VI

Diophantine Sets and Algorithmic Undecidability

1 The Basic Result

1.1. In §4 of Chapter V we showed that enumerable sets are the same thing as projections of level sets of primitive recursive functions. The projections of the level sets of a special kind of primitive recursive function—polynomials with coefficients in $\mathbf{Z}^+$—are called *Diophantine sets*. We note that this class does not become any larger if we allow the coefficients in the polynomial to lie in $\mathbf{Z}$. The basic purpose of this chapter is to prove the following deep result:

1.2. **Theorem** (M. Davis, H. Putnam, J. Robinson, Yu. Matiyasevič). *All enumerable sets are Diophantine.*

The plan of proof is described in §2. §§3–7 contain the intricate yet completely elementary constructions that make up the proof itself; these sections are not essential for understanding the subsequent material, and may be omitted if the reader so desires.

In §8 we use Theorem 1.2 to prove the existence of versal families of enumerable sets and functions. Recall that in §5 of Chapter V this result was shown to imply that enumerable sets exist that are undecidable, a fact we shall use in Section 1.3 below.

In §7, which stands somewhat apart from the rest of the chapter, we define the Kolmogorov complexity of recursive functions, establish the basic properties of this concept, and prove that the problem of computing the complexity is algorithmically undecidable.

In Chapter VII the following corollary of Theorem 1.2 will be used in an essential way: *enumerable sets are definable in* $\mathrm{L}_1\mathrm{Ar}$. In fact, by their very definition, Diophantine sets are defined by formulas of the form $\exists x_1 \cdots \exists x_n(p)$, where p is an atomic formula.

In the remainder of this section we describe the principal applications of Theorem 1.2: settling Hilbert's tenth problem, constructing polynomials that take only and all prime number values in $\mathbf{Z}^+$, and so on.

Yu. I. Manin, *A Course in Mathematical Logic for Mathematicians, Second Edition*, Graduate Texts in Mathematics 53, DOI 10.1007/978-1-4419-0615-1_6,

1.3. *Hilbert's tenth problem.* Hilbert stated it as follows:

> Suppose we are given a Diophantine equation with an arbitrary number of unknowns and with rational integer coefficients. Give a way in which it is possible to determine after a finite number of operations whether this equation is solvable in rational integers.

We show that the combination of Theorem 1.2, Theorem 5.8 of Chapter V (which follows from Theorem 1.2), and Church's thesis implies that this problem is undecidable.

First of all, any natural number is the sum of four integer squares (Lagrange). Hence $f(x_1, \dots, x_n) = 0$ is solvable in $(\mathbf{Z}^+)^n$ if and only if the equation $f(1 + \Sigma_{i=1}^4 y_{i1}^2, \dots, 1 + \Sigma_{i=1}^4 y_{in}^2) = 0$ is solvable in $(\mathbf{Z})^{4n}$. Consequently, it is sufficient to show that the mass problem "determining whether there are solutions in $(\mathbf{Z}^+)$" (see Section 2.6 of Chapter V) is algorithmically undecidable.

Let $E \subset \mathbf{Z}^+$ be an enumerable set that is not decidable. We represent E as the projection onto the t-coordinate of the 0-level of the polynomial $f_t = f(t; x_1, \dots, x_n)$, where $f \in Z[t, x_1, \dots, x_n]$. The equation $f_{t_0} = 0, t_0 \in \mathbf{Z}^+$, has a solution if and only if $t_0 \in E$. By the discussion in §2 of Chapter V, the corresponding mass problem for the family $\{f_t\}$ is algorithmically decidable if and only if the characteristic function of E is computable. But by our choice of E, this characteristic function is only semicomputable.

Thus, solvability in integers cannot be determined algorithmically even for a suitable one-parameter family of equations. The number of unknowns in the equation, and, in general, the codimension of the projection in Theorem 1.2, can be reduced to 13 (Matiyasevič, Robinson). The precise minimum is not known, although it is an interesting problem.

Finally, it should be noted that the construction of a Diophantine representation for any enumerable set E is completely effective in the sense that given a recursive description of f with $D(f) = E$ or of g with $g(\mathbf{Z}^+) = E$, we can write out the corresponding polynomial explicitly. The same holds for the construction of versal families, of an enumerable undecidable set, and so on. These are all constructive assertions, and not simple existence theorems.

1.4. *Polynomials that represent the prime numbers.* The search for "explicit formulas" for prime numbers was a traditional occupation of dedicated number theory enthusiasts for many centuries. Euler found the polynomial $x^2 + x + 41$, which takes a long series of only prime values. But it has long been known that the set of values at integer points of a polynomial f in $\mathbf{Z}[x_1, \dots, x_n]$ cannot consist entirely of prime numbers: for example, if p and q are two sufficiently large primes, then the congruence $f \equiv 0 \bmod pq$ can be solved (in infinitely many ways). On the other hand, the problem becomes solvable in the class of primitive recursive functions: the function $\{i \mapsto \text{the } i\text{th prime}\}$ is itself primitive recursive (see §1 of Chapter VII), but for trivial reasons.

The nontrivial statement of the problem and the problem's solution involve Theorem 1.2: the set of prime numbers is the set of all *positive values at points in* $(\mathbf{Z}^+)^n$ *of a certain polynomial in* $\mathbf{Z}[x_1, \dots, x_n]$ (or, if we prefer, n may be

replaced by $4n$; see the reduction step in 1.3). Matiyasevič showed that there is a suitable polynomial of degree 37 in 24 variables.

This is actually a general result that has nothing to do with the specific properties of prime numbers:

1.5. **Proposition.** *Let $E \subset \mathbf{Z}^+$ be a Diophantine set. Then there exists a polynomial $g \in \mathbf{Z}[x_0, \dots, x_n]$ such that E coincides with the set of positive values of g at points in $(\mathbf{Z}^+)^{n+1}$.*

Proof. Let E be the projection of the 0-level of the polynomial $f(x_0, x_1, \dots, x_n)$ onto the x_0-coordinate. We set

$$g = x_0[1 - f^2(x_0, x_1, \dots, x_n)].$$

Clearly, the positive values of g are precisely the elements of E. □

It remains only to use the fact that the set of prime numbers is decidable, and hence Diophantine by Theorem 1.2.

The following sets are also sets of positive integer values of polynomials:

1.6. *The sequences* $\{1, 10, 100, \dots, 10^k, \dots\}$ and $\{1, 2^2, 3^{3^3}, \dots, n^{n^{\cdot^{\cdot^{\cdot^n}}}}(n \text{ times}), \dots\}$.

It is amazing that the values of the corresponding polynomials can drop to zero and below in neighborhoods of points where these values are so large.

1.7. *The Fermat set* $\{n | n > 2$ *and* $x^n + y^n + z^n = 0$ *is solvable in* $\mathbf{Z}\}$. Thus, the variable n can be moved from the exponent to the coefficients of a Diophantine equation.

1.8. The set $\{10\varepsilon_1, 10^2\varepsilon_2, \dots, 10^n\varepsilon_n, \dots\}$, where ε_i is the ith digit after the decimal point in the decimal expansion of e (or π or $\sqrt[3]{2}$, or any other "computable" irrational number).

1.9. *The set of all partial fractions in the continued fraction expansion of e, or π, or $\sqrt[3]{2}$.*

We recall that in the case of $\sqrt[3]{2}$ it is not known whether this set is finite or infinite.

These examples show that many number-theoretic questions reduce to problems of the solvability of Diophantine equations. In Chapter VII we shall see that in a certain sense, "almost all of mathematics" reduces to such problems.

2 Plan of Proof

2.1. In this section we introduce some auxiliary notions and give the plan of proof for Theorem 1.2.

We shall temporarily introduce a class of sets that are intermediate between enumerable and Diophantine sets. In order to define this class, we

consider the map that to every subset $E \subset (\mathbf{Z}^+)^n$ associates the set $F \subset (\mathbf{Z}^+)^n$ that is given by the following rule:

$$\langle x_1, \ldots, x_n \rangle \in F \Leftrightarrow \forall k \in [1, x_n], \langle x_1, \ldots, x_{n-1}, k \rangle \in E.$$

We shall say that F is obtained from E by applying the bounded universal quantifier to the nth coordinate. We define similarly the operation of applying the bounded universal quantifier to any coordinate.

2.2. **Definition-Lemma**. *Consider the following three classes of subsets of* $(\mathbf{Z}^+)^n$ *for each n.*

(I) *Projections of level sets of primitive recursive functions.*

(II) *The least class of sets that contains the level sets of polynomials with integer coefficients and that is closed with respect to taking finite direct products, finite unions, finite intersections, projections, and applying the bounded universal quantifier.*

(III) *Projections of level sets of polynomials with integer coefficients.*

The following assertions hold for these classes:

(a) *The class* (I) *coincides with the class of enumerable sets, and the class* (III) *coincides with the class of Diophantine sets. We shall call sets in the class* (II) *D-sets.*

(b) (I) $\supset$ (II) $\supset$ (III).

Proof.

(a) In Theorem 4.3 of Chapter V we showed that the class of primitive enumerable sets coincides with the class of enumerable sets. The rest of (a) merely consists of definitions.

(b) Only the inclusion (II) $\subset$ (I) is not completely obvious. First of all, the m-level set of a polynomial f is the same as the 1-level set of the primitive recursive function $(f-m)^2+1$. Hence, to verify (II) $\subset$ (I) it suffices to show that the class (I) is closed with respect to (finite) direct product, union, intersection, and the bounded universal quantifier. All except for the last of these were established in Lemma 4.8 of Chapter V.

Finally, suppose F is the image of a primitive enumerable set E under the bounded universal quantifier:

$$\langle x_1, \ldots, x_{n-1}, x_n \rangle \in F \Leftrightarrow \forall k \leqslant x_n, \quad \langle x_1, \ldots, x_{n-1}, k \rangle \in E.$$

Starting with the function $f(x_1, \ldots, x_{n-1},\ x_n; y_1, \ldots, y_m)$ whose 1-level projects onto E, we want to construct a function g whose 1-level projects onto F. A natural idea is to consider as an approximation to g the product

$$\prod_{k=1}^{x_n} f(x_1, \ldots, x_{n-1}, k; y_{1k}, \ldots, y_{mk}),$$

where the y_{ik} are "independent variables." The only problem is that the number of arguments of this "function" increases with x_n. To deal with this, we apply

the Gödel function $\mathrm{Gd}(k,t)$, which was defined in Section 4.9 of Chapter V. The function g will now depend on $x_1,\dots,x_n$ and on m additional arguments $t_1,\dots,t_m$:

$$g(x_1,\dots,x_n;\ t_1,\dots,t_m) = \prod_{k=1}^{x_n} f(x_1,\dots,x_{n-1},k;\mathrm{Gd}(k,t_1),\dots,\mathrm{Gd}(k,t_m)).$$

This function is primitive recursive, because the kth factor is obtained from f and Gd by substitution and identifying arguments, and then g is constructed from such factors by recursion.

We now verify that the set F is the projection of the 1-level of g onto the $\langle x_1,\dots,x_n\rangle$-coordinates. In fact, if $g(x_1,\dots,t_m)=1$, then for all $1\leqslant k\leqslant x_n$ we have $f(x_1,\dots,x_{n-1},k,\mathrm{Gd}(k,t_1),\dots,\mathrm{Gd}(k,t_m))=1$, i.e., for all $1\leqslant k\leqslant x_n$ the point $\langle x_1,\dots,x_{n-1},k\rangle$ belongs to E. This means that $\langle x_1,\dots,x_n\rangle\in F$.

Conversely, if $\langle x_1,\dots,x_n\rangle\in F$, then for $1\leqslant k\leqslant x_n$ we can lift the point $\langle x_1,\dots,x_{n-1},k\rangle$ to the 1-level of f. Let the y-coordinates of the resulting point be $y_{1,k},\dots,y_{m,k}$. We solve the following system of equations for the t_i:

$$\mathrm{Gd}(k,t_i)=y_{i,k},\quad \text{for all } 1\leqslant k\leqslant x_n.$$

This is possible by the fundamental property of Gd. The resulting values for the t_i, along with $x_1,\dots,x_n$, make g equal to one. This completes the proof of Lemma 2.2. □

2.3. The plan for the rest of the proof of Theorem 1.2 is as follows. In §3 we show that the classes (I) and (II) coincide, and in §§4–7 we show that (II) and (III) coincide.

2.4. *Remark.* In the course of proving Lemma 2.2, we obtained the following facts, which should always be kept in mind in what follows:

(a) In the definitions of the classes (I)–(III) we may always replace "level sets" by "1-level sets" (by going from f to $(f-m)^2+1$).
(b) All of the classes (I)–(III) are closed with respect to (finite) products, intersections, unions, and also projections. (The proof of this for the class (I) in Lemma 4.8 of Chapter V is also applicable to the class (III).)

We encounter much greater difficulty in treating the bounded universal quantifier. Indeed, the most technical part of the proof in §§4–7 is concerned with showing that the class of Diophantine sets is closed with respect to the bounded universal quantifier.

3 Enumerable Sets Are *D*-Sets

Let $f:(\mathbf{Z}^+)^n\to\mathbf{Z}^+$ be a primitive recursive function. Its 1-level can be represented as the projection onto the first n coordinates of the set $\Gamma_f\cap[(\mathbf{Z}^+)^n\times\{1\}]$,

where Γ_f is the graph of f. Thus, an enumerable set can be obtained as a projection of the intersection of the graphs of two primitive recursive functions. Since, by definition, the class of D-sets is closed with respect to projections and intersections, the assertion in the title of this section follows from the following fact:

3.1. **Proposition**. *The graphs of primitive recursive functions are D-sets.*

Proof. The graphs of the basic functions are Diophantine. The stability of the property of graphs "being D-sets" relative to the composition and juxtaposition of functions is verified by the same arguments as in the proof of Lemma 4.8 of Chapter V. It remains to prove the stability under recursion. We shall first of all need information about the graph of Gödel's function. Here it is more convenient to use gd instead of Gd.

3.2. **Lemma**. *The graph of the Gödel function* $\mathrm{gd}(u,k,t) = \mathrm{rem}(1 + kt,\ u)$ *is Diophantine, and a fortiori, a D-set.*

Proof. The set

$$\Gamma_{\mathrm{gd}} = \{\langle u,k,t,\gamma\rangle | \gamma \text{ is the remainder when } u \text{ is divided by } 1 + kt\}$$

is the intersection of the following two sets in $(\mathbf{Z}^+)^4$:

$$E_1 : \gamma \leqslant 1 + kt;$$
$$E_2 : u - \gamma \geqslant 0 \quad \text{and is divisible by } 1 + kt.$$

Both E_1 and E_2 are Diophantine. In fact, E_1 is a projection of the 0-level of the polynomial $2 + kt - \gamma - y_1$, and E_2 is a projection of the 0-level of the polynomial $u - \gamma - (1 + kt)(y_2 - 1)$. The lemma is proved. □

3.3. **Corollary**. *Let f and g be functions of n and $n+2$ arguments, respectively, whose graphs are D-sets. Then the following equations determine D-sets in the $(x_1,\ldots,x_{n+1},u,t,\ldots)$-coordinate space (where any additional coordinates may follow the t):*

$$E : \mathrm{gd}(u,1,t) = f(x_1,\ldots,x_n);$$
$$F : \mathrm{gd}(u,x_{n+1}+1,t) = g(x_1,\ldots,x_{n+1},\mathrm{gd}(u,x_{n+1},t)).$$

Proof. Introducing extra coordinates after the t amounts to taking the direct product with $(\mathbf{Z}^+)^p$, and this, of course, takes D-sets to D-sets.

E can be represented as a projection of the intersection of the sets $\mathrm{gd}(u,k,t) = w$, $f(x_1,\ldots,x_n) = w$, and $k = 1$ (where k and w are auxiliary coordinates). Since Γ_{gd} and Γ_f are D-sets, the same is true for E.

Similarly, F can be represented as a projection of the intersection of the sets

$$\mathrm{gd}(u,x_{n+1}+1,t) = w_1,$$
$$\mathrm{gd}(u,x_{n+1},t) = w_2,$$
$$g(x_1,\ldots,x_{n+1},w_2) = w_1.$$

These are D-sets, because Γ_g and Γ_{gd} are D-sets. □

3.4. Proof of Proposition 3.1. Recall that it remains to verify the following assertion: Let h be the function defined recursively from functions f and g by the equations

$$h(x_1, \dots, x_n, 1) = f(x_1, \dots, x_n),$$
$$h(x_1, \dots, x_n, k+1) = g(x_1, \dots, x_n, k, h(x_1, \dots, x_n, k));$$

then the graph Γ_h,

$$\langle x_1, \dots, x_{n+1}, \eta \rangle \in \Gamma_h \Leftrightarrow \eta = h(x_1, \dots, x_{n+1}),$$

is a D-set whenever the graphs Γ_f and Γ_g are D-sets.

First step. We set $\Gamma_h = \Gamma^1 \cup \Gamma^2$, where $x_{n+1} = 1$ on Γ^1 and $x_{n+1} \geqslant 2$ on Γ^2. Since

$$\langle x_1, \dots, x_{n+1}, \eta \rangle \in \Gamma^1 \Leftrightarrow x_{n+1} = 1 \quad \text{and} \quad \eta = f(x_1, \dots, x_n),$$

it follows that Γ^1 is the intersection of $\Gamma_f \times \mathbf{Z}^+$ and a D-set, and therefore is a D-set. It remains to verify that Γ^2 is also a D-set.

Second step. In the $(x_1, \dots, x_{n+1}, \eta, u, t)$-coordinate space we consider the sets

$$E_1 : \eta = \mathrm{gd}(u, x_{n+1}, t),$$
$$E_2 : \mathrm{gd}(u, 1, t) = f(x_1, \dots, x_n),$$
$$E_3 : x_{n+1} > 1, \quad \mathrm{gd}(u, k, t) = g(x_1, \dots, x_n, k-1, \mathrm{gd}(u, k-1, t))$$
$$\text{for all } 2 \leqslant k \leqslant x_{n+1}.$$

It is easy to see that $\Gamma^2 = \mathrm{pr} \cap_{i=1}^3 E_i$. In fact, as in §4 of Chapter V, we obtain inclusion in one direction by comparing E_2 and E_3 with the inductive definition of h, and in the other direction by suitably choosing the parameters u and t in Gödel's function. Thus, it remains to show that the E_i are D-sets.

Third step. E_1 is the graph of gd with some additional coordinates. E_2 was shown to be a D-set in the proof of Corollary 3.3.

Finally, E_3 is "almost" obtained from the set F in Corollary 3.3 by applying the bounded universal quantifier to the x_{n+1}-coordinate. More precisely (for brevity, we ignore the η-coordinate);

$$\langle x_1, \dots, x_{n+1}, u, t \rangle \in E_3 \Leftrightarrow \forall k \in [2, x_{n+1}], \langle x_1, \dots, x_n, k-1, u, t, \rangle \in F$$
$$\Leftrightarrow \forall k \in [1, x_{n+1} - 1], \langle x_1, \dots, x_n, k, u, t, \rangle \in F.$$

Consequently, if we apply to F the bounded universal quantifier in the x_{n+1}-coordinate, we obtain a D-set that is the same as E_3 with the x_{n+1}-coordinates of all its points decreased by 1. So it remains to see that the operation of shifting back by 1 preserves the property of "being a D-set," and this follows easily from the definitions. The proof is complete. □

4 The Reduction

4.1. The next three sections are devoted to proving that the class of D-sets coincides with the class of Diophantine sets. As noted at the end of §2, it suffices to show that the class of Diophantine sets is closed with respect to the bounded universal quantifier.

Let $f(x_1, \dots, x_n, k, y_1, \dots, y_m)$ be any nonconstant polynomial with integer coefficients. f will be fixed for the duration of this section. Let d be the degree of f, and let c be the sum of the absolute values of its coefficients.

We define the set E by the condition

$$\langle x_1, \dots, x_n, y\rangle \in E \Leftrightarrow \forall k \leqslant y\, \exists \langle y_1, \dots, y_m\rangle,$$
$$f(x_1, \dots, x_n, k, y_1, \dots, y_m) = 0.$$

We want to show that E is Diophantine. In this section we prove the following reduction step, which is due to Davis, Putnam, and Robinson.

4.2. **Proposition.** *E is Diophantine if the following three sets are Diophantine*:

$$x_1 = x_2^{x_3};$$
$$x_1 = x_2!;$$
$$\frac{x_1}{x_2} = \binom{x_3/x_4}{x_5}, \qquad x_3 \geqslant x_4 x_5,$$

where $\binom{n}{k} = n(n-1)\cdots(n-k+1)/k!$ *is the "binomial coefficient."*

The proof of this and all subsequent propositions of this type follows a standard pattern. To show that E is Diophantine, we introduce auxiliary sets E_i with the following properties:

(a) $E = \bigcap_{i=1}^{N} E_i$;

(b) the E_i are Diophantine.

But usually we are not able to establish directly that all the E_i are Diophantine, so we apply the same procedure to certain of the E_i. Thus, the proof that E is Diophantine has a treelike pattern.

The exposition of each step will consist of the following stages: the introduction of auxiliary variables, which disappear when we project; explicit construction of the sets E_i; the proof of the inclusion $E \subset \operatorname{pr} \cap_{i=1}^{N} E_i$; and the proof of the inclusion $E \supset \operatorname{pr} \cap_{i=1}^{N} E_i$.

4.3. Proof of Proposition 4.2. We denote the auxiliary variables by the symbols Y, N, K, $Y_1, \dots, Y_m$. We introduce the sets E_i in the $\langle x_1, \dots, x_n, y, Y, N, K, Y_1, \dots, Y_m\rangle$-space by the following relations:

$$E_1: \quad N \geqslant c \cdot (x_1 \cdots x_n y Y)^d, \qquad Y < Y_1, \dots, Y < Y_m$$

(intuitively speaking, the right side of the first inequality gives a rough estimate for the value of the polynomial f at the point $\langle x_1, \ldots, x_n, y, y_1, \ldots, y_m \rangle$ if all $y_i \leqslant Y$).

$$E_2: \quad 1 + KN! = \prod_{k=1}^{y} (1 + kN!)$$

(this is a "large modulus"; $f = 0$ will be replaced by divisibility by this modulus).

$$E_3: f(x_1, \ldots, x_n, K, Y_1, \ldots, Y_m) \equiv 0 \bmod (1 + KN!);$$
$$E_{3+i}: \prod_{j<Y} (Y_j - j) \equiv 0 \bmod (1 + KN!), \quad i = 1, \ldots, m.$$

We define the set E' as $\cap_{i=1}^{m+3} E_i$.

Proof of the inclusion $E \subset \mathrm{pr}\, E'$. Given a point $\langle x_1, \ldots, x_n, y \rangle \in E$, we must choose values for the other coordinates so that the relations $E_1, \ldots, E_{m+3}$ are fulfilled.

By the definition of E, each point $\langle x_1, \ldots, x_n, k \rangle, k \leqslant y$, can be lifted to the 0-level of f:

$$f(x_1, \ldots, x_n, k, y_{1k}, \ldots, y_{mk}) = 0.$$

For Y we take the maximum of y and the y_{ik}. Then, as before, we find the Y_i and N by solving the system of Gödel equations

$$\mathrm{gd}(Y_i, k, N!) = y_{ik}, \quad \text{for all } 1 \leqslant k \leqslant y.$$

The proof of Gödel's lemma shows that the Y_i and N may be taken arbitrarily large, in particular, so as to satisfy E_1. The number K is uniquely determined by E_2.

All the choices have now been made. The relation E_{3+i} holds because by the definition of Y_i and gd, we can find a number $Y_i - j$ with $j \leqslant Y$, namely $j = y_{ik}$, such that $Y_i - j \equiv 0 \bmod (1 + kN!)$, for every $k \leqslant y$. Hence, the product on the left in E_{3+i} is divisible by all the $1 + kN!, 1 \leqslant k \leqslant y$, which are pairwise relatively prime, since $N \geqslant y$ by E_1. Therefore, this product is divisible by $1 + KN!$.

Finally, to verify E_3 we note that E_2 implies the congruence $K \equiv k \bmod (1 + kN!)$, $1 \leqslant k \leqslant y$, because $(1 + KN!) - (1 + kN!) \equiv 0 \bmod (1 + kN!)$. But then, since $y_{ik} \equiv Y_i \bmod (1 + kN!)$ by our choice of Y_i, we find that

$$\begin{aligned} f(x_1, \ldots, x_n, K, Y_1, \ldots, Y_m) &\equiv f(x_1, \ldots, x_n, k, y_{ik}, \ldots, y_{mk}) \\ &\equiv 0 \bmod (1 + kN!). \end{aligned}$$

Since the moduli $1 + kN!$ are pairwise relatively prime, this congruence implies E_3.

Proof of the inclusion $\mathrm{pr}\, E' \subset E$. Given a point

$$\langle x_1, \ldots, x_n, y, Y, N, K, Y_1, \ldots, Y_m \rangle$$

whose coordinates satisfy the relations $E_1, \dots, E_{m+3}$, we must find a vector $\langle y_{1k}, \dots, y_{mk} \rangle$ for each $k \leqslant y$ such that

$$f(x_1, \dots, x_n, k, y_{1k}, \dots, y_{mk}) = 0.$$

To do this we let p_k denote any prime divisor of $1 + kN!$, and we set

$$y_{ik} = \textit{the remainder when } Y_i \textit{ is divided by } p_k.$$

We claim that these y_{ik} give us the required equality. In fact, E_3 implies that $f(x_1, \dots, x_n, k, y_{1k}, \dots, y_{mk}) \equiv 0 \mod p_k$. It suffices to show that the number on the left is less than p_k. We have

$$\begin{aligned} p_k \text{ divides } \prod_{j \leqslant Y} (Y_i - j) \text{ by } E_{3+i} \\ \Rightarrow p_k \text{ divides } Y_i - j \text{ for some } j \leqslant Y \\ \Rightarrow y_{ik} = \text{the remainder when } Y_i \text{ is divided by } p_k \leqslant Y \\ \Rightarrow f(x_1, \dots, x_n, k, y_{1k}, \dots, y_{mk}) \leqslant c(x_1 \cdots x_n Y)^d \leqslant N < p_k, \end{aligned}$$

where the second inequality in the last line follows from E_1, and the third inequality follows because p_k divides $1 + kN!$.

Conclusion of the proof. It remains to show that the sets $E_1, \dots, E_{m+3}$ are Diophantine if the sets in Proposition 6.1 are Diophantine. In fact, if we trivially introduce new variables and make substitutions, we can first reduce the verification that all the E_i are Diophantine to showing that the following sets are Diophantine:

$$\begin{aligned} x_1 &= x_2!; \\ x_1 &= \prod_{k \leqslant x_2} (1 + kx_3); \\ x_1 &= \prod_{j \leqslant x_3} (x_2 - j), \qquad x_2 > x_3. \end{aligned}$$

It then remains to notice that the second of these relations can be written in the form

$$x_1 = x_3^{x_2} \left[\begin{matrix} \frac{1}{x_3} + x_2 \\ x_2 \end{matrix} \right],$$

and the third relation can be written as

$$x_1 = x_3! \begin{pmatrix} x_2 - 1 \\ x_3 \end{pmatrix}, \qquad x_2 > x_3.$$

This completes the proof of Proposition 4.2. □

5 Construction of a Special Diophantine Set

5.1. In this section we begin the proof that the three sets in Proposition 4.2 are Diophantine. In order that the reader may better appreciate this stage in the proof, we mention that the most troublesome obstacle here is the rapid growth of one of the coordinates in comparison to the others (for example, $x_1 = x_2!$). J. Robinson had the following key idea. She proved that if we know that *any* specific set in $(\mathbf{Z}^+)^2$ is Diophantine and has one coordinate that grows faster than any power of the other but slower than, say, x^x (for example, exponentially), we may then conclude that *all* enumerable sets are Diophantine. After this, Matiyasevič and Čudnovskiĭ were able to show that a certain set of that type (connected with Fibonacci numbers) is Diophantine. For a history of the question, see Matijasevič's article "Diophantine Sets" in *Uspehi Mat. Nauk*, vol. XXVII, No. 5 (1972) (translated in *Russian Math. Surveys*).

In this section we give a construction that is an improved version of the original construction. Its idea is based on the following observation. Let $x^2 - dy^2 = 1$ be Pell's equation (where $d \in \mathbf{Z}^+$ is not a perfect square). Its solutions $\langle x, y\rangle \in (\mathbf{Z}^+)^2$ form a semigroup with composition law

$$(x_1 + y_1\sqrt{d})(x_2 + y_2\sqrt{d}) = x_3 + y_3\sqrt{d}.$$

This is a cyclic semigroup. That is, let $\langle x_1, y_1\rangle$ be the solution with the least first coordinate. Then any other solution has the form $\langle x_n, y_n\rangle$, where $n \in \mathbf{Z}^+$, and

$$x_n + y_n\sqrt{d} = (x_1 + y_1\sqrt{d})^n.$$

We call n the *number* of the solution $\langle x_n, y_n\rangle$.

The coordinates x_n and y_n grow exponentially with n, so that the set of solutions of Pell's equation, and also the projections of this set on the x- and y-axes, are Diophantine sets having logarithmic density. This is not yet enough: we still have the problem of including the solution number n among the coordinates of a Diophantine set. Only then can we apply Robinson's technique. This is what will be done below.

5.2. *Notation.* We consider Pell's equation with variable d. Its first solution generally varies as a function of d in an uncontrollable fashion, so that it is convenient to choose only those d whose first solutions have the simple special form $\langle a, 1\rangle, a \in \mathbf{Z}^+$. Obviously, then $d = a^2 - 1$.

We shall call the equation $x^2 - (a^2 - 1)y^2 = 1$ the *a-equation*. We define the two sequences $x_n(a)$ and $y_n(a)$ as the coordinates of its nth solution:

$$x_n(a) + y_n(a)\sqrt{a^2 - 1} = \left(a + \sqrt{a^2 - 1}\right)^n.$$

For each n, a formal definition of $x_n(a)$ and $y_n(a)$ as polynomials in a can easily be given by induction on n. Then the expressions $x_n(a)$ and $y_n(a)$ will make sense for all $n \in \mathbf{Z}$ and $a \in \mathbf{C}$. In particular,

$$x_n(1) = 1, \quad y_n(1) = n;$$

and all the formulas given below remain true.

The basic result of this section is the following:

5.3. **Proposition.** *The set*

$$E: \quad y = y_n(a), \qquad a > 1;$$

in the $\langle y, n, a\rangle$-space is Diophantine.

The proof uses the elementary number-theoretic properties of the sequences $x_n(a)$ and $y_n(a)$, most of which will be verified at the end of the section (see 5.8). The idea for determining n in a Diophantine way from $\langle y, a\rangle$ is to observe that $y_n(a) \equiv n \bmod(a-1)$ (Lemma 5.4). This uniquely determines n as long as $n < a-1$. To pass to the general case, we introduce an auxiliary A-equation with A large, and find formulas for its nth solution (using y) in which n appears in only a Diophantine context.

Formally, the proof that E is Diophantine follows the pattern described in 4.2. In addition to the basic variables y, n, a, we introduce six auxiliary variables: x, x_1, y_1, A, x_2, y_2. We set

$$\begin{aligned}
&E_1: \ y \geqslant n, \qquad a > 1;\\
&E_2: \ x^2 - (a^2-1)y^2 = 1;\\
&E_3: \ y_1 \equiv 0 \bmod 2x^2y^2;\\
&E_4: \ x_1^2 - (a^2-1)y_1^2 = 1;\\
&E_5: \ A = a + x_1^2(x_1^2 - a);\\
&E_6: \ x_2^2 - (A^2-1)y_2^2 = 1;\\
&E_7: \ y_2 - y \equiv 0 \bmod x_1^2;\\
&E_8: \ y_2 \equiv n \bmod 2y.
\end{aligned}$$

Let $E' = \cap_{i=1}^{8} E_i$. We show that $\operatorname{pr} E' = E$.

The *inclusion* $E \subset \operatorname{pr} E'$. Given $\langle y, n, a\rangle \in E$, we must find values for the other variables such that $E_1, \ldots, E_8$ hold. As before, we shall not introduce any new symbols for these values; after we choose, say, a value for x, the letter x will become the name for this value.

E_1 is automatically satisfied: $y_n(a) \geqslant n$ for all $a \geqslant 1, n \geqslant 1$ (induction on n). We find x uniquely from E_2 : $x = x_n(a)$. We take $\langle x_1, y_1/2x^2y^2\rangle$ to be any solution of the Pell equation $X^2 - (a^2-1)(2x^2y^2)^2Y^2 = 1$; this gives E_4. A is found uniquely from E_5. We take $\langle x_2, y_2\rangle$ to be the nth solution of the A-equation. Now all choices have been made. To verify E_7 and E_8 we need two lemmas.

5.4. **Lemma.** $y_k(a) \equiv k \bmod(a-1)$.

5.5. **Lemma.** *If $a \equiv b \bmod c$, then $y_n(a) \equiv y_n(b) \bmod c$.*

These lemmas will be proved in 5.8.

We use these lemmas as follows. From E_5 we obtain

$$A = a + (1 + (a^2-1)y_1^2)(1 + (a^2-1)y_1^2 - a) \equiv 1 \bmod 2y,$$

because of E_3. Lemma 5.4 then gives $y_2 = y_n(A) \equiv n \bmod 2y$; this is E_8. Lemma 5.5 gives $y_n(A) \equiv y_n(a) \bmod x_1^2$ (because of E_5); this is E_7.

The inclusion pr $E' \subset E$. From the relations $E_1, \ldots, E_8$ we have only to prove that n is the number of the solution $\langle x, y\rangle$. Note that n occurs only in E_8.

For the time being we let N, N_1, and N_2 denote the numbers of the solutions $\langle x, y\rangle$, $\langle x_1, y_1\rangle$, and $\langle x_2, y_2\rangle$, respectively. We shall prove that

$$n \equiv N \quad \text{or} \quad n \equiv -N \bmod 2y.$$

Since we also have $y \geqslant n$ (by E_1) and $y \geqslant N$ (by the definition of N), it follows that $n = N$, as required. The number N_2 will be the "stepping stone" to get from n to N.

First of all, as before, it follows from E_5 that $A \equiv 1 \bmod 2y$, and then it follows from the definition of N_2 and Lemma 5.4 that $y_2 \equiv N_2 \bmod 2y$. But by E_8 we have $y_2 \equiv n \bmod 2y$; hence

$$N_2 \equiv n \bmod 2y.$$

Next, $A \equiv a \bmod x_1^2$ by E_5, and then $y_2 = y_{N_2}(A) \equiv y_{N_2}(a) \bmod x_1^2$ by Lemma 5.5. Using E_7, we have $y = y_N(a) \equiv y_2 \bmod x_1^2$. Hence

$$y_N(a) \equiv y_{N_2}(a) \bmod x_1^2.$$

We now need two more lemmas, which will be proved in 5.8.

5.6. **Lemma.** *If $y_i(a) \equiv y_j(a) \bmod x_n(a)$, where $a > 1$, then either $i \equiv j$ or $i \equiv -j \bmod 2n$.*

5.7. **Lemma.** *If $y_i(a)^2$ divides $y_j(a)$, then $y_i(a)$ divides j.*

If we apply Lemma 5.6 with N, N_2, and N_1 in place of i, j, and n, and use the last congruence proved, we obtain

$$N \equiv \pm N_2 \bmod 2N_1.$$

If we apply Lemma 5.7 with N and N_1 in place of i and j, and use E_3, we obtain $y | N_1$. Hence

$$N \equiv \pm N_2 \bmod 2y,$$

and since we have already shown that $N_2 \equiv n \bmod 2y$, this completes the proof. □

5.8. Proof of the Lemmas. We shall write x_n and y_n instead of $x_n(a)$ and $y_n(a)$. Using the formula

$$x_{nk} + y_{nk}\sqrt{a^2-1} = \left(x_n + y_n\sqrt{a^2-1}\right)^k,$$

we find that

$$y_{nk} = \sum_{\substack{j \leqslant k \\ j \equiv 1 (\mathrm{mod}\ 2)}} \binom{k}{j} x_n^{k-j} y_n^j (a^2-1)^{(j-1)/2}.$$

In particular,

$$y_{nk} \equiv k x_n^{k-1} y_n \mod (a^2-1),$$

which gives Lemma 5.4 if we set $n = 1$. In addition, we have

$$y_{nk} \equiv k x_n^{k-1} y_n \mod y_n^3.$$

If we replace nk, k, and n by n, n/k, and k, respectively, we obtain

$$y_n \equiv \frac{n}{k} x_k^{n/k-1} y_k \mod y_k^3.$$

Since x_k and y_k are relatively prime, we have

$$y_n \equiv 0 \mod y_k^2 \Rightarrow \frac{n}{k} \equiv 0 \mod y_k \Rightarrow n \equiv 0 \bmod y_k,$$

which gives Lemma 5.7.

If we write $y_n(a)$ as a polynomial in a with integer coefficients whose degree and coefficients depend only on n, we immediately obtain Lemma 5.5. It remains to prove Lemma 5.6.

First of all, the equation

$$x_{n\pm m} + \sqrt{a^2-1}\, y_{n\pm m} = \left(x_n + \sqrt{a^2-1}\, y_n\right)\left(x_m \pm \sqrt{a^2-1}\, y_m\right)$$

gives us

$$\begin{aligned} x_{n\pm m} &= x_n x_m \pm (a^2-1) y_n y_m, \\ y_{n\pm m} &= \pm x_n y_m + x_m y_n. \end{aligned}$$

Hence,

$$\begin{aligned} y_{2n\pm m} = y_{n+(n\pm m)} &\equiv x_{n\pm m} y_n \mod x_n \equiv \pm(a^2-1) y_n^2 y_m \mod x_n \\ &\equiv \mp y_m \mod x_n, \end{aligned}$$

and, similarly,

$$y_{4n\pm m} = y_{2n+(2n\pm m)} \equiv -y_{2n\pm m} \mod x_n \equiv y_{\pm m} \mod x_n.$$

This means that the class $y_k \bmod x_n$ has period $4n$ as a function of k, and within $[1, 4n]$ its behavior is determined by its values on the first quarter-period $[1, n]$:

$$y_{2n\pm m} \equiv \mp y_m, \quad y_{\pm m} \equiv \pm y_m, \quad \text{for } 1 \leqslant m \leqslant n.$$

If $a \geqslant 3$, it is clear that Lemma 5.6 follows from these facts and from the inequality $y_m < \frac{1}{2} x_n$ for $1 \leqslant m \leqslant n$, which, in turn, follows because

$$4y_m^2 < (a^2-1) y_n^2 + 1 = x_n^2.$$

If $a = 2$, then we only have $y_m < \frac{1}{2} x_n$ for $m \leqslant n-1$, but this is still enough to complete the proof of the lemma in this case. □

6 The Graph of the Exponential Is Diophantine

6.1. **Proposition.** *The set*

$$E: \quad m = a^n$$

in the $\langle m, a, n\rangle$-space is Diophantine.

Proof. It suffices to show that $E_0 = E \cap \{a | a > 1\}$ is Diophantine. If $a > 1$, we easily obtain by induction on n that

$$(2a-1)^n \leqslant y_{n+1}(a) \leqslant (2a)^n,$$

in the notation of §5. Hence, for any $N \geqslant 1$ we have

$$a^n \left(1 - \frac{1}{2Na}\right)^n = \frac{(2Na-1)^n}{(2N)^n} \leqslant \frac{y_{n+1}(Na)}{y_{n+1}(N)} \leqslant \frac{(2Na)^n}{(2N-1)^n} = a^n \left(1 - \frac{1}{2N}\right)^{-n}.$$

Thus, if we choose N large enough so that both

$$\left(1 - \frac{1}{2N}\right)^{-n} - 1 < \frac{1}{a^n} \quad \text{and} \quad 1 - \left(1 - \frac{1}{2Na}\right)^n < \frac{1}{a^n},$$

then we obtain $a^n = [y_{n+1}(Na)/y_{n+1}(N)]$ (where the brackets here and below denote the integral part of a number). E_0 is therefore a projection of the set E_1:

$$\begin{gathered} a > 1, \\ 0 \leqslant y_{n+1}(Na) - y_{n+1}(N)m < y_{n+1}(N), \\ N > ?, \end{gathered}$$

where a suitable lower bound for N must be inserted in place of ?, in such a way as to keep the last relation Diophantine. An elementary calculation shows that it suffices to set $N > 4n(y+1)$. The results in §5 then imply that E_1 is Diophantine if we trivially introduce the auxiliary relations

$$y' = y_{n+1}(N) \quad \text{and} \quad y'' = y_{n+1}(Na). \qquad \square$$

7 The Factorial and Binomial Coefficient Graphs Are Diophantine

In this section we carry out the last series of arguments.

7.1. **Proposition.** *The set*

$$E: \quad r = \binom{n}{k}, \quad n \geqslant k,$$

in the $\langle r, k, n\rangle$-space is Diophantine.

Here, by definition, $\binom{n}{k} = n(n-1)\cdots(n-k+1)/k!$. We shall need the following lemma.

7.2. **Lemma.** *If $u > n^k$, then $\binom{n}{k}$ = the remainder when $[(u+1)^n/u^k]$ is divided by u.*

Proof. We have

$$(u+1)^n/u^k = \sum_{i=k+1}^{n} \binom{n}{i} u^{i-k} + \binom{n}{k} + \sum_{i=0}^{k-1} \binom{n}{i} u^{i-k}.$$

The first sum is divisible by u, and the last sum is less than 1 if $u > n^k$. □

7.3. Proof of Proposition 7.1. We introduce the auxiliary variables u and v, and take the relations

$$\begin{aligned} E_1&: \quad u > n^k; \\ E_2&: \quad v = [(u+1)^n/u^k]; \\ E_3&: \quad r \equiv v \mod u; \\ E_4&: \quad r < u; \\ E_5&: \quad n \geqslant k. \end{aligned}$$

Lemma 7.2 immediately implies that $E = \mathrm{pr} \cap_{i=1}^{5} E_i$. E_1 is Diophantine because of Proposition 6.1; E_3, E_4, and E_5 are obviously Diophantine. It also becomes obvious that E_2 is Diophantine if we write E_2 in the form

$$(u+1)^n \leqslant u^k v < (u+1)^n + u^k$$

and again use Proposition 6.1. This completes the proof. □

7.4. **Proposition.** *The set $E : m = k!$ is Diophantine.*

7.5. **Lemma.** *If $k > 0$ and $n > (2k)^{k+1}$, then $k! = \left[n^k \Big/ \binom{n}{k}\right]$. (This is proved by some simple estimates.)*

Proof of Proposition 7.4. We take the auxiliary variable n and the relations

$$\begin{aligned} E_1&: \quad n > (2k)^{k+1}; \\ E_2&: \quad m = \left[n^k \Big/ \binom{n}{k}\right]. \end{aligned}$$

The rest is obvious (using Propositions 6.1 and 7.1). □

7.6. **Proposition.** *The set*

$$E: \quad \frac{x}{k} = \binom{p/q}{k}, \quad p > qk,$$

in the $\langle x, y, p, q, k\rangle$-space is Diophantine.

The proof that follows is a slightly more complicated version of the argument in 7.2 and 7.3.

7.7. **Lemma.** *Let $a > 0$ be an integer such that $a \equiv 0 \mod q^k k!$ and $a > 2^{p-1}p^{k+1}$. Then*

$$\binom{p/q}{k} = a^{-1}\left[a^{2k+1}(1+a^{-2})^{p/q}\right] - a\left[a^{2k-1}(1+a^{-2})^{p/q}\right].$$

This is proved using the binomial Taylor series for $(1 + a^{-2})^{p/q}$. The inequality $a > 2^{p-1}p^{k+1}$ allows us to throw away all the terms in the first sum starting with the $(k + 1)$th and all the terms in the second sum starting with the kth when we take the integral part. The congruence $a \equiv \mod q^k k!$ ensures that the partial sums are integers. □

7.8. Proof of Proposition 7.6. We use the auxiliary variables a, u_1, u_2, and v, and the following relations:

$$\begin{aligned}
&E_1: \ a \equiv 0 \mod q^k k!;\\
&E_2: \ a > 2^{p-1}p^{k+1};\\
&E_3: \ u_1/u_2 = a^{-1}\left[a^{2k+1}(1+a^{-2})^{p/q}\right];\\
&E_4: \ v = a\left[a^{2k-1}(1+a^{-2})^{p/q}\right];\\
&E_5: \ xu_2 = y(u_1 - vu_2).
\end{aligned}$$

It follows from Lemma 7.7 that $E = \operatorname{pr} \cap_{i=1}^{5} E_i$. E_1 and E_2 are immediately seen to be Diophantine from Propositions 6.1 and 7.1. E_3 and E_4 are shown to be Diophantine just as at the end of 7.3, except that this time we must raise the inequalities to the qth power after clearing denominators. E_5 is obviously Diophantine.

This concludes the proof of Theorem 1.2, that enumerable sets coincide with Diophantine sets. □

8 Versal Families

Versal families were defined and first used in Section 5.7 of Chapter V. The purpose of this section is to prove their existence, using the result that enumerable sets are Diophantine (Theorem 1.2).

8.1. **Theorem.** *For any $m \geqslant 0$, versal enumerable families of m-sets and m-functions over the base $\mathbf{Z}^+$ exist and can be effectively constructed.*

PROOF. We divide the proof into several steps. Recall that $\tau^{(2)} : (\mathbf{Z}^+)^2 \overset{\sim}{\Rightarrow} \mathbf{Z}^+$ is the primitive recursive isomorphism constructed in §4 of Chapter V, and $\langle t_1^{(2)}, t_2^{(2)} \rangle$ is its inverse. We shall write t_1 and t_2 for brevity.

(a) *A versal family of polynomials in* $\mathbf{Z}^+[x_1, x_2, x_3, \dots]$. We define polynomials $f[l] \in \mathbf{Z}^+[x_1, x_2, x_3, \dots]$ by recursion on $l \in \mathbf{Z}^+, l \geqslant 4$:

$$\begin{aligned} f[1] = f[2] = f[3] &= 1; \\ f[4k] &= k; \\ f[4k+1] &= x_k; \\ f[4k+2] &= f[t_1(k)] + f[t_2(k)]; \\ f[4k+3] &= f[t_1(k)] f[t_2(k)]. \end{aligned}$$

The definition is correct, since $t_1(k), t_2(k) < 4k+2$. The image of the map $k \mapsto f[k]$ coincides with all of $\mathbf{Z}^+[x_1, x_2, x_3, \dots]$, since it contains $\mathbf{Z}^+$ (in the $4k$-places) and all the x_k (in the $4k+1$-places), and, whenever it contains two polynomials $f[k_1]$ and $f[k_2]$, it contains their sum (in the $4\tau^{(2)}(k_1, k_2)+2$-place) and their product (in the $4\tau^{(2)}(k_1, k_2)+3$-place). (Compare with the numbering of constructible sets by ordinals in Chapter V.)

(b) *Construction of a versal* 1*-family over* $\mathbf{Z}^+$. Let E_k be the projection onto the x_1-coordinate of the 0-level of the polynomial $f[t_1(k)] - f[t_2(k)]$. Since all the elements of $\mathbf{Z}[x_1, x_2, x_3, \dots]$ can be represented as such a difference, it is clear that the family $\{E_k\}$ contains all enumerable sets.

(c) $\{E_k\}$ *is enumerable.* We must show that the total space $E = \{\langle i, j \rangle | i \in E_j\} \subset \mathbf{Z}^+ \times \mathbf{Z}^+$ is enumerable. We write the condition $i \in E_j$ in the form of an $\mathcal{L}_1$-type formula, in which all the quantified variables take values in $\mathbf{Z}^+$. We use the fact that $f[t_1(j)] - f[t_2(j)] \in \mathbf{Z}[x_1, \dots, x_j]$. We have

$$\begin{aligned} \langle i, j \rangle \in E \Leftrightarrow & i \in E_j \Leftrightarrow \exists x_1 \cdots \exists x_j (x_1 = i \wedge f[t_1(j)] = f[t_2(j)]) \\ \Leftrightarrow & \exists t ((\exists x_1 \cdots \exists x_j \, \forall k \leqslant j (f[k] = \mathrm{Gd}(k, t))) \\ & \wedge \mathrm{Gd}(5, t) = i \wedge \mathrm{Gd}(t_1(j), t) = \mathrm{Gd}(t_2(j), t)), \end{aligned}$$

where $\mathrm{Gd}(k, t)$ is Gödel's function (see §4 of Chapter V). Furthermore, by the definition of $f[k]$,

$$\begin{aligned} & \exists x_1 \cdots \exists x_j \forall k \leqslant j (f[k] = \mathrm{Gd}(k, t)) \\ & \quad \Leftrightarrow \forall k \leqslant j ((k \leqslant 3 \wedge \mathrm{Gd}(k, t) = 1) \vee \exists l ((k = 4l \wedge \mathrm{Gd}(k, t) = l) \\ & \qquad \vee (k = 4l+2 \wedge \mathrm{Gd}(k, t) = \mathrm{Gd}(t_1(l), t) + \mathrm{Gd}(t_2(l), t)) \\ & \qquad \vee (k = 4l+3 \wedge \mathrm{Gd}(k, t) = \mathrm{Gd}(t_1(l), t) \mathrm{Gd}(t_2(l), t)))). \end{aligned}$$

Here the part of the formula after $\exists l$ defines a decidable set in $\langle k, t, l \rangle$-space. The quantifier $\exists l$ projects this set onto the $\langle k, t \rangle$-coordinates, thereby taking it to an enumerable set, and the bounded quantifier $\forall k \leqslant j$ preserves enumerability (see §2). Returning to the formula that defines E, we find that the set we have constructed so far must be intersected with two other decidable sets and then projected along the t-axis, so that the result is again enumerable.

(d) *Construction of a versal m-family over* $\mathbf{Z}^+$. The case $m = 0$ is trivial, and the case $m = 1$ has already been discussed. The case $m \geqslant 2$ reduces to the case $m = 1$ using the isomorphism $\tau^{(m)} : (\mathbf{Z}^+)^m \tilde{\Rightarrow} \mathbf{Z}^+$. In fact, let $E_k = E_k^{(1)}$ be a versal 1-family, and set $E_k^{(m)} = (\tau^{(m)})^{-1}(E_k^{(1)})$. The family $\{E_k^{(m)}\}$ is enumerable because

$$E^{(m)} = \{\langle x, k\rangle | x \in E_k^{(m)}\} = \left\{\langle (\tau^{(m)})^{-1}(x), k\rangle | x \in E_k^{(1)}\right\}$$
$$= (\tau^{(m)}, \mathrm{pr}_1^1)^{-1} E^{(1)}.$$

(e) *Construction of a versal family of* 1*-functions.* We take a versal 2-family $\{E_k^{(2)}\}$ with total space

$$E^{(2)} = \{\langle x, y, k\rangle | \langle x, y\rangle \in E_k^{(2)}\} \subset (\mathbf{Z}^+)^3.$$

Let $g(x, y, k, z)$ be a primitive recursive function such that the projection of its 1-level onto the $\langle x, y, k\rangle$-coordinates coincides with $E^{(2)}$. We set

$$f(x, k) = t_1^{(2)}\left(\min\left\{u | g(x, t_1^{(2)}(u), t_2^{(2)}(u)) = 1\right\}\right).$$

We claim that $\{f_k | f_k(x) = f(x, k)\}$ is a versal family of 1-functions. The total function is obviously partial recursive. We need only verify that every partial recursive 1-function f occurs in the family.

Let Γ_f be the graph of f, and let $\Gamma_f = E_{k_0}^{(2)}$, where $k_0 \in \mathbf{Z}^+$. We show that $f = f_{k_0}$. In fact,

$$\langle x, f(x)\rangle \in \Gamma_f = E_{k_0}^{(2)} \Leftrightarrow \langle x, f(x), k_0\rangle \in E^{(2)} \Leftrightarrow \exists z \in \mathbf{Z}^+,$$
$$g(x, f(x), k_0, z) = 1.$$

Among the $z \in \mathbf{Z}^+$ that make $g(x, f(x),\ k_0, z) = 1$, we choose the z for which the number u given by $\langle f(x), z\rangle = \langle t_1^{(2)}(u), t_2^{(2)}(u)\rangle$ is minimal. For this u we have $f_{k_0}(x) = t_1^{(2)}(u) = f(x)$, which proves the claim.

(f) *Construction of a versal family of m-functions.* The case $m = 0$ is trivial. If $\{f_k^{(1)}\}$ is a versal family of 1-functions, then for $m \geqslant 2$ we set

$$f_k^{(m)}(x_1, \ldots, x_m) = f_k^{(1)}(\tau^{(m)}(x_1, \ldots, x_m)),$$

thereby obtaining a versal family of m-functions.

The theorem is proved. □

8.2. The choice of versal families is far from unique. If $m > 1$, there does not exist a versal family that contains each function or each set exactly once (i.e., a universal family). Nevertheless, there are important methods of extracting invariant information from data about the position of a function or set in a versal family. The next section is devoted to this question.

9 Kolmogorov Complexity

9.1. Let $u = \{u_k\}$ be an enumerable family of m-functions over $\mathbf{Z}^+$, and let f be a partial recursive m-function. We define the *complexity of f relative to the family u* as

$$C_u(f) = \begin{cases} \min\{k | u_k = f\}, & \text{if such a } k \text{ exists;} \\ \infty, & \text{otherwise.} \end{cases}$$

We call the enumerable family u (asymptotically) *optimal* if for any other enumerable family v, there exists a constant $c_{u,v} > 0$ such that for every partial recursive m-function f we have

$$C_u(f) \leqslant c_{u,v} C_v(f).$$

If we take v to be any versal family, we see that an optimal family must be versal, i.e., $C_u(f)$ never takes the value ∞.

9.2. **Theorem** (Kolmogorov)

(a) *For any $m \geq 0$, optimal families exist and can be effectively constructed.*
(b) *If u and v are optimal families of m-functions, then for any m-function f,*

$$c_{v,u}^{-1} \leqslant C_u(f)/C_v(f) \leqslant c_{u,v}.$$

9.3. *Remarks*

(a) The measure of complexity $C_u(f)$ involves the following intuitive ideas. In order to define any enumerable family u, it is necessary to give only a finite amount of information, for example, a program that semicomputes the total function of u. Therefore, in order to define a specific function f that occurs in the family u, it suffices to give no more than

$$\log_2 C_u(f) + \text{const}$$

bits of information, namely, the program for u and the number of f in u.

(b) A family being optimal means that it can be used to compute *any* m-function, and that the loss in using it rather than any other family to compute a function is bounded by a constant that does not depend on the function.

(c) Finally, the inequality 9.2(b), which follows trivially from the definition of an optimal family, shows that to within an additive term that is bounded in absolute value, the logarithmic measure of complexity

$$K_u(f) = [\log_2 C_u(f)] + 1 \quad (\text{where "[\ \ ]" = "integral part"})$$

does not depend on the choice of the optimal family u, and so is an asymptotic invariant of f.

9.4. Proof of Theorem 9.2. We first choose a recursive embedding $\theta : \mathbf{Z}^+ \times \mathbf{Z}^+ \to \mathbf{Z}^+$ that has a recursive inverse function and that satisfies the following linear growth condition in one of its arguments:

$$\theta(k, j) \leqslant k \cdot \phi(j), \quad \text{for all } k, j \in \mathbf{Z}^+ \text{ and some suitable } \phi : \mathbf{Z}^+ \to \mathbf{Z}^+.$$

For example, we could let $\theta_1(k, j) = (2k-1)2^j$ with $\phi_1(j) = 2^{j+1}$, or, following Kolmogorov, we could let

$$\theta_2(\overline{k_1 k_2 \cdots k_r}, \overline{j_1 j_2 \cdots j_s}) = \overline{j_1 j_1 \cdots j_s j_s 01 k_1 \cdots k_r},$$

where $k_\alpha, j_\beta \in \{0, 1\}$ and the bar denotes the binary expansion of a number. Here $\phi_2(j) < \text{ const} \cdot j^2$, so that this function grows more slowly. (See also Section 9.8 below.)

Now let U be any versal family of $(m+1)$-functions. We define a family u of m-functions by setting

$$u(x_1, \ldots, x_m, k) = U(x_1, \ldots, x_m, \theta^{-1}(k)).$$

We show that the family u is optimal, with the following bound for the constants:

$$c_{u,v} \leqslant \phi(C_U(v)).$$

In fact, let f be a recursive m-function. It suffices to consider the case in which f occurs in the family v. Then

$$\begin{aligned} f(x_1, \ldots, x_m) &= v(x_1, \ldots, x_m; C_v(f)) \\ &= U(x_1, \ldots, x_m, C_v(f); C_U(v)) \\ &= u(x_1, \ldots, x_m, \theta(C_v(f), C_U(v))), \end{aligned}$$

so that

$$C_u(f) \leqslant \theta(C_v(f), C_U(v)) \leqslant C_v(f)\phi(C_U(v)).$$

The theorem is proved. □

9.5. Example. A 0-function f can be identified with the single value it takes, i.e., with a positive integer n. In this case, Theorem 9.2 gives us an almost invariant complexity $C_u(n)$ for the integers. We have:

(a) $C_u(n) \leqslant \text{const} \cdot n$ for all n, since the function "n" appears in the nth place in the simplest versal family $u_n(\cdot) = n$.

(b) $C(n) \sim \min\{2^{j-1}(2k-1) | n \text{ is the } k\text{th value of the } j\text{th function in some versal family of 1-functions}\}$. (We write $f \sim g$ if f and g have the same domain of definition, and $f \leqslant \text{const} \cdot g$ and $g \leqslant \text{const} \cdot f$ for suitable constants. In relations of the type $C_u(f_k) \sim g(k)$, we often omit the designation of the optimal family u, which we take to be arbitrary, but fixed.)

It is clear from (b) that the complexity of the numbers p_n (the nth prime), n^2, or

$$n^{n^{n^{\cdot^{\cdot^{\cdot^{n}}}}}} \qquad (n \text{ times})$$

as $n \to \infty$ is asymptotically no greater than const $\cdot\, n$, since each of these is the nth value of a fixed recursive function. In 9.7(b) below, we shall lower this estimate to const $\cdot\, C(n)$.

Instead of integers, Kolmogorov and his collaborators considered finite binary sequences and constructed a theory that showed that the most complex binary sequences are those that approach random behavior. See the survey article by A. K. Zvonkin and L. A. Levin in *Uspehi Matem. Nauk*, vol. XXV, No. 6 (1970) (translated in *Russian Mathematical Surveys*), which contains a large bibliography.

9.6. **Proposition.**

(a) *Let*

$$F = f_0(f_1(x_1, \ldots, x_m), \ldots, f_n(x_1, \ldots, x_m), x_{m+1}, \ldots, x_p),$$

where the f_i are recursive functions. Then

$$C(F) \leqslant \text{ const} \cdot \prod_{i=1}^{n} C(f_i) \left(\log \prod_{i=1}^{n} C(f_i) \right)^{n-1}$$

if f_0 is fixed and f_i runs through all possible m-functions. Here const *depends on f_0 and on the families used to compute the complexity, but does not depend on $f_1, \ldots, f_n$.*

(b) *If f_0 is also allowed to vary, then $\prod_{i=1}^{n}$ must be replaced by $\prod_{i=0}^{n}$ and $\log^{n-1}$ must be replaced by $\log^n$ on the right.*

9.7. *Special cases*

(a) If, for example, we set $f_0 = \text{sum}_2$ or prod_2, then we have

$$C(f_1 + f_2), C(f_1 f_2) \leqslant \text{ const } C(f_1)C(f_2)\log(C(f_1)C(f_2)).$$

(b) If we set $n = 1$ and $m = 0$, we find that for any enumerable family $\{f_k\}$,

$$C(f(k, x_1, \ldots, x_p)) \leqslant \text{ const } C(k).$$

9.8. Proof of Proposition 9.6. First of all, for every $n \geqslant 1$ we define the following recursive bijection with a recursive inverse:

$$\theta^{(n)}(k_1, \ldots, k_n) = \begin{cases} \text{the index of the } n\text{-tuple } \langle k_1, \ldots, k_n \rangle \text{ if we order } n\text{-tuples} \\ \quad \text{according to increasing } \prod_{i=1}^{n} k_i, \text{ and in alphabetical order} \\ \quad\quad \text{for fixed } \prod_{i=1}^{n} k_i. \end{cases}$$

It is easy to see (by induction on n) that

$$\theta^{(n)}(k_1,\dots,k_n) \leqslant \text{ const } \prod_{i=1}^{n} k_i \left(\log \prod_{i=1}^{n} k_i\right)^{n-1}.$$

We define the function $\Theta : (\mathbf{Z}^+)^{n+1} \to \mathbf{Z}^+$ as follows:

$$\Theta(l_1,\dots,l_{n+1}) = \theta(\theta^{(n)}(l_1,\dots,l_n), l_{n+1}),$$

where θ is as described in 9.4.

We now consider two optimal families $v(x_1,\dots,x_p,l)$ and $u(x_1,\dots,x_m,k)$ of p-functions and m-functions, respectively. We use these two families to construct the families

$$\begin{aligned} &W(x_1,\dots,x_p;k_1,\dots,k_n,l) \\ &\qquad = v(u(x_1,\dots,x_m,k_1),\dots,u(x_1,\dots,x_m,k_n),x_{m+1},\dots,x_p,l), \\ &\qquad\quad w(x_1,\dots,x_p,k) = W(x_1,\dots,x_p,\Theta^{-1}(k)). \end{aligned}$$

The function F occurs in the

$$\theta\left(\theta^{(n)}(C_u(f_1),\dots,C_u(f_n)), C_v(f_0)\right)$$

place in the family w. Then the estimate $\theta(k,j) \leqslant k \cdot \phi(j)$, along with the estimate for $\theta^{(n)}$, gives assertion (a).

We similarly obtain (b) if we replace Θ by $\theta^{(n+1)}$ in the definition of w. □

Remark. The function $\theta^{(n)}$ gives us the most economical estimate for $C(F)$ that is symmetrical in the $C(f_1),\dots,C(f_n)$. In specific situations it might make sense to improve the estimate in certain of the $C(f_i)$ at the expense of worsening the estimate with respect to the others; this is done by suitably changing θ. For example, Kolmogorov's θ gives

$$C(f_1+f_2) \leqslant \text{ const } C(f_1)C(f_2)^2,$$

which is better than

$$\text{const } C(f_1)C(f_2)\log(C(f_1)C(f_2))$$

if $C(f_2)$ grows much more slowly than $C(f_1)$.

9.9. **Theorem.** *The function $C(f)$ is not computable. More precisely, let $g(k)$ be any unbounded partial recursive function, and let $\{f_k\}$ be any enumerable family. Then it is false that $C(f_k)|_{D(g)} \sim g(k)$.*

Thus, $C(f_k)$ can be computable (even up to $\sim$) only on a set of indices k such that there are only finitely many different functions among the functions f_k; otherwise, $C(f_k)$ is not bounded on this set.

PROOF. Suppose that $C(f_k)|_{D(g)} \sim g(k)$. We show that there exists a general recursive function $h : \mathbf{Z}^+ \to \mathbf{Z}^+$ whose image is contained in $D(g)$ and such that $g \circ h$ is monotonically increasing. We then obtain a contradiction as follows. By 9.7(b), for all k we have

$$C(f_{h(k)}) \leqslant \text{ const } C(k),$$

and, by our assumption and by the fact that $g \circ h$ is increasing,

$$C(f_{h(k)}) \geqslant \text{ const } g(h(k)) \geqslant \text{ const} \cdot k.$$

But these two inequalities are incompatible, because $\liminf C(k)/k = 0$ (for example, $C(k^2)/k^2 \leqslant \text{const}/k$).

It remains to construct h. We choose a general recursive bijection $h_1 : \mathbf{Z}^+ \xrightarrow{\sim} D(g)$, using Proposition 5.6 of Chapter V, and we set

$$E = \{k | \forall i < k, g(h_1(i)) < g(h_1(k))\}.$$

This set is decidable and infinite, and $g \circ h_1$ is an increasing function on E.

Let $h_2 : \mathbf{Z}^+ \to E$ be an increasing general recursive bijection (again using Proposition 5.6 of Chapter V). Then $h = h_1 \circ h_2$ has the necessary properties. The theorem is proved. □

9.10. *Remarks*

(a) Theorem 9.9 shows that computing complexity is a problem demanding creativity: even if we find the number of a place where f occurs in an optimal family $\{u_k\}$, there is no algorithm that could tell us whether this function occurs even sooner.

(b) Since $C(k) \neq C(l) \Rightarrow k \neq l$, it follows that for all x and B,

$$\text{card } \{y | y \leqslant x, C(y) \leqslant x/B\} \leqslant x/B,$$

i.e., most numbers have a large complexity.

Nevertheless, it is not possible to give effectively a sequence of numbers that asymptotically have maximal complexity. More precisely, let $\{k_i\}$ be any increasing sequence with $C(k_i) \geqslant k_i/B$ for some constant B. Then the set $\{k_i\}$ does not contain a single infinite enumerable set E. Otherwise, we would be able to find an increasing general recursive function $h : \mathbf{Z}^+ \to E$, and would obtain a contradiction, as in Theorem 9.9.

(c) Let $u = \{u_k\}$ be any optimal family of m-functions. The "moments of first appearance" $\{k | \forall i < k, u_i \neq u_k\}$ actually form a sequence of asymptotically maximal complexity, since, by the definition and by 9.7(b), they satisfy

$$k = C_u(u_k) \leqslant \text{ const} \cdot C(k).$$

Thus, we might say that in an optimal family the functions first appear "at random moments."

The problem of computing $C(u_k)$ is complicated by the fact that, at least in the specific families in the proof of Theorem 9.2, any function appears infinitely

often, so that if we are not lucky we might first notice the function arbitrarily far out from the place where it first appeared.

(d) Finally, we mention that at least one essential aspect of the complexity of computations has not been touched upon in our discussion of C_u. Namely, $\log_2 C(k)$ measures the *length of a program* that could compute k, but says nothing about the *time* it takes for such a program to work, let alone the possibilities for shortening the time by performing parallel computations, lengthening the program, and so on.

The concept of complexity is rather far removed from practical uses. But it seems to be such a fundamental idea that its role in theoretical mathematics is likely to grow.