$\mathrm{REC}(f) = \{h \in \mathbb{N}^{\mathbb{N}} \mid h \text{ is } f\text{-recursive}\} \, ,$$

then $f \leq_T g$ if and only if $\mathrm{REC}(f) \subseteq \mathrm{REC}(g)$. □

Definition 3.11.3. For $f, g \in \mathbb{N}^{\mathbb{N}}$ we say that f is *Turing equivalent* to g, abbreviated $f \equiv_T g$, if $f \leq_T g$ and $g \leq_T f$, i.e., $\mathrm{REC}(f) = \mathrm{REC}(g)$.

Proposition 3.11.4. $\equiv_T$ is an equivalence relation on $\mathbb{N}^{\mathbb{N}}$.

Proof. Immediate from Proposition 3.11.2. □

Definition 3.11.5 (Turing degrees). We let $\mathcal{D}_T$ denote the set of equivalence classes of $\mathbb{N}^{\mathbb{N}}$ modulo Turing reducibility:

$$\mathcal{D}_T \; = \; \mathbb{N}^{\mathbb{N}} / {\equiv_T} \, .$$

Elements of $\mathcal{D}_T$ are known as *Turing degrees* or *degrees of unsolvability* or sometimes just *degrees.* For any $f \in \mathbb{N}^{\mathbb{N}}$, the *Turing degree of* f is

$$\deg_T(f) \; = \; \{g \in \mathbb{N}^{\mathbb{N}} \mid f \equiv_T g\} \, .$$

We partially order $\mathcal{D}_T$ by letting $\deg_T(f) \leq \deg_T(g)$ if and only if $f \leq_T g$. Clearly this relation does not depend on the representative chosen from each equivalence class.

The structure of the partial ordering $(\mathcal{D}_T, \leq)$ has received much scrutiny in hundreds of research papers. In this section we mention only the most basic properties of $\mathcal{D}_T$.

Proposition 3.11.6. $\mathcal{D}_T$ has a least element, $\mathbf{0}$, which is just the set REC of recursive functions.

Proof. Straightforward, since any recursive function is g-recursive for all g. $\square$

Proposition 3.11.7. Every pair of Turing degrees $\mathbf{a}, \mathbf{b} \in \mathcal{D}_T$ has a least upper bound $\mathbf{a} \vee \mathbf{b} \in \mathcal{D}_T$. Thus $(\mathcal{D}_T, \leq)$ is an upper semilattice.

Proof. Let $\mathbf{a} = \deg_T(f)$ and $\mathbf{b} = \deg_T(g)$ be given, where $f \in \mathbb{N}^\mathbb{N}$ and $g \in \mathbb{N}^\mathbb{N}$. We define a function $f \oplus g \in \mathbb{N}^\mathbb{N}$ by letting $(f \oplus g)(2n) = f(n)$ and $(f \oplus g)(2n+1) = g(n)$ for all $n \in \mathbb{N}$. We claim that

$$\deg_T(f \oplus g) \;=\; \deg_T(f) \vee \deg_T(g)\,,$$

i.e., $\deg_T(f \oplus g)$ is the least upper bound of $\deg_T(f)$ and $\deg_T(g)$. Clearly $f \leq_T f \oplus g$ and $g \leq_T f \oplus g$. For any $h \in \mathbb{N}^\mathbb{N}$, if $f \leq_T h$ and $g \leq_T h$, then it is straightforward to show that $f \oplus g \leq_T h$. $\square$

Definition 3.11.8 (Turing degrees of sets). Given a set $A \subseteq \mathbb{N}$, we define $\deg_T(A) = \deg_T(\chi_A)$. In other words, the Turing degree of a set $A \subseteq \mathbb{N}$ is defined to be the Turing degree of its characteristic function $\chi_A : \mathbb{N} \to \mathbb{N}$. The following proposition shows that there is no loss in considering only Turing degrees of sets, rather than functions.

Proposition 3.11.9. Every Turing degree contains (the characteristic function of) a set.

Proof. Let $\deg_T(f)$ be an arbitrary Turing degree, where $f \in \mathbb{N}^\mathbb{N}$. It is straightforward to prove that $f \equiv_T \chi_A$ where $A = G_f = \{2^n 3^m \mid f(n) = m\}$. The proof uses the fact that f is a total function. $\square$

Definition 3.11.10. For $f \in \mathbb{N}^\mathbb{N}$, we let $f' = H^f$ be the Halting Problem relative to f, i.e.,

$$f' \;=\; H^f \;=\; \{e \mid \varphi_e^{(1),f}(0)\downarrow\}\,.$$

We can show that H^f is a complete $\Sigma_1^{0,f}$ set, i.e., H^f is $\Sigma_1^{0,f}$ and every $\Sigma_1^{0,f}$ set $A \subseteq \mathbb{N}$ is $\leq_m H^f$. This is the relativization to f of the fact that the Halting Problem $H = \{e \mid \varphi_e^{(1)}(0)\downarrow\}$ is a many-one complete r.e. set.

Proposition 3.11.11. For all $f, g \in \mathbb{N}^\mathbb{N}$ we have

1. $f <_T H^f$.
2. If $f \leq_T g$ then $H^f \leq_T H^g$.
3. $f \leq_T g$ if and only if $H^f \leq_m H^g$.

4. $H^f \equiv_T \neg H^f$ but $H^f \not\equiv_m \neg H^f$.

Proof. Straightforward. $\square$

Definition 3.11.12 (the Turing jump). For any Turing degree $\mathbf{a} = \deg_T(f)$ we let $\mathbf{a}' = \deg_T(f') = \deg_T(H^f)$. The degree $\mathbf{a}'$ is called the *Turing jump* of $\mathbf{a}$. By Proposition 3.11.11, $f \equiv_T g$ implies $H^f \equiv_T H^g$. Thus $\mathbf{a}'$ is well defined for all $\mathbf{a} \in \mathcal{D}_T$. The *Turing jump operator* $J : \mathcal{D}_T \to \mathcal{D}_T$ is defined by $J(\mathbf{a}) = \mathbf{a}'$.

Lemma 3.11.13. For any Turing degree $\mathbf{a}$ we have the strict inequality $\mathbf{a} < \mathbf{a}'$. For all $\mathbf{a}, \mathbf{b} \in \mathcal{D}_T$, if $\mathbf{a} \leq \mathbf{b}$ then $\mathbf{a}' \leq \mathbf{b}'$.

Proof. This is immediate from Proposition 3.11.11. $\square$

Definition 3.11.14. For $\mathbf{a} \in \mathcal{D}_T$ we define the *iterated Turing jumps* of $\mathbf{a}$ by induction as follows. Let $\mathbf{a}^{(0)} = \mathbf{a}$, and let $\mathbf{a}^{(n+1)} = (\mathbf{a}^{(n)})'$ for each $n \in \mathbb{N}$. Note that if $\mathbf{a} = \deg_T(A)$ then $\mathbf{a}^{(n)} = \deg_T(A^{(n)})$, where $A^{(0)} = A$ and $A^{(n+1)} = A^{(n)\prime}$.

An interesting relationship between Turing degrees and the arithmetical hierarchy is given by the following theorem due to Post.

Theorem 3.11.15 (Post). For $n \geq 1$, a set $A \subseteq \mathbb{N}$ is Σ^0_n if and only if it is r.e. relative to $0^{(n-1)}$. More generally, for $A, B \subseteq \mathbb{N}$, A is $\Sigma^{0,B}_n$ if and only if A is r.e. relative to $B^{(n-1)}$.

Proof. We omit the proof. See my Spring 2004 lecture notes [15]. $\square$

Corollary 3.11.16. For $n \geq 1$, A is Δ^0_n if and only if $A \leq_T 0^{(n-1)}$. More generally, A is $\Delta^{0,B}_n$ if and only if $A \leq_T B^{(n-1)}$.

Proof. A is $\Delta^{0,B}_n$ $\iff$ $A, \neg A$ are $\Sigma^{0,B}_n$ $\iff$ $A, \neg A$ are r.e. relative to $B^{(n-1)}$ $\iff$ A is recursive relative to $B^{(n-1)}$, i.e., $A \leq_T B^{(n-1)}$. $\square$

We note the following special case.

Corollary 3.11.17. $A \leq_T 0'$ if and only if A is Δ^0_2.

3.12 The Sacks Splitting Theorem and its Consequences

A substructure of $(\mathcal{D}_T, \leq)$ which has received a huge amount of attention is the recursively enumerable Turing degrees.

Definition 3.12.1. Let

$$\mathcal{E}_T = \{\deg_T(A) \mid A \text{ is recursively enumerable}\}.$$

The elements of $\mathcal{E}_T$ are called *recursively enumerable Turing degrees*, or *r.e. Turing degrees*, or sometimes just *r.e. degrees*.

The essential structure of $\mathcal{E}_T$ is given by the following proposition.

Proposition 3.12.2.

1. $\mathbf{0}, \mathbf{0}' \in \mathcal{E}_T$.
2. $\mathbf{0}$ is the bottom element of $\mathcal{E}_T$.
3. $\mathbf{0}'$ is the top element of $\mathcal{E}_T$.
4. $\mathcal{E}_T$ is closed under l.u.b. This means that if $\mathbf{a}, \mathbf{b} \in \mathcal{E}_T$ then $\mathbf{a} \vee \mathbf{b} \in \mathcal{E}_T$.

Proof. Statements 1 and 2 are obvious. If A is r.e., then by Proposition 3.2.5 $A \leq_m H = 0'$, hence $A \leq_T 0'$ and this gives statement 3. Alternatively, statement 3 follows from Corollary 3.11.17. For statement 4, note that if $A, B \subseteq \mathbb{N}$ are r.e. then so is

$$A \oplus B \;=\; \{2n \mid n \in A\} \cup \{2n+1 \mid n \in B\}$$

and $\chi_{A \oplus B} = \chi_A \oplus \chi_B$. $\square$

Remarks 3.12.3.

1. There are many Turing degrees $\mathbf{a}$ such that $\mathbf{a} \leq \mathbf{0}'$ yet $\mathbf{a} \notin \mathcal{E}_T$. Examples are provided by the Kleene/Post construction (see for instance [15]).
2. There are many sets $B \subseteq N$ such that $\deg_T(B) \in \mathcal{E}_T$ yet B is not r.e. For example, let B be the complement of a nonrecursive r.e. set.

At this moment, we have not yet proved that there exist any r.e. Turing degrees other than $\mathbf{0}$ and $\mathbf{0}'$. We shall prove the following theorem, which gives this and much more information concerning the structure of the r.e. degrees.

Theorem 3.12.4 (Sacks Splitting Theorem). Let $A \subseteq \mathbb{N}$ be a nonrecursive r.e. set. Let $C \subseteq \mathbb{N}$ be nonrecursive. There exist r.e. sets B_1 and B_2 such that

1. $A = B_1 \cup B_2$,
2. $B_1 \cap B_2 = \emptyset$,
3. $B_1 \not\leq_T B_2$,
4. $B_2 \not\leq_T B_1$,
5. $0 <_T B_1 <_T A$,
6. $0 <_T B_2 <_T A$,
7. $B_1 \oplus B_2 \equiv_T A$,
8. $C \not\leq_T B_1$,
9. $C \not\leq_T B_2$.

Before proving the Sacks Splitting Theorem, we note some of its corollaries. Say that $\mathbf{b}_1, \mathbf{b}_2 \in \mathcal{D}_T$ are *incomparable*, and write $\mathbf{b}_1 \mid \mathbf{b}_2$, if $\mathbf{b}_1 \not\leq \mathbf{b}_2$ and $\mathbf{b}_2 \not\leq \mathbf{b}_1$.

Corollary 3.12.5. For any $\mathbf{a} \in \mathcal{E}_T$ with $\mathbf{a} > \mathbf{0}$, there exist $\mathbf{b}_1, \mathbf{b}_2 \in \mathcal{E}_T$ such that $\mathbf{a} > \mathbf{b}_1 > \mathbf{0}$, $\mathbf{a} > \mathbf{b}_2 > \mathbf{0}$, $\mathbf{a} = \mathbf{b}_1 \vee \mathbf{b}_2$, and $\mathbf{b}_1 \mid \mathbf{b}_2$.

Proof. This is an immediate translation of parts 1–7 of the Sacks Splitting Theorem 3.12.4. □

Corollary 3.12.6 (Friedberg, Muchnik). There are incomparable r.e. Turing degrees in $\mathcal{E}_T$.

Proof. Apply Corollary 3.12.5 with $\mathbf{a} = \mathbf{0}'$. □

Corollary 3.12.7. There is an infinite strictly descending sequence of r.e. Turing degrees. There is an infinite set of pairwise incomparable r.e. Turing degrees.

Proof. Start with $\mathbf{a}_0 = \mathbf{0}'$. By Corollary 3.12.5 find r.e. degrees $\mathbf{a}_1, \mathbf{a}_2 < \mathbf{a}_0$ such that $\mathbf{a}_1 \mid \mathbf{a}_2$. By Corollary 3.12.5 again, find r.e. degrees $\mathbf{a}_3, \mathbf{a}_4 < \mathbf{a}_2$ such that $\mathbf{a}_3 \mid \mathbf{a}_4$. Continuing in this fashion, we see that $\mathbf{a}_0 > \mathbf{a}_2 > \mathbf{a}_4 > \cdots$ is an infinite descending sequence of r.e. degrees, while $\mathbf{a}_1, \mathbf{a}_3, \mathbf{a}_5, \ldots$ is an infinite sequence of pairwise incomparable r.e. degrees. □

Corollary 3.12.8. For any $\mathbf{a} > \mathbf{0}$ in $\mathcal{E}_T$, there exists $\mathbf{b}$ in $\mathcal{E}_T$ such that $\mathbf{a} > \mathbf{b} > \mathbf{0}$.

Proof. Apply Corollary 3.12.5 to $\mathbf{a}$. □

Corollary 3.12.9 (Friedberg, Muchnik). There exists a recursively enumerable Turing degree $\mathbf{a}$ which is *intermediate*, i.e., $\mathbf{0} < \mathbf{a} < \mathbf{0}'$. Equivalently, $\mathbf{a} \neq \mathbf{0}, \mathbf{0}'$.

Remark 3.12.10 (natural examples). The previous corollary is the solution to *Post's Problem* (see Rogers [11]). Another interesting and important problem, which remains open, is to find an example of a recursively enumerable Turing degree other than $\mathbf{0}$ and $\mathbf{0}'$ which is *mathematically natural*. The term "mathematically natural" has not been rigorously defined, but we would recognize such an example if we saw one.

3.13 Proof of the Sacks Splitting Theorem

We now turn to the proof of the Sacks Splitting Theorem. We begin by noting that many of the conclusions will follow automatically, if we can only prove that $A \not\leq_T B_1$ and $A \not\leq_T B_2$.

Lemma 3.13.1. Let A, B_1, B_2 be r.e. sets such that $A = B_1 \cup B_2$ and $B_1 \cap B_2 = \emptyset$. Then $A \equiv_T B_1 \oplus B_2$. Moreover, if $A \not\leq_T B_1$ and $A \not\leq_T B_2$, then $B_1 \not\leq_T B_2$ and $B_2 \not\leq_T B_1$, hence $0 <_T B_1 <_T A$ and $0 <_T B_2 <_T A$.

Proof. Assume first that A, B_1, B_2 are r.e. with $A = B_1 \cup B_2$ and $B_1 \cap B_2 = \emptyset$. We have $n \in A \iff (n \in B_1 \vee n \in B_2)$, so clearly $A \leq_T B_1 \oplus B_2$. For the converse, note that

$$n \in B_1 \iff (n \in A \wedge n \notin B_2)\,.$$

Since $\neg B_2$ is Π^0_1, it follows that B_1 is $\Pi^{0,A}_1$. But B_1 is Σ^0_1, hence $\Sigma^{0,A}_1$, so we actually have that B_1 is $\Delta^{0,A}_1$, i.e., B_1 is A-recursive, i.e., $B_1 \leq_T A$. A similar argument shows that $B_2 \leq_T A$. We now see that $B_1 \oplus B_2 \leq_T A$.

Now assume in addition that $A \not\leq_T B_1$ and $A \not\leq_T B_2$. We therefore have $B_1 <_T A$ and $B_2 <_T A$. Since $B_1 \oplus B_2 \equiv A$, it follows that B_1 and B_2 are Turing incomparable, hence nonrecursive. $\square$

We shall obtain the Sacks Splitting Theorem as a consequence of the following theorem of Binns, which is not only more general but also easier to state.

Theorem 3.13.2 (Binns Splitting Theorem). Let $P \subseteq \mathbb{N}^{\mathbb{N}}$ be Π^0_1 with no recursive members, i.e., $P \cap \mathrm{REC} = \emptyset$. Let A be an r.e. set. Then we can find r.e. sets B_1, B_2 such that $A = B_1 \cup B_2$, $B_1 \cap B_2 = \emptyset$, and there is no $f \in P$ such that $f \leq_T B_1$ or $f \leq_T B_2$.

Proof. We postpone the proof to Section 3.15. $\square$

Remark 3.13.3. The next lemma shows that the Binns Splitting Theorem implies its own generalization, replacing the Π^0_1 set $P \subseteq \mathbb{N}^{\mathbb{N}}$ by a Σ^0_3 set $S \subseteq \mathbb{N}^{\mathbb{N}}$. However, it fails for Π^0_3 sets. For example, it fails badly for the Π^0_3 set

$$\mathbb{N}^{\mathbb{N}} \setminus \mathrm{REC} = \{f \mid f \text{ is not recursive}\},$$

as shown by the Friedberg Splitting Theorem.

Definition 3.13.4 (weak equivalence).

1. For $S \subseteq \mathbb{N}^{\mathbb{N}}$, let $\widehat{S}$ be the *Turing upward closure* of S, i.e.,

$$\widehat{S} \;=\; \{g \in \mathbb{N}^{\mathbb{N}} \mid (\exists f \in S)\,(f \leq_T g)\}\,.$$

2. For $S_1, S_2 \subseteq \mathbb{N}^{\mathbb{N}}$, we say that S_1 and S_2 are *weakly equivalent*, and write $S_1 \equiv_w S_2$, if and only if $\widehat{S_1} = \widehat{S_2}$.

Lemma 3.13.5. Given a Σ^0_3 set $S \subseteq \mathbb{N}^{\mathbb{N}}$, we can find a Π^0_1 set $P \subseteq \mathbb{N}^{\mathbb{N}}$ such that $P \equiv_w S$.

Proof. The proof uses the technique of Skolem functions. Since S is Σ^0_3, we have

$$S \;=\; \{f \in \mathbb{N}^{\mathbb{N}} \mid \exists k\, \forall m\, \exists n\, R(f,k,m,n)\}$$

where R is recursive. Put

$$P \;=\; \{\langle k\rangle^\frown(f \oplus g) \mid \forall m\, R(f,k,m,g(m))\}\,.$$

Here $\langle k\rangle^\frown(f \oplus g)$ is our notation for the unique $h \in \mathbb{N}^{\mathbb{N}}$ such that $h(0) = k$ and $h(2m+1) = f(m)$, and $h(2m+2) = g(m)$ for all m. Clearly P is Π^0_1. Moreover, if $\langle k\rangle^\frown(f \oplus g) \in P$, then $f \in S$. Conversely, if $f \in S$, let k be such that $\forall m\, \exists n\, R(f,k,m,n)$, and define g by putting $g(m) =$ least n such that $R(f,k,m,n)$. Then $g \leq_T f$, hence $\langle k\rangle^\frown(f \oplus g) \leq_T f$ and $\in P$. □

Definition 3.13.6 (Π^0_2 singletons). We say that $f \in \mathbb{N}^{\mathbb{N}}$ is a Π^0_2 *singleton* if the singleton set $\{f\}$ is Π^0_2. This is equivalent to $\{f\}$ being Σ^0_3.

Lemma 3.13.7. If $f \leq_T 0'$ then f is a Π^0_2 singleton.

Proof. Assume $f \leq_T 0'$. By Corollary 3.11.17 f is Δ^0_2, i.e., the predicate $D(n,m) \equiv f(n) = m$ is Δ^0_2. Hence

$$P = \{g \in \mathbb{N}^{\mathbb{N}} \mid \forall n\, D(n, g(n))\} = \{g \in \mathbb{N}^{\mathbb{N}} \mid g = f\} = \{f\}$$

is Π^0_2 as a subset of $\mathbb{N}^{\mathbb{N}}$, i.e., f is a Π^0_2 singleton. □

Proposition 3.13.8. The Binns Splitting Theorem implies the Sacks Splitting Theorem.

Proof. Assume the Binns Splitting Theorem 3.13.2. We deduce the Sacks Splitting Theorem 3.12.4. Let A be a nonrecursive r.e. set, and let C be a nonrecursive set. Because A is r.e., we have $A \leq_T 0'$. If $C \leq_T 0'$ put $S = \{\chi_A, \chi_C\}$, otherwise put $S = \{\chi_A\}$. By Lemma 3.13.7 S is Σ^0_3. By Lemma 3.13.5 let $P \equiv_w S$ be Π^0_1. Apply the Binns Splitting Theorem to A and P to obtain r.e. sets B_1, B_2 such that $B_1 \cup B_2 = A$ and $B_1 \cap B_2 = \emptyset$ and there is no $f \in P$ such that $f \leq_T B_1$ or $f \leq_T B_2$. Since $P \equiv_w S$ and $B_1, B_2 \leq_T 0'$, we have $A \not\leq_T B_1$, $A \not\leq_T B_2$, $C \not\leq_T B_1$, $C \not\leq_T B_2$. The remaining conclusions of the Sacks Splitting Theorem now follow, in view of Lemma 3.13.1. □

It remains to prove the Binns Splitting Theorem.

3.14 Finite Approximations

For most proofs involving degrees of unsolvability, it is necessary to consider finite approximations to oracle computations. Intuitively, if an oracle computation $\varphi_e^{(1),f}(x)$ halts, then this computation can only use a finite amount of information from the oracle f, because it only performs a finite number of steps before halting. We state this insight formally as Proposition 3.14.3 below.

Notation 3.14.1 (finite sequences). We let $\mathrm{Seq} = \mathbb{N}^{<\mathbb{N}}$ denote the set of finite sequences of natural numbers. The *length* of $\sigma \in \mathrm{Seq}$ is denoted $\mathrm{lh}(\sigma)$. For $f \in \mathbb{N}^{\mathbb{N}}$ and $n \in \mathbb{N}$ we write

$$f[n] = \langle f(0), f(1), \ldots, f(n-1)\rangle \in \mathrm{Seq}\,.$$

Thus $\mathrm{lh}(f[n]) = n$. We write $f \supset \sigma$ if f *extends* σ, i.e., if $f[n] = \sigma$ where $n = \mathrm{lh}(\sigma)$. For $\sigma, \tau \in \mathrm{Seq}$ we write $\sigma \subset \tau$ if σ is an *initial segment* of τ, i.e., $\mathrm{lh}(\sigma) < \mathrm{lh}(\tau)$ and $\sigma(i) = \tau(i)$ for all $i < \mathrm{lh}(\sigma)$.

Definition 3.14.2 (finite approximations). For $e, s, x, y \in \mathbb{N}$ and $\sigma \in \text{Seq}$, we write

$$\varphi_{e,s}^{(1),\sigma}(x) \simeq y$$

if and only if $e, x, y < s$ and for some (equivalently, all) $f \in \mathbb{N}^{\mathbb{N}}$ extending σ, the oracle computation $\varphi_e^{(1),f}(x)$ halts in fewer than s steps with output y, and during this computation, no oracle information from f is used except the part of f which is in σ.

Proposition 3.14.3. We have:

1. $\varphi_e^{(1),f}(x) \simeq y$ if and only if $\exists n\, \exists s\; \varphi_{e,s}^{(1),f[n]}(x) \simeq y$.
2. $\varphi_e^{(1),f}(x) \simeq y$ if and only if $\exists s\; \varphi_{e,s}^{(1),f[s]}(x) \simeq y$.
3. If $s \leq t$ and $\sigma \subseteq \tau$, then $\varphi_{e,s}^{(1),\sigma}(x) \simeq y$ implies $\varphi_{e,t}^{(1),\tau}(x) \simeq y$.
4. The 5-place relation $\varphi_{e,s}^{(1),\sigma}(x) \simeq y$ is primitive recursive.

Proof. Straightforward. ☐

Definition 3.14.4 (use functions).

1. $u(f, e, x) =$ the supremum of all n such that the oracle information $f(n)$ is used in the computation of $\varphi_e^{(1),f}(x)$.
2. $u(\sigma, e, x, s) =$ the supremum of all $n < \text{lh}(\sigma)$ such that the oracle information $\sigma(n)$ is used in the first s steps of the computation of $\varphi_{e,s}^{(1),\sigma}(x)$.

Proposition 3.14.5.

1. $u(f, e, x) = \lim_s u(f[s], e, x, s)$.
2. The 4-place function $u(\sigma, e, x, s)$ is primitive recursive.

Proof. Straightforward. ☐

Definition 3.14.6 (trees). A *tree* is a set $T \subseteq \text{Seq}$ such that $\sigma \subset \tau$, $\tau \in T$ implies $\sigma \in T$. If T is a tree, a *path* through T is any $f \in \mathbb{N}^{\mathbb{N}}$ such that $f[n] \in T$ for all n. The set of all paths through T is denoted $[T]$.

Proposition 3.14.7. Given a Π_1^0 set $P \subseteq \mathbb{N}^{\mathbb{N}}$, we can find a primitive recursive tree $T \subseteq \text{Seq}$ such that $P = [T]$.

Proof. Let $P \subseteq \mathbb{N}^{\mathbb{N}}$ be Π_1^0. Then $P = \{f \mid \forall n\, R(f, n)\}$ where R is recursive. Let e be an index of χ_R, i.e., $\varphi_e^{(1),f}(n) = \chi_R(f, n)$ for all $f \in \mathbb{N}^{\mathbb{N}}$ and $n \in \mathbb{N}$. Define

$$T = \left\{ \sigma \in \text{Seq} \;\middle|\; \forall n < \text{lh}(\sigma) \;\; \varphi_{e,\text{lh}(\sigma)}^{(1),\sigma}(n) \not\simeq 0 \right\}.$$

Then T is a primitive recursive tree, and P is the set of paths through T. ☐

Notation 3.14.8. For $A \subseteq \mathbb{N}$ we write

1. $\varphi_e^{(1),A}(x) \simeq \varphi_e^{(1),\chi_A}(x)$,
2. $u(A, e, x) \simeq u(\chi_A, e, x)$.
3. $A[n] = \chi_A[n]$,

3.15 Proof of the Binns Splitting Theorem

We now restate and prove the Binns Splitting Theorem 3.13.2.

Theorem 3.15.1. Let $P \subseteq \mathbb{N}^\mathbb{N}$ be Π_1^0 with $P \cap \mathrm{REC} = \emptyset$. For any r.e. set A we can find r.e. sets B_1, B_2 such that $A = B_1 \cup B_2$, $B_1 \cap B_2 = \emptyset$, and $\neg\, \exists f \in P\, (f \leq_T B_1 \vee f \leq_T B_2)$.

Proof. The general framework for the proof is as for the Friedberg Splitting Theorem. Let f be a one-to-one recursive function such that $A = \mathrm{range}(f)$. We write $A^s = \{f(0), \ldots, f(s-1)\}$. Our construction will be such that at stage $s+1$ we have already defined B_1^s and B_2^s with $A^s = B_1^s \cup B_2^s$ and $B_1^s \cap B_2^s = \emptyset$ and at this stage we decide whether to put $f(s)$ into B_1 or B_2.

Our requirements for the Binns Splitting Theorem are

$$R(e, i) : \ \varphi_e^{(1),B_i} \notin P\,.$$

As usual, we define a priority ordering of the requirements by putting

$$(e', i') < (e, i)$$

if and only if $2e'+i' < 2e+i$. Our strategy for satisfying $R(e, i)$ will be to preserve computations tending to put $\varphi_e^{B_i}$ into P. This may seem counterintuitive, since $R(e, i)$ requires that $\varphi_e^{B_i} \notin P$. However, by preserving finite approximations to $\varphi_e^{B_i}$, we will eventually force $\varphi_e^{B_i}$ to be either recursive or not total, hence $\notin P$.

By Proposition 3.14.7 let T be a primitive recursive tree such that $P = [T]$, the set of paths through T. Thus $\varphi_e^{B_i} \in P$ if and only if $\forall x\, \varphi_e^{(1),B_i}(x)\downarrow$ and $\forall y\, \langle \varphi_e^{B_i}(x) \mid x < y \rangle \in T$. Note also that if $\varphi_e^{(1),B_i}(x)\downarrow$ then $\varphi_e^{(1),B_i}(x) = \lim_s \varphi_{e,s}^{(1),B_i^s[s]}(x)$ and $u(B_i, e, x) = \lim_s u(B_i^s[s], e, x, s)$. We define the *length function* by

$$l(e, i, s) \,=\, \sup \left\{ y \mid \forall x < y \ \ \varphi_{e,s}^{(1),B_i^s[s]}(x)\downarrow \ \text{ and } \ \left\langle \varphi_{e,s}^{(1),B_i^s[s]}(x) \mid x < y \right\rangle \in T \right\}.$$

We define the *restraint function* by

$$r(e, i, s) \,=\, \sup \left\{ u(B_i^s[s], e, x, s) \mid x \leq l(e, i, s) \text{ and } \varphi_{e,s}^{(1),B_i^s[s]}(x)\downarrow \right\}.$$

Roughly speaking, the length function $l(e, i, s)$ measures the amount of agreement between $\varphi_{e,s}^{(1),B_i^s[s]}$ and the tree T, while the restraint function $r(e, i, s)$

tells us how much oracle information needs to be preserved, in order to keep this agreement, for the sake of requirement $R(e, i)$.

Our construction is as follows.

Stage 0: $B_1^0 = B_2^0 = \emptyset$.

Stage $s+1$: Let $x = f(s)$. Choose the least (e, i) such that $x \leq r(e, i, s)$. If $i = 1$ or (e, i) is undefined, enumerate x into B_2. If $i = 2$, enumerate x into B_1.

This completes the construction.

The *injury set* $I(e, i)$ is defined by

$$I(e, i) \;=\; \{s+1 \mid f(s) \in B_i^{s+1} \setminus B_i^s \text{ and } f(s) \leq r(e, i, s)\}\,.$$

This is the set of stages at which $R(e, i)$ is injured. By construction, if $s+1 \in I(e, i)$ then some $(e', i') < (e, i)$ was chosen at stage $s+1$. In other words, a requirement can be injured only for the sake of requirements of higher priority.

Lemma 3.15.2. The following hold for all e, i.

1. $I(e, i)$ is finite.
2. $\varphi_e^{(1),B_i} \notin P$.
3. $r(e, i) = \lim_s r(e, i, s)$ exists and is finite.

Proof. We prove 1, 2, and 3 by simultaneous induction on (e, i). Assume that 1, 2, and 3 hold for all $(e', i') < (e, i)$. Put $r = \max\{r(e', i') \mid (e', i') < (e, i)\}$. Let s_1 be such that $A^s[r+1] = A[r+1]$ and $r(e', i', s) = r(e', i')$ for all $(e', i') < (e, i)$ and $s \geq s_1$. Then by construction $I(e, i) \subseteq \{0, \ldots, s_1+1\}$, so this injury set is finite. This proves 1.

To prove 2, assume for a contradiction that $\varphi_e^{(1),B_i} \in P$. Then, given y, we can effectively find $s > s_1$ such that $\forall x \leq y\ \varphi_{e,s}^{(1),B_i^s[s]}(x)\downarrow$ and

$$\left\langle \varphi_{e,s}^{(1),B_i^s[s]}(x) \;\middle|\; x \leq y \right\rangle \in T\,.$$

For all such s we have $y \leq l(e, i, s)$, hence $u(B_i^s[s], e, y, s) \leq r(e, i, s)$. Moreover $s+1 \notin I(e, i)$, hence by construction $\varphi_{e,s+1}^{(1),B_i^{s+1}[s+1]}(y) = \varphi_{e,s}^{(1),B_i^s[s]}(y)$. It follows that $\varphi_e^{(1),B_i}(y) = \varphi_{e,s}^{(1),B_i^s[s]}(y)$ for all such s. Thus $\varphi_e^{(1),B_i}$ is recursive, contradicting our assumption that $P \cap \mathrm{REC} = \emptyset$. This proves 2.

To prove 3, consider the least y such that either $\varphi_e^{(1),B_i}(y)\uparrow$ or

$$\left\langle \varphi_e^{(1),B_i}(x) \;\middle|\; x \leq y \right\rangle \notin T\,.$$

Choose $s_2 > s_1$ such that $\varphi_{e,s}^{(1),B_i^s[s]}(x) \downarrow= \varphi_e^{(1),B_i}(x)$ for all $x < y$ and all $s \geq s_2$. Note that for all $s \geq s_2$ we have

$$\left\langle \varphi_{e,s}^{(1),B_i^s}(x) \;\middle|\; x < y \right\rangle \;=\; \left\langle \varphi_e^{(1),B_i}(x) \;\middle|\; x < y \right\rangle \;\in\; T\,,$$

hence $l(e, i, s) \geq y$ and $r(e, i, s) \geq u(B_i^s[s], e, x, s)$ for all $x < y$.

Case 1: $\varphi_{e,s}^{(1),B_i^s[s]}(y)\uparrow$ for all $s \geq s_2$. Then for all $s \geq s_2$ we have $l(e,i,s) = y$ and $r(e,i,s+1) = r(e,i,s)$.

Case 2: $\varphi_{e,s}^{(1),B_i^s[s]}(y)\downarrow$ for some $s \geq s_2$. Then for any such s we have

$$r(e,i,s) \ \geq \ u(B_i^s[s], e, y, s)\,,$$

hence by construction $\varphi_{e,s+1}^{(1),B_i^{s+1}[s+1]}(y) = \varphi_{e,s}^{(1),B_i^s[s]}(y)$. It follows that for all such s we have $\varphi_e^{(1),B_i}(y) = \varphi_{e,s}^{(1),B_i^s[s]}(y)$, hence

$$\left\langle \varphi_{e,s}^{(1),B_i^s[s]}(x) \;\middle|\; x \leq y \right\rangle \notin T\,,$$

hence again $l(e,i,s) = y$ and $r(e,i,s+1) = r(e,i,s)$.

In either case we have $r(e,i,s+1) = r(e,i,s)$ for all sufficiently large s. Thus 3 holds, and our lemma is proved. □

This completes the proof of the Binns Splitting Theorem. □

Exercise 3.15.3. A Turing degree $\mathbf{b}$ is said to be *low* if $\mathbf{b}' = \mathbf{0}'$. Prove that the r.e. Turing degrees $\mathbf{b}_1 = \deg_T(B_1)$ and $\mathbf{b}_2 = \deg_T(B_2)$ constructed in the proof of the Binns Splitting Theorem 3.15.1 are low.

3.16 Some Additional Results

In this section we mention some additional results and problems concerning r.e. Turing degrees.

More to come ......

Chapter 4

Randomness

It seems appropriate to call an infinite sequence of 0's and 1's "random" if it is the result of an infinite sequence of independent coin tosses using a fair or unbiased coin. The purpose of this chapter is to define and discuss a mathematically rigorous, recursion-theoretic concept of randomness which corresponds to this intuitive, non-mathematical notion. References for this material are Downey/Hirschfeldt [7] and Simpson [16].

4.1 Measure-Theoretic Preliminaries

In this section we present the measure-theoretic background material which we shall need.

Definition 4.1.1 (the Cantor space).

1. The *Cantor space* is the set
$$2^{\mathbb{N}} = \{0,1\}^{\mathbb{N}} = \{X : \mathbb{N} \to \{0,1\}\}\,.$$
Note that each $X \in 2^{\mathbb{N}}$ is an infinite sequence of 0's and 1's, namely $X = \langle X(0), X(1), \ldots, X(n), \ldots \rangle$.

2. We write $\mathrm{Seq}_2 = 2^{<\mathbb{N}}$ = the set of finite sequences of 0's and 1's. According to our Notation 3.14.1, for all $\sigma \in \mathrm{Seq}_2$ and $X \in 2^{\mathbb{N}}$ we have $X \supset \sigma$ if and only if $X[\mathrm{lh}(\sigma)] = \sigma$. For $\sigma \in \mathrm{Seq}_2$ we put
$$N_\sigma = \{X \in 2^{\mathbb{N}} \mid X \supset \sigma\}\,.$$
We view the Cantor space $2^{\mathbb{N}}$ as a topological space with basic open sets N_σ, $\sigma \in \mathrm{Seq}_2$. Thus $U \subseteq 2^{\mathbb{N}}$ is said to be *open* if there exists $G \subseteq \mathrm{Seq}_2$ such that $U = \bigcup_{\sigma \in G} N_\sigma$.

Remark 4.1.2. It is easy to see that the above-defined topology on the Cantor space is the same as the product topology on $2^{\mathbb{N}} = \prod_{n\in\mathbb{N}}\{0,1\}$, where the two-point space $\{0,1\}$ has the discrete topology. Therefore, by Tychonoff's Theorem, $2^{\mathbb{N}}$ is compact.

Remark 4.1.3. The points of $2^{\mathbb{N}}$ are just the infinite sequences of 0's and 1's. Eventually we are going to define what it means for a point of $2^{\mathbb{N}}$ to be *random*. The standard method of formalizing informal notions such as independence and probability is by means of *measure theory*, as we shall now explain.

Definition 4.1.4 (probability measures). Let I be a nonempty set. A *σ-algebra* on I is a set $\mathcal{S} \subseteq P(I)$, the powerset of I, such that $\emptyset \in \mathcal{S}$ and $I \in \mathcal{S}$ and $\mathcal{S}$ is closed under the operations of countable union, countable intersection, and complementation. A *probability measure* on I is a function $\mu : \mathcal{S} \to [0,1]$, where $\mathcal{S}$ is a σ-algebra on I, such that $\mu(\emptyset) = 0$ and $\mu(I) = 1$ and

$$\mu(\bigcup_{n=1}^{\infty} S_n) = \sum_{n=1}^{\infty} \mu(S_n)$$

for all sequences of pairwise disjoint sets $S_1, S_2, \ldots \in \mathcal{S}$. This last property is known as *countable additivity*. A *probability space* is an ordered triple $(I, \mathcal{S}, \mu)$ as above. For $S \in \mathcal{S}$, the *measure of* S is the real number $\mu(S)$.

Remark 4.1.5. Let $(I, \mathcal{S}, \mu)$ be a probability space. A set $S \subseteq I$ such that $S \in \mathcal{S}$ is called an *event*. For $S \in \mathcal{S}$, the measure of S is thought of as the *probability* of the event S, i.e., the likelihood that a "random" or "randomly chosen" element of I will belong to S. Note that we have not yet rigorously defined the concept "random."

Definition 4.1.6 (Borel sets, regularity).

1. Let I be a topological space. The *Borel sets* of I are the smallest σ-algebra on I containing the open sets of I. A *Borel probability measure* on I is a probability measure μ on I such that the domain of μ consists of the Borel sets of I.

2. Let $(I, \mathcal{S}, \mu)$ be a probability space such that I is also a topological space, and $\mathcal{S}$ includes the Borel sets of I. We say that μ is *regular* (with respect to the given topology on I) if, for all $S \in \mathcal{S}$,

$$\mu(S) = \inf\{\mu(U) \mid S \subseteq U \text{ and } U \text{ is open}\}.$$

Theorem 4.1.7 (the fair coin measure). There exists a Borel probability measure μ on $2^{\mathbb{N}}$ such that, for all $\sigma \in \mathrm{Seq}_2$, $\mu(N_\sigma) = 1/2^{\mathrm{lh}(\sigma)}$. Note that μ is unique with these properties. We refer to μ as the *fair coin* measure on $2^{\mathbb{N}}$, because it arises by viewing $X \in 2^{\mathbb{N}}$ as the result of a sequence of independent tosses of a fair coin. It can be shown that μ is regular.

Proof. We omit the proof. See any measure theory textbook. □

Definition 4.1.8 (null sets). In any probability space, a *null set* is any subset of a set of measure 0. Thus $T \subseteq I$ is null if and only if $T \subseteq S$ for some $S \in \mathcal{S}$ such that $\mu(S) = 0$. Note also that, if μ is regular, then T is null if and only if $\forall \epsilon > 0$ $\exists$ open set U such that $T \subseteq U$ and $\mu(U) \leq \epsilon$. Equivalently, $T \subseteq \bigcap_n U_n$ where each U_n is open and $\mu(U_n) \leq 1/2^n$.

4.2 Effective Randomness

In this section we define what it means for a point $X \in 2^{\mathbb{N}}$ to be "effectively random." In this context, "effectively" means "recursion-theoretically."

We first consider what it means for a subset of $2^{\mathbb{N}}$ to be "effectively open." From now on, let μ be the fair coin measure on $2^{\mathbb{N}}$.

Definition 4.2.1. A set $U \subseteq 2^{\mathbb{N}}$ is said to be Σ^0_1 if $U = \{X \in 2^{\mathbb{N}} \mid \exists k\, R(X,k)\}$ where R is recursive. A sequence of sets $U_n \subseteq 2^{\mathbb{N}}$, $n \in \mathbb{N}$, is said to be *uniformly* Σ^0_1 if $U_n = \{X \in 2^{\mathbb{N}} \mid \exists k\, R(X,k,n)\}$ for some fixed recursive predicate R.

Proposition 4.2.2 (effective openness). $U \subseteq 2^{\mathbb{N}}$ is Σ^0_1 if and only if U is *effectively open*, i.e., $U = \bigcup_{\sigma \in G} N_\sigma$ for some recursively enumerable $G \subseteq \mathrm{Seq}_2$. Moreover, we may assume that G is primitive recursive and *pairwise incompatible*, i.e., there are no $\sigma, \tau \in G$ such that $\sigma \subset \tau$. A similar result holds for uniformly Σ^0_1 sequences of sets of $U_n \subseteq 2^{\mathbb{N}}$, $n \in \mathbb{N}$.

Proof. Assume that U is Σ^0_1, say $U = \{X \in 2^{\mathbb{N}} \mid \exists k\, R(X,k)\}$ where R is recursive. Let e be such that $\varphi^{(1),X}_e(0) \simeq$ least k such that $R(X,k)$. Then

$$U \;=\; U_e \;=\; \{X \in 2^{\mathbb{N}} \mid \varphi^{(1),X}_e(0)\downarrow\}\,.$$

A e with this last property is called an *index* of U, or a Σ^0_1 *index* of U. Thus we have a uniform Σ^0_1 indexing of all Σ^0_1 subsets of $2^{\mathbb{N}}$.

We now use the idea of finite approximations from Section 3.14. Given an index e of $U \subseteq 2^{\mathbb{N}}$, define $G \subseteq \mathrm{Seq}_2$ by

$$G \;=\; G_e \;=\; \{\sigma \in \mathrm{Seq}_2 \mid \varphi^{(1),\sigma}_e(0)\downarrow \wedge \neg\exists\tau \subset \sigma\ \varphi^{(1),\tau}_e(0)\downarrow\}\,.$$

Then $U = \bigcup_{\sigma \in G} N_\sigma$ and G is primitive recursive and pairwise incompatible. Conversely, if $U = \bigcup_{\sigma \in G} N_\sigma$ where G is r.e., then clearly U is Σ^0_1. The uniform version is proved similarly. □

Remark 4.2.3. As usual, we relativize as follows. Given an oracle $f \in \mathbb{N}^{\mathbb{N}}$, we say that $U \subseteq 2^{\mathbb{N}}$ is $\Sigma^{0,f}_1$ if $U = \{X \in 2^{\mathbb{N}} \mid \exists k\, R(f \oplus X, k)\}$ where R is recursive. It can be shown that $U \subseteq 2^{\mathbb{N}}$ is open if and only if U is $\Sigma^{0,f}_1$ for some f. Thus we see a close analogy between open sets and r.e. sets. This analogy can be pushed much farther.

The following is a recursion-theoretic analog of Definition 4.1.8.

Definition 4.2.4 (effectively null sets). A set $T \subseteq 2^{\mathbb{N}}$ is said to be *effectively null* if there exists a uniformly Σ^0_1 sequence of sets $U_n \subseteq 2^{\mathbb{N}}$, $n \in \mathbb{N}$, such that $T \subseteq \bigcap_n U_n$ and $\mu(U_n) \leq 1/2^n$ for all n.

We are now ready to define our concept of recursion-theoretic randomness.

Definition 4.2.5 (randomness). A point $X \in 2^{\mathbb{N}}$ is said to be *effectively random*, or just *random*, if X does not belong to any effectively null set. An equivalent condition is that the singleton set $\{X\}$ is not effectively null.

Corollary 4.2.6. $\mu(\{X \in 2^{\mathbb{N}} \mid X \text{ is random}\}) = 1$.

Proof. There are only countably many effectively null sets. By countable additivity, their union is null. The corollary is a restatement of this. $\square$

Remark 4.2.7. The great mathematician Kolmogorov invented probability theory. He also invented a theory of algorithmic randomness, known as *Kolmogorov complexity*. Our concept of randomness was originally formulated in 1966 by Martin-Löf, a former Ph. D. student of Kolmogorov. It can be shown that this concept of randomness is closely related to Kolmogorov complexity.

Proposition 4.2.8. If $X \in 2^{\mathbb{N}}$ is recursive, then X is not random.

Proof. If X is recursive, then the sets $U_n = N_{X[n]}$ are uniformly Σ^0_1 of measure $1/2^n$. Hence $\{X\} = \bigcap_n U_n$ is effectively null, hence X is not random. $\square$

Proposition 4.2.9. If a Π^0_1 set $P \subseteq 2^{\mathbb{N}}$ is of measure 0, then it is effectively null. It follows that no $X \in P$ is random.

Proof. Let e be a Σ^0_1 index of $2^{\mathbb{N}} \setminus P$. Then $P = \{X \in 2^{\mathbb{N}} \mid \varphi_e^{(1),X}(0)\uparrow\}$. Put $V_s = \{X \in 2^{\mathbb{N}} \mid \varphi_{e,s}^{(1),X[s]}(0)\uparrow\}$. Clearly the sets V_s, $s \in \mathbb{N}$ are uniformly Σ^0_1, and $V_0 \supseteq V_1 \supseteq \cdots \supseteq V_s \supseteq \cdots$ and $P = \bigcap_s V_s$. Hence by countable additivity $\lim_s \mu(V_s) = \mu(P) = 0$. The function $s \mapsto \mu(V_s)$ is primitive recursive, so let $h(n) =$ the least $s > n$ such that $\mu(V_s) \leq 1/2^n$. Then h is recursive, hence the sets $U_n = V_{h(n)}$, $n \in \mathbb{N}$, are uniformly Σ^0_1. Moreover $\mu(U_n) \leq 1/2^n$ and $P = \bigcap_n U_n$, so P is effectively null. $\square$

Remark 4.2.10. Let us say that $X \in 2^{\mathbb{N}}$ is *weakly random* if X does not belong to any Π^0_1 subset of $2^{\mathbb{N}}$ of measure 0. We have just shown that if X is random then X is weakly random. The converse does not hold. For example, any Cohen generic $X \in 2^{\mathbb{N}}$ is weakly random but not random.

We pause to mention the following useful property of random sequences. Note that this property is in fact equivalent to randomness.

Theorem 4.2.11 (Solovay). Assume that $A \in 2^{\mathbb{N}}$ is random. If $U_n \subseteq 2^{\mathbb{N}}$, $n \in \mathbb{N}$, are uniformly Σ^0_1 such that

$$\sum_{n=0}^{\infty} \mu(U_n) < \infty\,,$$

then $\{n \in \mathbb{N} \mid A \in U_n\}$ is finite.

Proof. Let c be such that $\sum_{n=0}^{\infty} \mu(U_n) \leq c < \infty$. For each $k \geq 1$ put

$$W_k = \{X \in 2^{\mathbb{N}} \mid \exists^{\geq k} n\ (X \in U_n)\}\,.$$

Note that the sets W_k, $k \geq 1$, are uniformly Σ^0_1. We claim that $\mu(W_k) \leq c/k$. To see this, write

$$W_{k,s} = \{X \in 2^{\mathbb{N}} \mid (\exists^{\geq k} n \leq s)\ (X \in U_n)\}$$

and note that $W_k = \bigcup_s W_{k,s}$. We have

$$\begin{aligned} c \ \geq \ \sum_{n=0}^{\infty} \mu(U_n) \ &\geq \ \sum_{n=0}^{s} \mu(U_n) \\ &= \ \sum_{n=0}^{s} \int_X U_n(X)\, dX \ = \ \int_X \sum_{n=0}^{s} U_n(X)\, dX \\ &\geq \ \int_X k\, W_{k,s}(X)\, dX \ = \ k\, \mu(W_{k,s}) \end{aligned}$$

so $\mu(W_{k,s}) \leq c/k$. It follows that $\mu(W_k) = \sup_s \mu(W_{k,s}) \leq c/k$ as claimed. Let h be primitive recursive such that $h(n) \geq 2^n c$ for all n. Then $W_{h(n)}$ is uniformly Σ^0_1 of measure $\leq 1/2^n$. Since A is random, we have $A \notin \bigcap_n W_{h(n)} = \bigcap_k W_k$, so $A \notin W_k$ for some k, i.e., $\exists^{<k} n\ A \in U_n$. Thus $\{n \mid A \in U_n\}$ is finite. $\square$

We end this section with the following interesting result.

Theorem 4.2.12 (Martin-Löf). The union of all effectively null sets is effectively null.

Corollary 4.2.13. We can write $\{X \in 2^{\mathbb{N}} \mid X \text{ is not random}\} = \bigcap_n U_n$ where U_n is uniformly Σ^0_1 and $\mu(U_n) \leq 1/2^n$ for all n.

Proof. Put $S = \{X \in 2^{\mathbb{N}} \mid X \text{ is not random}\}$. By the definition of randomness, S is the union of all effectively null sets. It follows by Theorem 4.2.12 that S itself is effectively null. Thus $S \subseteq \bigcap_n U_n$ where U_n is uniformly Σ^0_1 and $\mu(U_n) \leq 1/2^n$ for all n. But clearly $\bigcap_n U_n$ itself is effectively null, hence $S = \bigcap_n U_n$. $\square$

Remark 4.2.14 (tests for randomness, effectively null G_δ sets). The following terminology is sometimes used. Define a *test* or *test for randomness* to be an *effectively null G_δ set*, i.e., a set of the form $\bigcap_n U_n$ where U_n is uniformly Σ^0_1 and $\mu(U_n) \leq 1/2^n$ for all n. We say that $X \in 2^{\mathbb{N}}$ *passes the test* if $X \notin \bigcap_n U_n$. By the definition of randomness, X is random if and only if X passes all tests. Corollary 4.2.13 tells us that there is a *universal* test, i.e., a test such that if X passes that test then it passes all tests and is therefore random.

We may reformulate Corollary 4.2.13 by saying that the set $\{X \in 2^{\mathbb{N}} \mid X$ is not random$\}$ is effectively null G_δ. Moreover, it is the largest effectively null G_δ set, which is the same as the largest effectively null set.

Corollary 4.2.15. $\{X \in 2^{\mathbb{N}} \mid X \text{ is random}\}$ is Σ^0_2.

Proof. By the definition of the arithmetical hierarchy, $S \subseteq 2^{\mathbb{N}}$ is Π^0_2 if and only if $S = \bigcap_n U_n$ where U_n is uniformly Σ^0_1. In particular, by Corollary 4.2.13, $\{X \in 2^{\mathbb{N}} \mid X \text{ is not random}\}$ is Π^0_2, hence $\{X \in 2^{\mathbb{N}} \mid X \text{ is random}\}$ is Σ^0_2. $\square$

Corollary 4.2.16. For all $\epsilon > 0$ we can find a Π^0_1 set $P \subseteq 2^{\mathbb{N}}$ such that $\forall X\, (X \in P \Rightarrow X \text{ is random})$ and $\mu(P) > 1 - \epsilon$.

Proof. Put $P = 2^{\mathbb{N}} \setminus U_n$, where U_n is as in Corollary 4.2.13, and $\epsilon > 1/2^n$. $\square$

Remark 4.2.17. We cannot improve Corollary 4.2.16 to say that there exists a Π^0_1 set $P \subseteq 2^{\mathbb{N}}$ such that $\forall X\,(X \in P \Rightarrow X$ is random) and $\mu(P) = 1$. This is obvious, because any nonempty Σ^0_1 subset of $2^{\mathbb{N}}$ is of positive measure.

We now prove Martin-Löf's Theorem 4.2.12.

Notation 4.2.18 (the tilde notation). Recall from the proof of Proposition 4.2.2 that for every Σ^0_1 set $U \subseteq 2^{\mathbb{N}}$ there exists an index e such that

$$U \;=\; U_e \;=\; \{X \in 2^{\mathbb{N}} \mid \varphi_e^{(1),X}(0)\downarrow\} \;=\; \{X \in 2^{\mathbb{N}} \mid \exists s\; \varphi_{e,s}^{(1),X[s]}(0)\downarrow\}\,.$$

Moreover, the sets U_e for all $e \in \mathbb{N}$ are uniformly Σ^0_1. We put

$$U_{e,s} \;=\; \left\{X \in 2^{\mathbb{N}} \;\middle|\; \varphi_{e,s}^{(1),X[s]}(0)\downarrow\right\}$$

and note that $U_e = \bigcup_s U_{e,s}$ and $U_{e,0} \subseteq U_{e,1} \subseteq \cdots \subseteq U_{e,s} \subseteq \cdots$. Note also that $\mu(U_{e,s})$ is a rational number, and the function $(e,s) \mapsto \mu(U_{e,s})$ is primitive recursive. Now, given an index e and a rational number r, define

$$\widetilde{U} \;=\; \widetilde{U}_r \;=\; \widetilde{U}_{e,r} \;=\; \{X \mid \exists s\,(X \in U_{e,s} \wedge \mu(U_{e,s}) \le r)\}\,.$$

Note that $\widetilde{U}_{e,r}$ is uniformly Σ^0_1 for all $e \in \mathbb{N}$ and all rational r. The following properties of $\widetilde{U}_{e,r}$ are easily verified.

1. $\widetilde{U}_{e,r} \subseteq U_e$.
2. $\mu\left(\widetilde{U}_{e,r}\right) \le r$.
3. If $\mu(U_e) \le r$ then $\widetilde{U}_{e,r} = U_e$.

We may describe $\widetilde{U}_{e,r}$ as "U_e enumerated so long as its measure is $\le r$," or "the Σ^0_1 subset of $2^{\mathbb{N}}$ with index e, enumerated so long as its measure is $\le r$."

Proof of Theorem 4.2.12. Define Σ^0_1 sets $V_e \subseteq 2^{\mathbb{N}}$, $e \in \mathbb{N}$, as follows. Given e, compute $\varphi_e^{(1)}(e)$. If $\varphi_e^{(1)}(e)\uparrow$, let $V_e = \emptyset$. If $\varphi_e^{(1)}(e) \simeq i$, let $V_e =$ the Σ^0_1 subset of $2^{\mathbb{N}}$ with index i enumerated so long as its measure is $\le 1/2^e$. Note that V_e is uniformly Σ^0_1 and $\mu(V_e) \le 1/2^e$. Define $S = \bigcap_n U_n$ where $U_n = \bigcup_{e=n+1}^{\infty} V_e$. Note that U_n is uniformly Σ^0_1. Moreover, $\mu(U_n) \le \sum_{e=n+1}^{\infty} \mu(V_e) \le \sum_{e=n+1}^{\infty} 1/2^e = 1/2^n$, so S is effectively null. We claim that $S \supseteq T$ for all effectively null sets T. To see this, suppose T is effectively null, say $T \subseteq \bigcap_n W_n$ where $W_n \subseteq 2^{\mathbb{N}}$ is uniformly Σ^0_1 and $\mu(W_n) \le 1/2^n$. By the Parametrization Theorem, let h be a primitive recursive function such that, for all n, $h(n)$ is a Σ^0_1 index of W_n. Let e be an index of h qua recursive function, i.e., let e be such that $\varphi_e^{(1)}(n) \simeq h(n)$ for all n. Then $\varphi_e^{(1)}(e) \simeq h(e)$ is a Σ^0_1 index of W_e. Since $\mu(W_e) \le 1/2^e$, it follows that $V_e = W_e$, hence $T \subseteq V_e$. Since h has infinitely many indices qua recursive function, we see that $T \subseteq V_e$ for infinitely many e. It follows that $T \subseteq U_n$ for all n, i.e., $T \subseteq S$. This completes the proof. $\square$

4.3 Randomness Relative to an Oracle

In this section we relativize our concept of randomness to a Turing oracle, and we prove some theorems concerning the relativized concept.

Definition 4.3.1. Let $f \in \mathbb{N}^{\mathbb{N}}$ be an oracle. We say that $A \in 2^{\mathbb{N}}$ is *f-random*, or *random over f*, or *random relative to f*, if there is no uniformly $\Sigma_1^{0,f}$ sequence of sets $U_n^f \subseteq 2^{\mathbb{N}}$, $n \in \mathbb{N}$, such that $A \in \bigcap_n U_n^f$ and $\mu(U_n^f) \leq 1/2^n$ for all n.

Remark 4.3.2. Recall that if A is recursive then A is not random. This relativizes to the following statement: if $A \leq_T f$ then A is not f-random.

Notation 4.3.3. For $A, B \in 2^{\mathbb{N}}$ recall that $A \oplus B \in 2^{\mathbb{N}}$ is defined by

$$(A \oplus B)(2n) = A(n)\,, \qquad (A \oplus B)(2n+1) = B(n)\,,$$

for all n. Thus $2^{\mathbb{N}} \times 2^{\mathbb{N}} \cong 2^{\mathbb{N}}$ via the mapping $(A, B) \mapsto A \oplus B$.

Theorem 4.3.4. If $A \oplus B$ is random, then A is random over B, and B is random over A.

Proof. Suppose B is not random over A, say $B \in \bigcap_n V_n^A$ where V_n^A is uniformly $\Sigma_1^{0,A}$ of measure $\leq 1/2^n$. For an arbitrary $X \in 2^{\mathbb{N}}$, let V_n^X be V_n^A with A replaced by X. (More precisely, let V_n^X be the $\Sigma_1^{0,X}$ set with $\Sigma_1^{0,X}$ index $h(n)$ where h is a fixed primitive recursive function such that V_n^A is the $\Sigma_1^{0,A}$ set with $\Sigma_1^{0,A}$ index $h(n)$.) Define

$$W_n \;=\; \left\{ X \oplus Y \;\middle|\; X \in 2^{\mathbb{N}} \text{ and } Y \in \widetilde{V}_n^X \right\}$$

where $\widetilde{V}_n^X$ is V_n^X enumerated so long as its measure is $\leq 1/2^n$. (See our Notation 4.2.18.) Note that W_n is uniformly Σ_1^0. By Fubini's Theorem, $\mu(W_n) \leq 1/2^n$. Since $\widetilde{V}_n^A = V_n^A$ and $B \in V_n^A$, it follows that $A \oplus B \in W_n$ for all n. Thus $A \oplus B$ is not random. We have now proved that if $A \oplus B$ is random then B is random over A. The proof that A is random over B is similar. □

Corollary 4.3.5. If $A \oplus B$ is random, then $A \not\leq_T B$ and $B \not\leq_T A$.

Corollary 4.3.6. If $A \oplus B$ is random, then A and B are random.

The preceding theorem has a converse due to van Lambalgen 1987.

Theorem 4.3.7 (van Lambalgen). If A is random, and if B is random over A, then $A \oplus B$ is random.

Proof. Suppose $A \oplus B$ is not random, say $A \oplus B \in \bigcap_n W_n$ where W_n is uniformly Σ_1^0 and $\mu(W_n) \leq 1/2^n$. By passing to a subsequence, we may assume that $\mu(W_n) \leq 1/2^{2n}$. For all $X \in 2^{\mathbb{N}}$ and $n \in \mathbb{N}$ put

$$V_n^X \;=\; \{Y \in 2^{\mathbb{N}} \mid X \oplus Y \in W_n\}$$

and note that V_n^X is uniformly $\Sigma_1^{0,X}$. For all $n \in \mathbb{N}$ put

$$U_n \;=\; \left\{ X \in 2^{\mathbb{N}} \;\middle|\; \mu(V_n^X) > \frac{1}{2^n} \right\}$$

and note that U_n is uniformly Σ_1^0. By Fubini's Theorem we have

$$\begin{aligned} \mu(W_n) \;&=\; \int_X \int_Y W_n(X \oplus Y)\, dY\, dX \;=\; \int_X \int_Y V_n^X(Y)\, dY\, dX \\ &\geq\; \int_X U_n(X)\, V_n^X(Y)\, dY\, dX \;\geq\; \int_X U_n(X)\, \frac{1}{2^n}\, dX \\ &=\; \frac{1}{2^n} \int_X U_n(X)\, dX \;=\; \frac{1}{2^n}\, \mu(U_n)\,, \end{aligned}$$

hence $\mu(U_n) \leq 2^n \mu(W_n) \leq 2^n/2^{2n} = 1/2^n$. Since $\mu(U_n) \leq 1/2^n$ and A is random, it follows by Solovay's Theorem 4.2.11 that $\{n \mid A \in U_n\}$ is finite. Thus $\mu(V_n^A) \leq 1/2^n$ for all but finitely many n. Moreover V_n^A is uniformly $\Sigma_1^{0,A}$ and $B \in \bigcap_n V_n^A$. Thus B is not random over A. This proves the theorem. $\square$

The following result is due to Joseph Miller 2004.

Theorem 4.3.8 (J. Miller). Assume that A is random, $A \leq_T B$, and B is random over C. Then A is random over C.

Proof. Fix e such that $A = \varphi_e^{(1),B}$. For each $\sigma \in \mathrm{Seq}_2$ put

$$V_\sigma \;=\; \{Y \in 2^{\mathbb{N}} \mid \sigma \subseteq \varphi_e^{(1),Y}\}\,.$$

Note that $B \in V_{A[n]}$ for all n. Moreover, V_σ is uniformly Σ_1^0, and $\sigma \mid \tau \Rightarrow V_\sigma \cap V_\tau = \emptyset$.

Lemma 4.3.9. There is a constant c such that $\mu(V_{A[n]}) \leq c/2^n$ for all n.

Proof. For each rational number c put

$$U_c \;=\; \left\{ X \in 2^{\mathbb{N}} \;\middle|\; \exists n\; \mu\big(V_{X[n]}\big) > \frac{c}{2^n} \right\}.$$

Note that U_c is uniformly Σ_1^0. Let G_c be the set of $\sigma \in \mathrm{Seq}_2$ such that $\mu(V_\sigma) > c/2^{\mathrm{lh}(\sigma)}$ and σ is minimal with this property. Thus $U_c = \bigcup_{\sigma \in G_c} N_\sigma$. Since G_c is pairwise incompatible, the sets V_σ, $\sigma \in G_c$, are pairwise disjoint. We have $\mu(U_c) = \sum_{\sigma \in G_c} 1/2^{\mathrm{lh}(\sigma)}$, hence

$$1 \;=\; \mu(2^{\mathbb{N}}) \;\geq\; \mu\left(\bigcup_{\sigma \in G_c} V_\sigma \right) \;=\; \sum_{\sigma \in G_c} \mu(V_\sigma) \;\geq\; \sum_{\sigma \in G_c} \frac{c}{2^{\mathrm{lh}(\sigma)}} \;=\; c\, \mu(U_c)\,.$$

Thus $\mu(U_c) \leq 1/c$ for all c. In particular, $\mu(U_{2^n}) \leq 1/2^n$ for all n. Since A is random, it follows that $A \notin U_{2^n}$ for some n, hence $A \notin U_c$ for some c. The lemma follows. $\square$

Let c be a rational number as in the lemma. Let $\widetilde{V}_\sigma = V_\sigma$ enumerated so long as its measure is $\leq c/2^{\mathrm{lh}(\sigma)}$. Then $\widetilde{V}_\sigma$ is uniformly Σ^0_1 of measure $\leq c/2^{\mathrm{lh}(\sigma)}$. By the lemma we have $\widetilde{V}_{A[n]} = V_{A[n]}$ for all n.

Now suppose A is not random over C, say $A \in \bigcap_i U_i^C$ where U_i^C is uniformly Σ^0_1 and $\mu(U_i^C) \leq 1/2^i$. As in the proof of Proposition 4.2.2, let $G_i^C \subseteq \mathrm{Seq}_2$ be uniformly $\Sigma^{0,C}_1$ and pairwise incompatible such that $U_i^C = \bigcup_{\sigma \in G_i^C} N_\sigma$. Put $W_i^C = \bigcup_{\sigma \in G_i^C} \widetilde{V}_\sigma$. Then W_i^C is uniformly $\Sigma^{0,C}_1$ and

$$\mu(W_i^C) \;=\; \sum_{\sigma \in G_i^C} \mu(\widetilde{V}_\sigma) \;\leq\; \sum_{\sigma \in G_i^C} \frac{c}{2^{\mathrm{lh}(\sigma)}} \;=\; c\,\mu(U_i^C) \;\leq\; \frac{c}{2^i}\;.$$

Moreover, since $A \in U_i^C$ and $B \in V_{A[n]} = \widetilde{V}_{A[n]}$ for all n, we have $B \in W_i^C$ for all i. Thus B is not random over C. This proves the theorem. □

Definition 4.3.10 (n-randomness). For $n \geq 1$, $A \in 2^{\mathbb{N}}$ is said to be *n-random* if A is random relative to $0^{(n-1)}$. Thus 1-randomness is just randomness, while 2-randomness is randomness relative to the Halting Problem, etc.

Corollary 4.3.11. Assume A is random, $A \leq_T B$, and B is n-random. Then A is n-random.

Proof. This is a special case of Theorem 4.3.8. □

Definition 4.3.12. $A \in 2^{\mathbb{N}}$ is said to be *arithmetically random* if A is n-random for all $n \geq 1$.

Corollary 4.3.13. If A is random, $A \leq_T B$, and B is arithmetically random, then A is arithmetically random.

Bibliography

[1] Stål Aanderaa. A Trachtenbrot inseparability theorem for groups (abstract). *Bulletin of Symbolic Logic*, 1:117, 1995.

[2] Stål Aanderaa and Daniel E. Cohen. Modular machines I, II. In [3], pages 1–18, 19–28, 1980.

[3] S. I. Adian, W. W. Boone, and G. Higman, editors. *Word Problems II: The Oxford Book*. Studies in Logic and the Foundations of Mathematics. North-Holland, 1980. X + 578 pages.

[4] G. Baumslag and C. F. Miller, editors. *Algorithms and Classification in Combinatorial Group Theory*. Mathematical Sciences Research Institute Publications. Springer-Verlag, 1992. VII + 232 pages.

[5] Martin Davis. Hilbert's Tenth Problem is unsolvable. *American Mathematical Monthly*, 80:233–269, 1973.

[6] Martin Davis. *Computability and Unsolvability*. Dover Publications, New York, 1982. XXV + 248 pages.

[7] Rodney G. Downey and Denis Hirschfeldt. *Algorithmic Randomness and Complexity*. 2004. Preprint, May 22, 2004, 305 pages, in preparation.

[8] Charles F. Miller III. Decision problems for groups: survey and reflections. In [4], pages 1–59, 1992.

[9] O. G. Kharlampovich. The universal theory of the class of finite nilpotent groups is undecidable. *Mathematical Notes*, 33:254–263, 1983. The original Russian version appeared in Mat. Zametki 33, 1983, pp. 499–516.

[10] Alexander Nabutovsky. Einstein structures: existence versus uniqueness. *Geometric and Functional Analysis*, 5:76–91, 1995.

[11] Hartley Rogers, Jr. *Theory of Recursive Functions and Effective Computability*. McGraw-Hill, 1967. XIX + 482 pages.

[12] Joseph J. Rotman. *The Theory of Groups*. Allyn and Bacon, second edition, 1973. X + 342 pages.

[13] Joseph R. Shoenfield. *Mathematical Logic*. Addison–Wesley, 1967. VII + 344 pages.

[14] Stephen G. Simpson. Foundations of Mathematics. Unpublished lecture notes, Department of Mathematics, Pennsylvania State University, 86 pages, 1995–2004.

[15] Stephen G. Simpson. Topics in Mathematical Logic: Spring 2004. Unpublished lecture notes, Department of Mathematics, Pennsylvania State University, 105 pages, 2004.

[16] Stephen G. Simpson. Mass problems and randomness. *Bulletin of Symbolic Logic*, 11:1–27, 2005.

[17] A. M. Slobodskoi. Undecidability of the universal theory of finite groups. *Algebra and Logic*, 20:139–156, 1981. The original Russian version appeared in Algebra i Logika 20, 1981, pp. 207–230, 251.

[18] Robert I. Soare. *Recursively Enumerable Sets and Degrees*. Perspectives in Mathematical Logic. Springer-Verlag, 1987. XVIII + 437 pages.