ration: Please
ne MONTHLY.
keep one for

ssroom Notes

partment of
ng Associate
ssociation of
PTIONS: to
ngton, D. C.

ER
H, JR.

of January.
er.

HILBERT'S TENTH PROBLEM IS UNSOLVABLE

MARTIN DAVIS, Courant Institute of Mathematical Science

When a long outstanding problem is finally solved, every mathematician would like to share in the pleasure of discovery by following for himself what has been done. But too often he is stymied by the abstruseness of so much of contemporary mathematics. The recent negative solution to Hilbert's tenth problem given by Matiyasevič (cf. [23], [24]) is a happy counterexample. In this article, a complete account of this solution is given; the only knowledge a reader needs to follow the argument is a little number theory: specifically basic information about divisibility of positive integers and linear congruences. (The material in Chapter 1 and the first three sections of Chapter 2 of [25] more than suffices.)

Hilbert's tenth problem is to give a computing algorithm which will tell of a given polynomial Diophantine equation with integer coefficients whether or not it has a solution in integers. Matiyasevič proved that *there is no such algorithm.*

Hilbert's tenth problem is the tenth in the famous list which Hilbert gave in his 1900 address before the International Congress of Mathematicians (cf. [18]). The way in which the problem has been resolved is very much in the spirit of Hilbert's address in which he spoke of the conviction among mathematicians "that every definite mathematical problem must necessarily be susceptible of a precise settlement, either in the form of an actual answer to the question asked, or by *the proof of the impossibility of its solution* ..." (italics added). Concerning such impossibility proofs Hilbert commented:

"Sometimes it happens that we seek the solution under unsatisfied hypotheses or in an inappropriate sense and are therefore unable to reach our goal. Then the task arises of proving the impossibility of solving the problem under the given hypotheses and in the sense required. Such impossibility proofs were already given by the ancients, in showing, e.g., that the hypotenuse of an isosceles right triangle has an irrational ratio to its leg. In modern mathematics the question of the impossibility of certain solutions has played a key role, so that we have acquired the knowledge that such old and difficult problems as to prove the parallel axiom, to square the circle, or to solve equations of the fifth degree in radicals have no solution in the originally intended sense, but nevertheless have been solved in a precise and completely satisfactory way."

Martin Davis received his Princeton Ph. D. under Alonzo Church. He has held positions at Univ. of Illinois, IAS, Univ. of Calif.-Davis, Ohio State Univ., Rensselaer Poly, Yeshiva Univ. and New York Univ., and he spent a leave at Westfield College, London. He has done research in various aspects of the foundations of mathematics, and is the author of *Computability and Unsolvability* (McGraw-Hill, 1958), *The Undecidable* (editor, Raven Press, 1965), *Lectures on Modern Mathematics* (Gordon and Breach, 1967), and *First Course in Functional Analysis* (Gordon and Breach, 1967). *Editor.*

Matiyasevič's negative solution of Hilbert's tenth problem is of just this character. It is not a solution in Hilbert's "originally intended sense" but rather a "precise and completely satisfactory" proof that no such solution is possible. The methods needed to make it possible to prove the non-existence of algorithms had not been developed in 1900. These methods are part of the theory of recursive (or computable) functions, developed by logicians much later ([6] is an exposition of recursive function theory). In this article no previous knowledge of recursive function theory is assumed. The little that is needed is developed in the article itself.

What will be proved in the body of this article is that no algorithm exists for testing a polynomial with integer coefficients to determine whether or not it has *positive integer* solutions (Hilbert inquired about arbitrary integer solutions). But then it will follow at once that there can be no algorithm for integer solutions either. For one could test the equation

$$P(x_1, \cdots, x_n) = 0$$

for possession of positive solutions $\langle x_1, \cdots, x_n \rangle$ by testing

$$P(1 + p_1^2 + q_1^2 + r_1^2 + s_1^2, \cdots, 1 + p_n^2 + q_n^2 + r_n^2 + s_n^2) = 0$$

for possession of integer solutions $\langle p_1, q_1, r_1, s_1, \cdots, p_n, q_n, r_n, s_n \rangle$. This is because (by a well-known theorem of Lagrange) every non-negative integer is the sum of four squares. (Just this once the stated prerequisite is exceeded! Cf. [17], p. 302.) In the body of this article, only positive integers will be dealt with—except when the contrary is explicitly stated.

When Matiyasevič announced his beautiful and ingenious solution in January 1970, it had been known for a decade that the unsolvability of Hilbert's tenth problem would follow if one could construct a Diophantine equation whose solutions were such that one of its components grew roughly exponentially with another of its components. (In §9, this is explained more precisely.) Matiyasevič showed how the Fibonacci numbers could be used to construct such an equation. In this article the historical development of the subject will not be followed; the aim has rather been to give as smooth and straightforward an account of the main results as seems currently feasible. A brief appendix gives the history.

1. Diophantine Sets. In this article the usual problem of Diophantine equations will be inverted. Instead of being given an equation and seeking its solutions, one will begin with the set of "solutions" and seek a corresponding Diophantine equation. More precisely:

DEFINITION. A set S of ordered n-tuples of positive integers is called **Diophantine** if there is a polynomial $P(x_1, \cdots, x_n, y_1, \cdots, y_m)$, where $m \geqq 0$, with integer coefficients such that a given n-tuple $\langle x_1, \cdots, x_n \rangle$ belongs to S if and only if there exist positive integers $y_1, \cdots, y_m$ for which

$$P(x_1, \cdots, x_n, y_1, \cdots, y_m) = 0.$$

Borrowing from logic the symbols "$\exists$" for "there exists" and "$\Leftrightarrow$" for "if and only if", the relation between the set S and the polynomial P can be written succinctly as:

$$\langle x_1, \cdots, x_n \rangle \in S \Leftrightarrow (\exists\, y_1, \cdots, y_m)\, [P(x_1, \cdots, x_n, y_1, \cdots, y_m) = 0],$$

or equivalently:

$$S = \{\langle x_1, \cdots, x_n \rangle \mid (\exists\, y_1, \cdots, y_m)\ [P(x_1, \cdots, x_n, y_1, \cdots, y_m) = 0]\}.$$

Note that P may (and in non-trivial cases always will) have negative coefficients. *The word* "**polynomial**" *should always be so construed in the article* except where the contrary is explicitly stated. Also *all numbers in this article are positive integers* unless the contrary is stated.

The main question which will be discussed (and settled) in this article is:

Which sets are Diophantine? A vague paraphrase of the eventual answer is: *any set which could possibly be Diophantine is Diophantine.* What does the phrase "which could possibly be Diophantine" mean? And how is all this related to Hilbert's tenth problem? These quite reasonable questions will only be answered much later. In the meantime, the task will be developing techniques for showing that various sets are indeed Diophantine.

A few very simple examples:

(i) *the numbers which are not powers of* 2:

$$x \in S \Leftrightarrow (\exists\, y, z)[x = y(2z + 1)],$$

(ii) *the composite numbers*:

$$x \in S \Leftrightarrow (\exists\, y, z)\, [x = (y + 1)(z + 1)],$$

(iii) *the ordering relation on the positive integers; that is the sets* $\{\langle x, y \rangle \mid x < y\}$, $\{\langle x, y \rangle \mid x \leqq y\}$:

$$x < y \Leftrightarrow (\exists\, z)\ (x + z = y),$$

$$x \leqq y \Leftrightarrow (\exists\, z)\ (x + z - 1 = y),$$

(iv) *the divisibility relation; that is* $\{\langle x, y \rangle \mid x \mid y\}$:

$$x \mid y \Leftrightarrow (\exists\, z)\ (xz = y).$$

Examples (i) and (ii) suggest, as other sets to consider, the set of powers of 2 and of primes respectively. As we shall eventually see, these sets are Diophantine; but the proof is not at all easy.

Another example:

(v) *the set* W *of* $\langle x, y, z \rangle$ *for which* $x \mid y$ *and* $x < z$: Here

$$x \mid y \Leftrightarrow (\exists\, u)\ (y = xu) \text{ and } x < z \Leftrightarrow (\exists\, v)\ (z = x + v).$$

Hence,

$$\langle x, y, z\rangle \in W \Leftrightarrow (\exists u, v)\,[(y - xu)^2 + (z - x - v)^2 = 0].$$

Note that the technique just used is perfectly general. So, in defining a Diophantine set one may use a *simultaneous system* $P_1 = 0,\ P_2 = 0, \cdots, P_k = 0$ of polynomial equations since this system can be replaced by the equivalent single equation:

$$P_1^2 + P_2^2 + \cdots + P_k^2 = 0.$$

By a "function" a positive integer valued function of one or more positive integer arguments will always be understood.

DEFINITION. A function f of n arguments is called **Diophantine** if

$$\{\langle x_1, \cdots, x_n, y\rangle \mid y = f(x_1, \cdots, x_n)\}$$

is a Diophantine set, (i.e., f is Diophantine if its "graph" is Diophantine).

Another question that will be answered here is: *which functions are Diophantine*?

An important Diophantine function is associated with the *triangular numbers*, that is numbers of the form:

$$T(n) = 1 + 2 + \cdots + n = \frac{n(n+1)}{2}.$$

Since $T(n)$ is an increasing function, for each positive integer z, there is a unique $n \geqq 0$ such that

$$T(n) < z \leqq T(n+1) = T(n) + n + 1.$$

Hence each z is *uniquely representable* as:

$$z = T(n) + y; \quad y \leqq n + 1,$$

or equivalently, uniquely representable as:

$$z = T(x + y - 2) + y.$$

In this case, one writes $x = L(z)$, $y = R(z)$; also one sets

$$P(x, y) = T(x + y - 2) + y - 1.$$

Note that $L(z)$, $R(z)$ and $P(x, y)$ are Diophantine functions since

$$\begin{aligned} z &= P(x, y) \Leftrightarrow 2z = (x + y - 2)(x + y - 1) + 2y \\ x &= L(z) \Leftrightarrow (\exists y)\,[2z = (x + y - 2)(x + y - 1) + 2y] \\ y &= R(z) \Leftrightarrow (\exists x)\,[2z = (x + y - 2)(x + y - 1) + 2y]. \end{aligned}$$

The function $P(x, y)$ maps the set of ordered pairs of positive integers one-one

onto the set of positive integers. And, for each z, the ordered pair which is mapped into z by $P(x, y)$ is $(L(z), R(z))$. ("P" is for "pair", "L" for "left", and "R" for "right".) Note also that $L(z) \leqq z$, $R(z) \leqq z$. To summarize:

THEOREM 1.1 (Pairing Function Theorem[1]). *There are Diophantine functions $P(x, y)$, $L(z)$, $R(z)$ such that*

(1) *for all x, y, $L(P(x, y)) = x$, $R(P(x, y)) = y$, and*

(2) *for all z, $P(L(z), R(z)) = z$, $L(z) \leqq z$, $R(z) \leqq z$.*

Another useful Diophantine function is related to the Chinese Remainder Theorem, stated below:

DEFINITION. The numbers $m_1, \cdots, m_N$ are called an **admissible sequence of moduli** if $i \neq j$ implies that m_i and m_j are relatively prime.

THEOREM 1.2 (Chinese Remainder Theorem). *Let $a_1, \cdots, a_N$ be any positive integers and let $m_1 \cdots, m_N$ be an admissible sequence of moduli. Then there is an x such that:*

$$\begin{aligned} x &\equiv a_1 \mod m_1 \\ x &\equiv a_2 \mod m_2 \\ &\cdots\cdots \\ x &\equiv a_N \mod m_N. \end{aligned}$$

The Chinese remainder theorem is proved for example in [25], p. 33. (That x can be assumed positive is not ordinarily stated. But since the product of the moduli added to a solution gives another solution, this is obvious.)

Now let the function $S(i, u)$ be defined as follows:

$$S(i, u) = w,$$

where w is the unique positive integer for which:

$$w \equiv L(u) \mod 1 + iR(u)$$

$$w \leqq 1 + i\,R(u).$$

Here w is simply the least *positive* remainder when $L(u)$ is divided by $1 + i\,R(u)$.

THEOREM 1.3 (Sequence Number Theorem). *There is a Diophantine function $S(i, u)$ such that*

(1) *$S(i, u) \leqq u$, and*

(2) *for each sequence $a_1, \cdots, a_N$, there is a number u such that*

$$S(i, u) = a_i \text{ for } 1 \leqq i \leqq N.$$

Proof. The first task is to show that $S(i, u)$ as defined just above, is a Diophantine

function. The claim is that $w = S(i,u)$ if and only if the following system of equations has a solution:

$$2u = (x+y-2)(x+y-1)+2y$$
$$x = w + z(1+iy)$$
$$1+iy = w+v-1.$$

This is because (by the discussion leading to the Pairing Function Theorem), the first equation is equivalent to:

$$x = L(u) \text{ and } y = R(u).$$

Then (using a technique already noted) one needs only sum the squares of the three equations to see that $S(i,u)$ is Diophantine.

Now $S(i,u) \leq L(u) \leq u$. So finally, let $a_1, \cdots, a_N$ be given numbers. Choose y to be some number greater than each of $a_1, \cdots, a_N$ and divisible by each of $1, 2, \cdots, N$. Then the numbers $1+y, 1+2y, \cdots, 1+Ny$ are an admissible sequence of moduli. (For, if $d \mid 1+iy$ and $d \mid 1+jy$, $i<j$, then $d \mid [j(1+iy)-i(1+jy)]$, i.e., $d \mid j-i$ so that $d \leq N$; but this is impossible unless $d=1$ because $d \mid y$.) This being the case, the Chinese Remainder Theorem can be applied to obtain a number x such that

$$x \equiv a_1 \mod 1+y$$
$$x = a_2 \mod 1+2y$$
$$\cdots\cdots\cdots$$
$$x = a_N \mod 1+Ny.$$

Let $u = P(x,y)$, so that $x = L(u)$ and $y = R(u)$. Then, for $i = 1, 2, \cdots, N$

$$a_i \equiv L(u) \mod 1+iR(u)$$

and $a_i < y = R(u) < 1+iR(u)$. But then by definition, $a_i = S(i,u)$.

A striking characterization of Diophantine sets of positive integers (cf. [26]) is given by:

THEOREM 1.4. *A set S of positive integers is Diophantine if and only if there is a polynomial P such that S is precisely the set of* **positive** *integers in the range of P.*

Proof. If S is related to $P(x_1, \cdots, x_m)$ as in the theorem then

$$x \in S \Leftrightarrow (\exists x_1, \cdots, x_m)\,[x = P(x_1, \cdots, x_m)].$$

Conversely, let

$$x \in S \Leftrightarrow (\exists x_1, \cdots, x_m)\,[Q(x, x_1, \cdots, x_m) = 0].$$

Let $P(x, x_1, \cdots, x_m) = x[1 - Q^2(x, x_1, \cdots, x_m)]$. Then, if $x \in S$, choose $x_1, \cdots, x_m$ such

that $Q(x, x_1, \cdots, x_m) = 0$. Then $P(x, x_1, \cdots, x_m) = x$; so x is in the range of P. On the other hand, if $z = P(x, x_1, \cdots, x_m)$, $z > 0$, then $Q(x, x_1, \cdots, x_m)$ must vanish (otherwise $1 - Q^2 \leqq 0$) so that $z = x$ and $x \in S$.

2. Twenty-four easy lemmas. The first major task is to prove that the exponential function $h(n, k) = n^k$ is Diophantine. This is the hardest thing we shall have to do. The proof is in §3. In this section we develop the methods we shall need, using the so-called Pell equation:

$$\left.\begin{aligned} &x^2 - dy^2 = 1, \quad x, y \geqq 0, \\ &d = a^2 - 1, \qquad a > 1. \end{aligned}\right\} \qquad (*)$$

where

Although this is a famous equation with a considerable literature,[2] a self-contained treatment is given. Note the obvious solutions to (*):

$$x = 1 \qquad y = 0$$

$$x = a \qquad y = 1.$$

LEMMA 2.1. *There are no integers x, y, positive, negative, or zero, which satisfy (*) for which $1 < x + y\sqrt{d} < a + \sqrt{d}$.*

Proof. Let x, y satisfy (*). Since

$$1 = (a + \sqrt{d})(a - \sqrt{d}) = (x + y\sqrt{d})(x - y\sqrt{d}),$$

the inequality implies (taking negative reciprocals) $-1 < -x + y\sqrt{d} < -a + \sqrt{d}$. Adding the inequalities: $0 < 2y\sqrt{d} < 2\sqrt{d}$, i.e., $0 < y < 1$, a contradiction.

LEMMA 2.2. *Let x, y and x', y' be integers, positive, negative, or zero which satisfy (*). Let*

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

Then, x'', y'' satisfies (*).

Proof. Taking conjugates: $x'' - y''\sqrt{d} = (x - y\sqrt{d})(x' - y'\sqrt{d})$. Multiplying gives:

$$(x'')^2 - d(y'')^2 = (x^2 - dy^2)((x')^2 - d(y')^2) = 1.$$

DEFINITION. $x_n(a)$, $y_n(a)$ are defined for $n \geqq 0$, $a > 1$, by setting

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

Where the context permits, the dependence on a is not explicitly shown, writing x_n, y_n.

LEMMA 2.3. *x_n, y_n satisfy (*).*

Proof. This follows at once by induction using Lemma 2.2.

LEMMA 2.4. *Let x, y be a non-negative solution of (*). Then for some n, $x = x_n$, $y = y_n$.*

Proof. To begin with $x + y\sqrt{d} \geqq 1$. On the other hand the sequence $(a + \sqrt{d})^n$ increases to infinity. Hence for some $n \geqq 0$,

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

If there is equality, the result is proved; so suppose otherwise:

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

Since $(x_n + y_n\sqrt{d})\,(x_n - y_n\sqrt{d}) = 1$, the number $x_n - y_n\sqrt{d}$ is positive. Hence, $1 < (x + y\sqrt{d})\,(x_n - y_n\sqrt{d}) < a + \sqrt{d}$. But this contradicts Lemmas 2.1 and 2.2.

The defining relation:

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

is a formal analogue of the familiar formula:

$$(\cos u) + (\sin u)\sqrt{-1} = e^{iu} = (\cos 1 + (\sin 1)\sqrt{-1})^u,$$

with x_n playing the role of cos, y_n playing the role of sin and d playing the role of -1. Thus, the familiar trigonometric identities have analogues in which -1 is replaced by d at appropriate places. For example the Pell equation itself

$$x_n^2 - dy_n^2 = 1$$

is just the analogue of the Pythagorean identity. Next analogues of the familiar addition formulas are obtained.

LEMMA 2.5. *$x_{m\pm n} = x_m x_n \pm d y_n y_m$ and $y_{m\pm n} = x_n y_m \pm x_m y_n$.*

Proof.

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

Hence,

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

Similarly, $(x_{m-n} + y_{m-n}\sqrt{d})\,(x_n + y_n\sqrt{d}) = x_m + y_m\sqrt{d}$. So

$$x_{m-n} + y_{m-n}\sqrt{d} = (x_m + y_m\sqrt{d})\,(x_n - y_n\sqrt{d}),$$

and one proceeds as above.

LEMMA 2.6. $y_{m\pm1} = a\,y_m \pm x_m$, *and* $x_{m\pm1} = ax_m \pm dy_m$.

Proof. Take $n = 1$ in Lemma 2.5.

The familiar notation (x, y) is used to symbolize the g.c.d. of x and y.

LEMMA 2.7. $(x_n, y_n) = 1$.

Proof. If $d \mid x_n$ and $d \mid y_n$, then $d \mid x_n^2 - dy_n^2$, i.e., $d \mid 1$.

LEMMA 2.8. $y_n \mid y_{nk}$.

Proof. This is obvious when $k = 1$. Proceeding by induction, using the addition formula (Lemma 2.5),

$$y_{n(m+1)} = x_n y_{nm} + x_{nm} y_n.$$

By the induction hypothesis $y_n \mid y_{nm}$. Hence, $y_n \mid y_{n(m+1)}$.

LEMMA 2.9. $y_n \mid y_t$ *if and only if* $n \mid t$.

Proof. Lemma 2.8 gives the implication in one direction. For the converse suppose $y_n \mid y_t$ but $n \nmid t$. So one can write $t = nq + r$, $0 < r < n$. Then,

$$y_t = x_r y_{nq} + x_{nq} y_r.$$

Since (by Lemma 2.8) $y_n \mid y_{nq}$, it follows that $y_n \mid x_{nq} y_r$. But $(y_n, x_{nq}) = 1$. (If $d \mid y_n$, $d \mid x_{nq}$, then by Lemma 2.8 $d \mid y_{nq}$ which, by Lemma 2.7, implies $d = 1$.) Hence $y_n \mid y_r$. But, since $r < n$, we have $y_r < y_n$ (e.g., by Lemma 2.6). This is a contradiction.

LEMMA 2.10. $y_{nk} \equiv k\,x_n^{k-1} y_n \bmod (y_n)^3$.

Proof.

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

So,

$$y_{nk} = \sum_{\substack{j=1 \\ j \text{ odd}}}^{k} \binom{k}{j} x_n^{k-j} y_n^j d^{(j-1)/2}.$$

But all terms of this expansion for which $j > 1$ are $\equiv 0 \bmod (y_n)^3$.

LEMMA 2.11. $y_n^2 \mid y_{ny_n}$.

Proof. Set $k = y_n$ in Lemma 2.10.

LEMMA 2.12. *If* $y_n^2 \mid y_t$, *then* $y_n \mid t$.

Proof. By Lemma 2.9, $n \mid t$. Set $t = nk$. Using Lemma 2.10, $y_n^2 \mid k\, x_n^{k-1} y_n$, i.e., $y_n \mid k x_n^{k-1}$. But by Lemma 2.7, $(y_n, x_n) = 1$. So, $y_n \mid k$ and hence $y_n \mid t$.

LEMMA 2.13. *$x_{n+1} = 2ax_n - x_{n-1}$ and $y_{n+1} = 2ay_n - y_{n-1}$.*

Proof. By Lemma 2.6,

$$x_{n+1} = ax_n + dy_n, \qquad y_{n+1} = ay_n + x_n,$$
$$x_{n-1} = ax_n - dy_n, \qquad y_{n-1} = ay_n - x_n.$$

So, $x_{n+1} + x_{n-1} = 2ax_n$, $y_{n+1} + y_{n-1} = 2ay_n$.

These second order difference equations, together with the initial values $x_0 = 1$, $x_1 = a$, $y_0 = 0$, $y_1 = 1$, determine the values of all the x_n, y_n. Various properties of these sequences are easily established by checking them for $n = 0, 1$ and using these difference equations to show that the property for $n + 1$ can be inferred from its holding for n and $n - 1$. Some simple (but important) examples follow:

LEMMA 2.14. $y_n \equiv n \bmod a - 1$.

Proof. For $n = 0, 1$ equality holds. Proceeding inductively, using $a \equiv 1$, mod $a-1$:

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

LEMMA 2.15. *If $a \equiv b \bmod c$, then for all n,*

$$x_n(a) \equiv x_n(b), \quad y_n(a) \equiv y_n(b) \mod c.$$

Proof. Again for $n = 0, 1$ the congruence is an equality. Proceeding by induction:

$$\begin{aligned} y_{n+1}(a) &= 2ay_n(a) - y_{n-1}(a) \\ &\equiv 2by_n(b) - y_{n-1}(b) \mod c \\ &= y_{n+1}(b). \end{aligned}$$

LEMMA 2.16. *When n is even y_n is even and when n is odd y_n is odd.*

Proof. $y_{n+1} = 2ay_n - y_{n-1} \equiv y_{n-1} \bmod 2$. So when n is even, $y_n \equiv y_0 = 0 \mod 2$, and when n is odd, $y_n \equiv y_1 = 1 \mod 2$.

LEMMA 2.17. $x_n(a) - y_n(a)(a - y) \equiv y^n \bmod 2ay - y^2 - 1$.

Proof. $x_0 - y_0(a - y) = 1$ and $x_1 - y_1(a - y) = y$, so the result holds for $n = 0$ and 1. Using Lemma 2.13 and proceeding by induction:

$$\begin{aligned} x_{n+1} - y_{n+1}(a - y) &= 2a[x_n - y_n(a - y)] - [x_{n-1} - y_{n-1}(a - y)] \\ &\equiv 2ay^n - y^{n-1} \end{aligned}$$

$$= y^{n-1}(2ay - 1)$$
$$\equiv y^{n-1} y^2$$
$$= y^{n+1}.$$

LEMMA 2.18. *For all* n, $y_{n+1} > y_n \geqq n$.

Proof. By Lemma 2.6, $y_{n+1} > y_n$. Since $y_0 = 0 \geqq 0$, it follows by induction that $y_n \geqq n$ for all n.

LEMMA 2.19. *For all* n, $x_{n+1}(a) > x_n(a) \geqq a^n$; $\quad x_n(a) \leqq (2a)^n$.

Proof. By Lemmas 2.6 and 2.13 $a\, x_n(a) \leqq x_{n+1}(a) \leqq (2a)x_n(a)$. The result follows by induction.

Next some periodicity properties of the sequence x_k are obtained.

LEMMA 2.20. $x_{2n \pm j} \equiv -x_j \bmod x_n$.

Proof. By the addition formulas (Lemma 2.5)

$$\begin{aligned} x_{2n \pm j} &= x_n x_{n \pm j} + d y_n y_{n \pm j} \\ &\equiv d y_n (y_n x_j \pm x_n y_j) \mod x_n \\ &\equiv d y_n^2 x_j \mod x_n \\ &= (x_n^2 - 1) x_j \\ &\equiv -x_j \mod x_n. \end{aligned}$$

LEMMA 2.21. $x_{4n \pm j} \equiv x_j \mod x_n$.

Proof. By Lemma 2.20

$$x_{4n \pm j} \equiv -x_{2n \pm j} \equiv x_j \mod x_n.$$

LEMMA 2.22. *Let* $x_i \equiv x_j \bmod x_n$, $i \leqq j \leqq 2n$, $n > 0$. *Then* $i = j$, *unless* $a = 2$, $n = 1$, $i = 0$ *and* $j = 2$.

Proof. First suppose x_n is odd and let $q = (x_n - 1)/2$. Then the numbers $-q, -q+1, -q+2, \cdots, -1, 0, 1, \cdots, q-1, q$ are a complete set of mutually incongruent residues modulo x_n. Now by Lemma 2.19,

$$1 = x_0 < x_1 < \cdots < x_{n-1}.$$

Using Lemma 2.6, $x_{n-1} \leqq x_n/a \leqq \frac{1}{2} x_n$; so $x_{n-1} \leqq q$. Also by Lemma 2.20, the numbers

$$x_{n+1}, x_{n+2}, \cdots, x_{2n-1}, x_{2n}$$

are congruent modulo x_n respectively to:

$$-x_{n-1}, -x_{n-2}, \cdots, -x_1, -x_0 = -1.$$

Thus the numbers $x_0, x_1, x_2, \cdots, x_{2n}$ are mutually incongruent modulo x_n. This gives the result.

Next suppose x_n is even and let $q = x_n/2$. In this case, it is the numbers

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

which are a complete set of mutually incongruent residues modulo x_n. (For, $-q \equiv q \mod x_n$.) As above, $x_{n-1} \leqq q$. So the result will follow as above, unless $x_{n-1} = q = x_n/2$, so that $x_{n+1} \equiv -q \mod x_n$, in which case $i = n-1$, $j = n+1$ would contradict our result. But, by Lemma 2.6,

$$x_n = ax_{n-1} + dy_{n-1},$$

so that $x_n = 2x_{n-1}$ implies $a = 2$ and $y_{n-1} = 0$, i.e., $n = 1$. So the result can fail only for $a = 2$, $n = 1$ and $i = 0$, $j = 2$.

LEMMA 2.23. *Let* $x_j \equiv x_i \mod x_n$, $n > 0$, $0 < i \leqq n$, $0 \leqq j < 4n$, *then either* $j = i$ *or* $j = 4n - i$.

Proof. First suppose $j \leqq 2n$. Then by Lemma 2.22, $j = i$ unless the exceptional case occurs. Since $i > 0$, this can only happen if $j = 0$. But then

$$i = 2 > 1 = n.$$

Otherwise, let $j > 2n$ and set $\bar{j} = 4n - j$ so $0 < \bar{j} < 2n$. By Lemma 2.21, $x_j \equiv x_{\bar{j}} \equiv x_i \mod x_n$. Again $\bar{j} = i$ unless the exceptional case of Lemma 2.22 occurs. But this last is out of the question because $i, \bar{j} > 0$.

LEMMA 2.24. *If* $0 < i \leqq n$ *and* $x_j \equiv x_i \mod x_n$, *then* $j \equiv \pm i \mod 4n$.

Proof. Write $j = 4nq + \bar{j}$, $0 \leqq \bar{j} < 4n$. By Lemma 2.21,

$$x_i \equiv x_{\bar{j}} \equiv x_j \mod x_n.$$

By Lemma 2.23 $i = \bar{j}$ or $i = 4n - \bar{j}$. So, $j \equiv \bar{j} \equiv \pm i \mod 4n$.

3. The exponential function. Consider the system of Diophantine equations:

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

(II) $u^2 - (a^2-1)v^2 = 1$

(III) $s^2 - (b^2-1)t^2 = 1$

(IV) $v = ry^2$

(V) $b = 1 + 4py = a + qu$

(VI) $s = x + cu$

(VII) $t = k + 4(d-1)y$

(VIII) $$y = k + e - 1.$$

Then it is possible to prove:

THEOREM 3.1. *For given a, x, k, $a > 1$, the system* I-VIII *has a solution in the remaining arguments $y, u, v, s, t, b, r, p, q, c, d, e$ if and only if $x = x_k(a)$.*

Proof. First let there be given a solution of I-VIII. By V, $b > a > 1$. Then I, II, III imply (by Lemma 2.4) that there are $i, j, n > 0$ such that

$$x = x_i(a),\ y = y_i(a),\ u = x_n(a),\ v = y_n(a),\ s = x_j(b),\ t = y_j(b).$$

By I V, $y \leqq v$ so that $i \leqq n$. V and VI yield the congruences

$$b \equiv a \mod x_n(a);\quad x_j(b) \equiv x_i(a) \mod x_n(a)$$

and by Lemma 2.15 one gets also

$$x_j(b) \equiv x_j(a) \mod x_n(a).$$

Thus,

$$x_i(a) \equiv x_j(a) \mod x_n(a).$$

By Lemma 2.24,

(1) $$j = \pm i \mod 4n.$$

Next, equation IV yields

$$(y_i(a))^2 \mid y_n(a).$$

so that by Lemma 2.12,

$$y_i(a) \mid n$$

and (1) yields:

(2) $$j = \pm i \mod 4y_i(a).$$

By equation V

$$b \equiv 1 \mod 4y_i(a),$$

so by Lemma 2.14,

(3) $$y_j(b) \equiv j \mod 4y_i(a).$$

By equation VII,

(4) $$y_j(b) \equiv k \mod 4y_i(a).$$

Combining (2), (3), (4),

(5) $$k \equiv \pm i \mod 4y_i(a).$$

Equation VIII yields

$$k \leqq y_i(a)$$

and by Lemma 2.18,

$$i \leqq y_i(a).$$

Since the numbers

$$-2y+1,\ -2y+2,\cdots,-1,0,1,\cdots,2y$$

form a complete set of mutually incongruent residues modulo $4y = 4y_i(a)$, these inequalities show that (5) implies $k = i$. Hence

$$x = x_i(a) = x_k(a).$$

Conversely, let $x = x_k(a)$. Set $y = y_k(a)$ so that I holds. Let $m = 2ky_k(a)$ and let $u = x_m(a)$, $v = y_m(a)$. Then II is satisfied. By Lemmas 2.9 and 2.11 $y^2 \mid v$. Hence one can choose r satisfying IV. Moreover by Lemma 2.16, v is even so that u is odd. By Lemma 2.7, $(u,v) = 1$. Hence $(u, v\ 4y) = 1$. (If p is a prime divisor of u and of $4y$, then $p \mid y$ because u is odd, and hence $p \mid v$ since $y \mid v$.) So by the Chinese Remainder Theorem (Theorem 1.2), one can find b_0 such that

$$b_0 \equiv 1 \mod 4y$$

$$b_0 \equiv a \mod u.$$

Since $b_0 + 4juy$ will also satisfy these congruences, b, p, q satisfying V can be found. III is satisfied by setting $s = x_k(b)$, $t = y_k(b)$. Since $b > a$, $s = x_k(b) > x_k(a) = x$. By Lemma 2.15 (using V), $s \equiv x \mod u$. So c can be chosen to satisfy VI. By Lemma 2.18, $t \geqq k$ and by Lemma 2.14, $t \equiv k \mod b - 1$ and hence using V, $t \equiv k \mod 4y$. So d can be chosen to satisfy VII. By Lemma 2.18 again, $y \geqq k$, so VIII can be satisfied by setting $e = y - k + 1$.

COROLLARY 3.2. *The function*

$$g(z,k) = x_k(z+1)$$

is Diophantine.

Proof. Adjoin to the system I-VIII:

(A) $$a = z + 1.$$

By the theorem, the system (A), I-VIII has a solution if and only if $x = x_k(a) = g(z,k)$. Thus a Diophantine definition of g can be obtained in the usual way by summing the squares of 9 polynomials.

Now at last it is possible to prove:

THEOREM 3.3. *The exponential function $h(n,k) = n^k$ is Diophantine.*

First, a simple inequality:

LEMMA 3.4. *If* $a > y^k$, *then* $2ay - y^2 - 1 > y^k$.

Proof. Set $g(y) = 2ay - y^2 - 1$. Then (since $a \geqq 2$) $g(1) = 2a - 2 \geqq a$. For $1 \leqq y < a$, $g'(y) = 2a - 2y > 0$. So $g(y) \geqq a$ for $1 \leqq y < a$. Then for $a > y^k \geqq y$, $2ay - y^2 - 1 \geqq a > y^k$.

Now, adjoin to equations I-VIII:

IX $\quad (x - y(a - n) - m)^2 = (f - 1)^2(2an - n^2 - 1)^2$

X $\quad m + g = 2an - n^2 - 1$

XI $\quad w = n + h = k + l$

XII $\quad a^2 - (w^2 - 1)(w - 1)^2 z^2 = 1.$

Theorem 3.3 then follows at once from:

LEMMA 3.5. $m = n^k$ *if and only if equations* I-XII *have a solution in the remaining arguments.*

Proof. Suppose I-XII hold. By XI, $w > 1$. Hence $(w - 1)z > 0$ and so by XII $a > 1$. So Theorem 3.1 applies and it follows that $x = x_k(a)$, $y = y_k(a)$. By IX and Lemma 2.17,

$$m \equiv n^k \mod 2an - n^2 - 1.$$

XI yields

$$k, n < w.$$

By XII (using Lemma 2.4), for some j, $a = x_j(w)$, $(w - 1)z = y_j(w)$. By Lemma 2.14,

$$j \equiv 0 \mod w - 1$$

so that $j \geqq w - 1$. So by Lemma 2.19,

$$a \geqq w^{w-1} > n^k.$$

Now by X, $m < 2an - n^2 - 1$, and by Lemma 3.4,

$$n^k < 2an - n^2 - 1.$$

Since m and n^k are congruent and both less than the modulus, they must be equal.

Conversely, suppose that $m = n^k$. Solutions must be found for I-XII. Choose any number w such that $w > n$ and $w > k$. Set $a = x_{w-1}(w)$ so that $a > 1$. By Lemma 2.14,

$$y_{w-1}(w) \equiv 0 \mod w - 1.$$

So one can write

$$y_{w-1}(w) = z(w-1);$$

thus XII is satisfied. XI can be satisfied by setting

$$h = w - n, \quad l = w - k.$$

As before, $a > n^k$ so that again by Lemma 3.4,

$$m = n^k < 2an - n^2 - 1$$

and X can be satisfied. Setting $x = x_k(a)$, $y = y_k(a)$, Lemma 2.17 permits one to define f such that

$$x - y(a-n) - m = \pm (f-1)(2an - n^2 - 1),$$

so that IX is satisfied. Finally, I-VIII can be satisfied by Theorem 3.1.

4. The language of Diophantine predicates. Now that it has been proved that the *exponential function is* Diophantine, many other functions and sets can be handled. As an example, let

$$h(u,v,w) = u^{v^w}.$$

The claim is that h is a Diophantine function. For:

$$y = u^{v^w} \Leftrightarrow (\exists z)\,(y = u^z \,\&\, z = v^w),$$

where "&" is the logician's symbol for "and". Using Theorem 3.3, there is a polynomial P such that:

$$y = u^z \Leftrightarrow (\exists r_1, \cdots, r_n)\,[P(y,u,z,r_1,\cdots,r_n) = 0],$$
$$z = v^w \Leftrightarrow (\exists s_1, \cdots, s_n)\,[P(z,v,w,s_1,\cdots,s_n) = 0].$$

Then,

$$y = u^{v^w} \Leftrightarrow (\exists z, r_1, \cdots, r_n, s_1, \cdots, s_n)\,[P^2(y,u,z,r_1,\cdots,r_n) + P^2(z,v,w,s_1,\cdots,s_n) = 0].$$

Now this procedure is perfectly general: Expressions which are already known to yield Diophantine sets may be combined freely using the logical operations of "&" and "$(\exists)$"; the resulting expression will again define a Diophantine set. (Such expressions are sometimes called *Diophantine predicates*.) In this "language" it is also permissible to use the logician's "$\vee$" for "or", since:

$$(\exists r_1, \cdots, r_n)\,[P_1 = 0] \vee (\exists s_1, \cdots, s_m)\,[P_2 = 0]$$
$$\Leftrightarrow (\exists r_1, \cdots, r_n, s_1, \cdots, s_m)\,[P_1 P_2 = 0].$$

Three important Diophantine functions are given by:

THEOREM 4.1. *The following functions are Diophantine*:

(1) $$f(n,k) = \binom{n}{k}$$

(2) $$g(n) = n!$$

(3) $$h(a,b,y) = \prod_{k=1}^{y} (a+bk).$$

In proving this theorem the familiar notation $[\alpha]$, where α is a real number, will be used to mean the unique integer such that

$$[\alpha] \leqq \alpha < [\alpha] + 1.$$

LEMMA 4.1. *For* $0 < k \leqq n$, $u > 2^n$

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

Proof.

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

where

$$S = \sum_{i=k}^{n} \binom{n}{i} u^{i-k} \qquad R = \sum_{i=0}^{k-1} \binom{n}{i} u^{i-k}.$$

Then S is an integer and

$$\begin{aligned} R &< u^{-1} \sum_{i=0}^{k-1} \binom{n}{i} \\ &< u^{-1} \sum_{i=0}^{n} \binom{n}{i} \\ &= u^{-1}(1+1)^n \\ &< 1. \end{aligned}$$

So,

$$S \leqq (u+1)^n/u^k < S + 1$$

which gives the result.

LEMMA 4.2. *For* $0 < k \leqq n$, $u > 2^n$,

$$[(u+1)^n/u^k] \equiv \binom{n}{k} \mod u.$$

Proof. In Lemma 4.1 all terms of the sum for which $i > k$ are divisible by u.

LEMMA 4.3. $f(n,k) = \binom{n}{k}$ *is Diophantine.*

Proof. Since

$$\binom{n}{k} \leqq \sum_{i=0}^{n} \binom{n}{i} = 2^n < u,$$

Lemma 4.2 determines $\binom{n}{k}$ as the unique positive integer congruent to $[(u+1)^n/u^k]$ modulo u and $< u$. Thus,

$$z = \binom{n}{k} \Leftrightarrow (\exists u, v, w)\,(v = 2^n \,\&\, u > v$$

$$\&\, w = [(u+1)^n/u^k] \,\&\, z \equiv w \mod u \,\&\, z < u).$$

To see that $\binom{n}{k}$ is Diophantine, it then suffices to note that each of the above expressions separated by "&" are Diophantine predicates; $v = 2^n$ is of course Diophantine by Theorem 3. The inequality $u > v$ is of course Diophantine since $u > v \Leftrightarrow (\exists x)(u = v + x)$. Also,

$$z \equiv w \mod u \;\&\; z < u \Leftrightarrow (\exists x, y)\,(w = z + (x-1)u \,\&\, u = z + y).$$

Finally

$$w = [(u+1)^n/u^k]$$

$$\Leftrightarrow$$

$$(\exists x, y, t)\,(t = u + 1 \;\&\, x = t^n \,\&\, y = u^k \,\&\, w \leqq x/y < w + 1),$$

and $w \leqq x/y < w + 1 \Leftrightarrow wy \leqq x < (w+1)y$.

LEMMA 4.4. *If* $r > (2x)^{x+1}$ *then*

$$x! = \left[r^x \Big/ \binom{r}{x}\right].$$

Proof. Let $r > (2x)^{x+1}$. Then,

$$r^x \Big/ \binom{r}{x} = \frac{r^x x!}{r(r-1)\cdots(r-x+1)}$$

$$= x! \left\{ \frac{1}{\left(1 - \frac{1}{r}\right)\cdots\left(1 - \frac{x-1}{r}\right)} \right\}$$

$$< x! \cdot \frac{1}{\left(1 - \frac{x}{r}\right)^x}.$$

Now,

$$\begin{aligned}\frac{1}{1 - \frac{x}{r}} &= 1 + \frac{x}{r} + \left(\frac{x}{r}\right)^2 + \cdots \\ &= 1 + \frac{x}{r}\left\{1 + \frac{x}{r} + \left(\frac{x}{r}\right)^2 + \cdots\right\} \\ &< 1 + \frac{x}{r}\{1 + \tfrac{1}{2} + \tfrac{1}{4} + \cdots\} \\ &= 1 + \frac{2x}{r}.\end{aligned}$$

And,

$$\begin{aligned}\left(1 + \frac{2x}{r}\right)^x &= \sum_{j=0}^{x} \binom{x}{j}\left(\frac{2x}{r}\right)^j \\ &< 1 + \frac{2x}{r}\sum_{j=1}^{x}\binom{x}{j} \\ &< 1 + \frac{2x}{r} \cdot 2^x.\end{aligned}$$

So,

$$\begin{aligned}r^x / \binom{r}{x} &< x! + \frac{2x}{r} \cdot x!\, 2^x \\ &< x! + \frac{2^{x+1} x^{x+1}}{r} \\ &< x! + 1.\end{aligned}$$

LEMMA 4.5. *$n!$ is a Diophantine function.*

Proof. $m = n! \Leftrightarrow$

$$(\exists r, s, t, u, v)\{s = 2x + 1 \,\&\, t = x + 1 \,\&\, r = s^t$$

$$\&\, u = r^n \,\&\, v = \binom{r}{n} \,\&\, mv \leqq u < (m+1)v\}.$$

LEMMA 4.6. *Let $bq \equiv a \bmod M$. Then,*

$$\prod_{k=1}^{y}(a + bk) \equiv b^y y!\binom{q+y}{y} \mod M.$$

Proof.

$$b^y y! \binom{q+y}{y} = b^y (q+y)(q+y-1)\cdots(q+1)$$

$$= (bq+yb)(bq+(y-1)b)\cdots(bq+b)$$

$$\equiv (a+yb)(a+(y-1)b)\cdots(a+b) \pmod{M}.$$

LEMMA 4.7. *$h(a,b,y) = \prod_{k=1}^{y}(a+bk)$ is a Diophantine function.*

Proof. In Lemma 4.6 choose $M = b(a+by)^y + 1$. Then, $(M,b) = 1$ and $M > \prod_{k=1}^{y}(a+bk)$. Hence the congruence $bq \equiv a \bmod M$ is solvable for q and then $\prod_{k=1}^{y}(a+bk)$ is determined as the unique number which is congruent modulo M to $b^y y! \binom{q+y}{y}$ and is also $< M$. I.e.,

$$z = \prod_{k=1}^{y}(a+bk) \leftrightarrow (\exists M,p,q,r,s,t,u,v,w,x)$$

$$\Big\{ r = a+by \,\&\, s = r^y \,\&\, M = bs+1$$

$$\&\, bq = a + Mt \,\&\, u = b^y \,\&\, v = y! \,\&\, z < M$$

$$\&\, w = q+y \,\&\, x = \binom{w}{y} \,\&\, z + Mp = uvx \Big\}.$$

Using the previous expressions for the exponential function, for $v = y!$ and for $x = \binom{w}{y}$, we obtain the result.

The assertion of Theorem 4.1 is contained in Lemmas 4.3, 4.5, and 4.7.

5. Bounded quantifiers. The language of Diophantine predicates permits use of &, $\vee$, and $\exists$. Other operations used by logicians are:

$\sim$ for "not"

$(\forall x)$ for "for all x"

$\rightarrow$ for "if $\cdots$, then $\cdots$"

However, as will be clear later, the use of any of these other operations can lead to expressions which define sets that are not Diophantine. There are also the *bounded existential quantifiers*:

"$(\exists y)_{\leq x}\cdots$" which means "$(\exists y)(y \leq x \,\&\, \cdots)$"

and the *bounded universal quantifiers*:

"$(\forall y)_{\leq x}\cdots$" which means "$(\forall y)(y > x \vee \cdots)$".

It turns out that these operations may be adjoined to the language of Diophantine predicates; that is, the sets defined by expressions of this extended language will still be Diophantine. I.e.,

THEOREM 5.1. *If P is a polynomial,*

$$R = \{\langle y, x_1, \cdots, x_n\rangle \mid (\exists z)_{\leq y}(\exists y_1, \cdots, y_m)\,[P(y, z, x_1, \cdots, x_n, y_1, \cdots, y_m) = 0]\}$$

and

$$S = \{\langle y, x_1, \cdots, x_n\rangle \mid (\forall z)_{\leq y}(\exists y_1, \cdots, y_m)\,[P(y, z, x_1, \cdots, x_n, y_1, \cdots, y_m) = 0]\},$$

then R and S are Diophantine.

That R is Diophantine is trivial. Namely,

$$\langle y, x_1, \cdots, x_n\rangle \in R \Leftrightarrow (\exists z, y_1, \cdots, y_m)\,(z \leqq y \,\&\, P = 0).$$

The proof of the other half of the theorem is far more complicated.

LEMMA 5.1.

$$(\forall k)_{\leq y}(\exists y_1, \cdots, y_m)\,[P(y, k, x_1, \cdots, x_n, y_1, \cdots, y_m) = 0]$$

$$\Leftrightarrow$$

$$(\exists u)\,(\forall k)_{\leq y}(\exists y_1, \cdots, y_m)_{\leq u}[P(y, k, x_1, \cdots, x_n, y_1, \cdots, y_m) = 0].$$

Proof. The right side of the equivalence trivially implies the left side. For the converse, suppose the left side is true for given $y, x_1, \cdots, x_n$. Then for each $k = 1, 2, \cdots, y$ there are definite numbers $y_1^{(k)}, \cdots, y_m^{(k)}$ for which:

$$P(y, k, x_1, \cdots, x_n, y_1^{(k)}, \cdots, y_m^{(k)}) = 0.$$

Taking u to be the maximum of the my numbers

$$\{y_j^{(k)} \mid j = 1, \cdots, m;\ k = 1, 2, \cdots, y\},$$

it follows that the right side of the equivalence is likewise true.

LEMMA 5.2. *Let $Q(y, u, x_1, \cdots, x_n)$ be a polynomial with the properties:*

(1) $Q(y, u, x_1, \cdots, x_n) > u$, (2) $Q(y, u, x_1, \cdots, x_n) > y$,

(3) $k \leq y$ *and* $y_1, \cdots, y_m \leq u$ *imply* $|P(y, k, x_1, \cdots, x_n, y_1, \cdots, y_m)| \leqq Q(y, u, x_1, \cdots, x_n)$.

Then,

$$(\forall k)_{\leq y}(\exists y_1, \cdots, y_m)_{\leq u}[P(y, k, x_1, \cdots, x_n, y_1, \cdots, y_m) = 0]$$

$$\Leftrightarrow$$

$$(\exists c, t, a_1, \cdots, a_m)\ \Big[1 + ct - \prod_{k=1}^{y} (1 + kt)$$

$$\& t = Q(y, u, x_1, \cdots, x_n)! \,\&\, 1 + ct \,\Big|\, \prod_{j=1}^{u} (a_1 - j)$$

$$\& \cdots \& 1 + ct \,\Big|\, \prod_{j=1}^{u} (a_m - j)$$

$$\& P(y, c, x_1, \cdots, x_n, a_1, \cdots, a_m) \equiv 0 \mod 1 + ct].$$

The point of this lemma is that while the right side of the equivalence seems the more complicated of the two, it is free of bounded universal quantifiers.

Proof. First the implication in the $\Leftarrow$ direction:

For each $k = 1, 2, \cdots, y$, let p_k be a prime factor of $1 + kt$. Let $y_i^{(k)}$ be the remainder when a_i is divided by p_k $(k = 1, 2, \cdots, y;\ i = 1, 2, \cdots, m)$. It will follow that for each k, i:

(a) $$1 \leqq y_i^{(k)} \leqq u$$

(b) $$P(y, k, x_1, \cdots, x_n, y_1^{(k)}, \cdots, y_m^{(k)}) = 0.$$

To demonstrate (a), note that $p_k \mid 1 + kt$, $1 + kt \mid 1 + ct$ and $1 + ct \mid \prod_{j=1}^{u} (a_i - j)$. I.e., $p_k \mid \prod_{j=1}^{u} (a_i - j)$. Since p_k is a prime, $p_k \mid a_i - j$ for some $j = 1, 2, \cdots, u$. That is

$$j \equiv a_i \equiv y_i^{(k)} \mod p_k.$$

Since $t = Q(y, u, x_1, \cdots, x_n)!$, (2) implies that every divisor of $1 + kt$ must be $> Q(y, u, x_1, \cdots, x_n)$. So $p_k > Q(y, u, x_1, \cdots, x_n)$ and by (1), $p_k > u$. Hence $j \leqq u < p_k$. Since $y_i^{(k)}$ is the *remainder* when a_i is divided by p_k, also $y_i^{(k)} < p_k$. So,

$$y_i^{(k)} = j.$$

To demonstrate (b), first note that

$$1 + ct \equiv 1 + kt \equiv 0 \mod p_k.$$

Hence

$$k + kct \equiv c + kct \mod p_k,$$

i.e., $k \equiv c \mod p_k$. We have already obtained

$$y_i^{(k)} \equiv a_i \mod p_k.$$

Thus,

$$P(y, k, x_1, \cdots, x_n, y_1^{(k)}, \cdots, y_m^{(k)}) \equiv P(y, c, x_1, \cdots, x_n, a_1, \cdots, a_m)$$

$$\equiv 0 \mod p_k.$$

Finally

$$\left| P(y, k, x_1, \cdots, x_n, y_1^{(k)}, \cdots, y_m^{(k)}) \right| \leqq Q(y, u, x_1, \cdots, x_n) < p_k.$$

This proves (b) and completes the proof of the $\Leftarrow$ implication.

To prove the $\Rightarrow$ implication, let

$$P(y, k, x_1, \cdots, x_n, y_1^{(k)}, \cdots, y_m^{(k)}) = 0,$$

for each $k = 1, 2, \cdots, t$, where each $y_j^{(k)} \leqq u$. We set $t = Q(y, u, x_1, \cdots, x_n)!$, and since $\prod_{k=1}^{y}(1 + kt) \equiv 1 \bmod t$, we can find c such that

$$1 + ct = \prod_{k=1}^{y} (1 + kt).$$

Now, it is claimed that for $1 \leqq k < l \leqq y$,

$$(1 + kt, 1 + lt) = 1.$$

For, let $p \mid 1 + kt, p \mid 1 + lt$. Then $p \mid l - k$, so $p < y$. But since $Q(y, u, x_1, \cdots, x_n) > y$ this implies $p \mid t$ which is impossible. Thus the numbers $1 + kt$ form an *admissible sequence of moduli* and the Chinese Remainder Theorem (Theorem 1.2) may be applied to yield, for each i, $1 \leqq i \leqq m$, a number a_i such that

$$a_i \equiv y_i^{(k)} \bmod 1 + kt, \quad k = 1, 2, \cdots, y.$$

As above, $k \equiv c \bmod 1 + kt$. So

$$\begin{aligned} P(y, c, x_1, \cdots, x_n, a_1, \cdots, a_m) &\equiv P(y, k, x_1, \cdots, x_n, y_1^{(k)}, \cdots, y_m^{(k)}) \bmod 1 + kt, \\ &= 0. \end{aligned}$$

Since the numbers $1 + kt$ are relatively prime in pairs and each divides $P(y, c, x_1, \cdots, x_n, a_1, \cdots, a_m)$ so does their product. I.e.,

$$P(y, c, x_1, \cdots, x_n, a_1, \cdots, a_m) \equiv 0 \bmod 1 + ct.$$

Finally,

$$a_i \equiv y_i^{(k)} \bmod 1 + kt,$$

i.e.,

$$1 + kt \mid a_i - y_i^{(k)}.$$

Since $1 \leqq y_i^{(k)} \leqq u$,

$$1 + kt \mid \prod_{j=1}^{u} (a_i - j).$$

And again since the $1 + kt$'s are relatively prime to one another,

$$1 + ct \mid \prod_{j=1}^{u} (a_i - j).$$

Now it is easy to complete the proof of Theorem 5.1 using Lemmas 5.1 and 5.2. First find a polynomial Q satisfying (1), (2), (3) of Lemma 5.2. This is easy to do: Write

$$P(y, k, x_1, \cdots, x_n, y_1, \cdots, y_m) = \sum_{r=1}^{N} t_r$$

where each t_r has the form

$$t_r = \left| c \right| y^a k^b x_1^{q_1} x_2^{q_2} \cdots x_n^{q_n} y_1^{s_1} y_2^{s_2} \cdots y_m^{s_m}$$

for c an integer positive or negative. Set $u_r = c\, y^{a+b} x_1^{q_1} x_2^{q_2} \cdots x_n^{q_n} u^{s_1+s_2+\cdots+s_m}$ and let

$$Q(y, u, x_1, \cdots, x_n) = u + y + \sum_{r=1}^{N} u_r.$$

Then (1), (2), and (3) of Lemma 5.2 hold trivially. Thus:

$$(\forall k)_{\leq y}(\exists y_1, \cdots, y_m)\,[P(y, k, x_1, \cdots, x_n, y_1, \cdots, y_m) = 0]$$

$$\Leftrightarrow$$

$$(\exists u, c, t, a_1, \cdots, a_m)\Big[1 + ct = \prod_{k=1}^{y} (1 + kt)$$

$$\& t = Q(y, u, x_1, \cdots, x_n)! \,\&\, 1 + ct \Big| \prod_{j=1}^{u} (a_1 - j)$$

$$\& \cdots \& 1 + ct \Big| \prod_{j=1}^{u} (a_m - j)$$

$$\& P(y, c, x_1, \cdots, x_n, a_1, \cdots, a_m) = 0 \mod 1 + ct\Big]$$

$$\Leftrightarrow$$

$$(\exists u, c, t, a_1, \cdots, a_n, e, f, g_1, \cdots, g_m, h_1, \cdots, h_m, l)$$

$$\Big[e = 1 + ct \,\&\, e = \prod_{k=1}^{y} (1 + kt) \,\&\, f = Q(y, u, x_1, \cdots, x_n)$$

$$\& t = f! \,\&\, g_1 = a_1 - u - 1 \,\&\, g_2 = a_2 - u - 1 \,\&\, \cdots \,\&\, g_m = a_m - u - 1$$

$$\& h_1 = \prod_{k=1}^{u} (g_1 + k) \,\&\, h_2 = \prod_{k=1}^{u} (g_2 + k)$$

$$\& \cdots \& h_m = \prod_{k=1}^{u} (g_m + k) \,\&\, e \Big| h_1 \,\&\, e \Big| h_2 \,\&\, \cdots \,\&\, e \Big| h_m$$

$$\& l = P(y, c, x_1, \cdots, x_n, a_1, \cdots, a_n) \,\&\, e \Big| l\Big]$$

and this is Diophantine by Theorem 4.1.

6. Recursive functions. So far one trick after another has been used to show that various sets are Diophantine. But now very powerful methods are available: it turns out that the expanded version of the language of Diophantine predicates, permitting the use of bounded quantifiers (sanctioned by Theorem 5.1) together with the Sequence Number Theorem (Theorem 1.3) enables one to show in quite a straightforward way that almost any set we please is Diophantine.

Some examples are in order:

(i) *the set P of prime numbers*:

$$x \in P \Leftrightarrow x > 1 \,\&\, (\forall y, z)_{\leq x}[yz < x \vee yz > x \vee y = 1 \vee z = 1].$$

Another Diophantine definition of the primes is:

$$x \in P \Leftrightarrow x > 1 \,\&\, ((x-1)!\,, x) = 1$$

$$\Leftrightarrow x > 1 \,\&\, (\exists y, z, u, v)\,[y = x - 1 \,\&\, z = y! \,\&\, (uz - vx)^2 = 1];$$

but the first definition is the more natural one.

From Theorem 1.4 it follows that *there is a "prime-representing" polynomial* P, i.e., a positive integer is prime if and only if it is in the range of P. For an explicit construction of such a polynomial P, cf. [23a].

(ii) *the function* $g(y) = \prod_{k=1}^{y} (1 + k^2)$. Here we use the Sequence Number Theorem to "encode" the sequence $g(1), g(2), \cdots, g(y)$ into a single number u, i.e., so that

$$S(i, u) = g(i), \qquad i = 1, 2, \cdots, y.$$

Thus, $z = g(y)$

$$\Leftrightarrow (\exists u)\ \{S(1, u) = 2 \,\&\, (\forall k)_{\leq y}[k - 1 \vee (S(k, u) = (1 + k^2)S(k - 1, u))] \,\&\, z = S(y, u)\}$$

$$\Leftrightarrow (\exists u)\,\{S(1, u) = 2 \,\&\, (\forall k)_{\leq y}[k = 1 \vee (\exists a, b, c)\,(a = k - 1$$

$$\&\, b = S(a, u) \,\&\, c = S(k, u) \,\&\, c = (1 + k^2)b)] \,\&\, z = S(y, u)\}.$$

By now it is clear that the available methods are quite general. They are so powerful that the question becomes: how can any "reasonable" set or function escape these methods, i.e., not be Diophantine?

The strength of the methods can be tested by considering the class of all *computable* or *recursive* functions. These are the functions which can be computed by a finite program or computing machine having arbitrarily large amounts of time and memory at its disposal. Many rigorous definitions of this class (all of them equivalent) are available. One of the simplest is as follows:

The *recursive functions*[3] are all those functions obtainable from the initial functions

$$c(x) = 1,\ s(x) = x + 1;\ U_i^n(x_1, \cdots, x_n) = x_i, \qquad 1 \leqq i \leqq n;$$

$S(i, u)$ (The sequence number function)[4]

iteratively applying the three operations: *composition*, *primitive recursion*, and *minimalization* defined below:

COMPOSITION yields the function

$$h(x_1, \cdots, x_n) = f(g_1(x_1, \cdots, x_n), \cdots, g_m(x_1, \cdots, x_n))$$

from the given functions $g_1, \cdots, g_m$ and $f(t_1, \cdots, t_m)$.

PRIMITIVE RECURSION yields the function $h(x_1, \cdots, x_n, z)$ which satisfies the equations:

$$h(x_1, \cdots, x_n, 1) = f(x_1, \cdots, x_n)$$

$$h(x_1, \cdots, x_n, t + 1) = g(t, h(x_1, \cdots, x_n, t), x_1, \cdots, x_n),$$

rom the given functions f, g.

When $n = 0$, f becomes a constant so that h is obtained directly from g.

MINIMALIZATION yields the function:

$$h(x_1, \cdots, x_n) = \min_y[f(x_1, \cdots, x_n, y) = g(x_1, \cdots, x_n, y)]$$

from the given functions f, g assuming that f, g are such that for each $x_1, \cdots, x_n$ there is at least one y satisfying the equation $f(x_1, \cdots, x_n, y) = g(x_1, \cdots, x_n, y)$: (i.e., h must be everywhere defined).

The main result of this article is:

THEOREM 6.1. *A function is Diophantine if and only if it is recursive.*

To begin with, consider the following short list of recursive functions:

(1) $x + y$ is recursive since

$$x + 1 = s(x),$$

$$x + (t + 1) = s(x + t) = g(t, x + t, x),$$

where $g(u, v, w) = s(U_2^3(u, v, w))$.

(2) $x \cdot y$ is recursive since

$$x \cdot 1 = U_1^1(x)$$

$$x \cdot (t + 1) = (x \cdot t) + x = g(t, x \cdot t, x),$$

where $g(u, v, w) = U_2^3(u, v, w) + U_3^3(u, v, w)$.

(3) For each fixed k, the constant function $c_k(x) = k$ is recursive, since $c_1(x)$ is one of the initial functions and $c_{k+1}(x) = c_k(x) + c(x)$.

(4) Any polynomial $P(x_1, \cdots, x_n)$ with *positive* integer coefficients is recursive, since any such function can be expressed by a finite iteration of additions and multiplications of variables and $c(x)$. E.g.,

$$2x^2y + 3xz^3 + 5 = c_2(x) \cdot x \cdot x \cdot y + c_3(x) \cdot x \cdot z \cdot z \cdot z + c_5(x).$$

So (1), (2), (3), and composition gives the result.

Now it is easy to see that every Diophantine function is recursive:

Let f be Diophantine, and write:

$$y = f(x_1, \cdots, x_n) \Leftrightarrow (\exists t_1, \cdots, t_m)\,[P(x_1, \cdots, x_n, y, t_1, \cdots, t_m) = Q(x_1, \cdots, x_n, y, t_1, \cdots, t_m)],$$

where P, Q are polynomials with *positive* integer coefficients. Then, by the sequence number theorem:

$$f(x_1, \cdots, x_n) = S(1, \min_u[P(x_1, \cdots, x_n, S(1,u), S(2,u), \cdots, S(m+1,u)) = Q(x_1, \cdots, x_n, S(1,u), S(2,u), \cdots, S(m+1,u))]).$$

Since P, Q, $S(i,u)$ are recursive, so is f (using composition and minimalization).

To obtain the converse: $S(i,u)$ is known to be Diophantine; the other initial functions are trivially Diophantine. Hence it suffices to prove that the Diophantine functions are closed under composition, primitive recursion and minimalization.

Composition: If $h(x_1, \cdots, x_n) = f(g_1(x_1, \cdots, x_n), \cdots, g_m(x_1, \cdots, x_n))$, where $f, g_1, \cdots, g_m$ are Diophantine, then so is h since

$$y = h(x_1, \cdots, x_n) \Leftrightarrow (\exists t_1, \cdots, t_m)\,[t_1 = g_1(x_1, \cdots, x_n) \,\&\, \cdots \,\&\, t_m = g_m(x_1, \cdots x_n) \,\&\, y = f(t_1, \cdots, t_m)].$$

Primitive Recursion: If

$$h(x_1, \cdots, x_n, 1) = f(x_1, \cdots, x_n)$$

$$h(x_1, \cdots, x_n, t+1) = g(t, h(x_1, \cdots, x_n, t), x_1, \cdots, x_n),$$

and f, g are Diophantine, then (using the sequence number theorem to "code" the numbers $h(x_1, \cdots, x_n, 1), h(x_1, \cdots, x_n, 2), \cdots, h(x_1, \cdots, x_n, z)$:

$$y = h(x_1, \cdots, x_n, z) \Leftrightarrow$$

$$(\exists u)\,\{(\exists v)\,[v = S(1,u) \,\&\, v = f(x_1, \cdots, x_n)]$$

$$\&\, (\forall t)_{\le z}[(t = z) \vee (\exists v)\,(v = S(t+1,u)$$

$$\&\, v = g(t, S(t,u), x_1, \cdots, x_n))] \,\&\, y = S(z,u)\}$$

so that (using Theorem 5.1) h is Diophantine.

Minimalization: If

$$h(x_1, \cdots, x_n) = \min_y[f(x_1, \cdots, x_n, y) = g(x_1, \cdots, x_n, y)],$$

where f, g are Diophantine, then so is h since,

$$\begin{aligned} y = h(x_1, \cdots, x_n) \Leftrightarrow \\ (\exists z)[z = f(x_1, \cdots, x_n, y) \,\&\, z = g(x_1, \cdots, x_n, y)] \\ \&\,(\forall t)_{\leq y}[(t = y) \vee (\exists u, v)\,(u = f(x_1, \cdots, x_n, t) \\ \&\, v = g(x_1, \cdots, x_n, t) \,\&\, (u < v \vee v < u)]. \end{aligned}$$

7. A universal Diophantine set. An explicit enumeration of all the Diophantine sets of positive integers will now be described. Any polynomial with positive integer coefficients can be built up from 1 and variables by successive additions and multiplications. We fix the alphabet

$$x_0, x_1, x_2, x_3, \cdots$$

of variables and then set up the following enumeration of all such polynomials (using the pairing functions):

$$\begin{aligned} P_1 &= 1 \\ P_{3i-1} &= x_{i-1} \\ P_{3i} &= P_{L(i)} + P_{R(i)} \\ P_{3i+1} &= P_{L(i)} \cdot P_{R(i)}. \end{aligned}$$

Write $P_i = P_i(x_0, x_1, \cdots, x_n)$, where n is large enough so that all variables occurring in P_i are included. (Of course P_i will not in general depend on all of these variables.) Finally, let

$$D_n = \{x_0 \mid (\exists x_1, \cdots, x_n)\,[P_{L(n)}(x_0, x_1, \cdots, x_n) = P_{R(n)}(x_0, x_1, \cdots, x_n)]\}.$$

Here, $P_{L(n)}$ and $P_{R(n)}$ do not actually involve all of the variables $x_0, x_1, \cdots, x_n$—but clearly cannot involve any others. (Recall that $L(n), R(n) \leq n$.) By the way the sequence P_i has been constructed, it is seen that the sequence of sets:

$$D_1, D_2, D_3, D_4, \cdots$$

includes all Diophantine sets. Moreover:

THEOREM 7.1 (Universality Theorem[5]).

$$\{\langle n, x\rangle \mid x \in D_n\} \text{ is Diophantine.}$$

Proof. Once again using the sequence number theorem, it is claimed that:

$$x \in D_n \Leftrightarrow (\exists u)\{S(1,u) = 1 \,\&\, S(2,u) = x$$
$$\&\,(\forall i)_{\leq n}[S(3i,u) = S(L(i),u) + S(R(i),u)]$$
$$\&\,(\forall i)_{\leq n}[S(3i+1,u) = S(L(i),u) \cdot S(R(i),u)]$$
$$\&\, S(L(n),u) = S(R(n),u)\}.$$

It is clear enough that the predicate on the right-hand side of this equivalence is Diophantine, so it is only necessary to verify the claim:

Let $x \in D_n$ for given x, n. Then there are numbers $t_1, \cdots, t_n$ such that $P_{L(n)}(x, t_1, \cdots, t_n) = Q_{L(n)}(x, t_1, \cdots, t_n)$. Choose u (by the sequence number theorem) so that

$$(*) \qquad S(j,u) = P_j(x, t_1, \cdots, t_n), \qquad j = 1, 2, \cdots, 3n+2.$$

Then in particular $S(2,u) = x$ and $S(3i-1,u) = t_{i-1}$, $i = 2, 3, \cdots, n+1$. Thus the right-hand side of the equivalence is true.

Conversely, let the right-hand side hold for given n, x. Set

$$t_1 = S(5,u),\ t_2 = S(8,u),\ \cdots, t_n = S(3n+2,u).$$

Then, (*) must be true. Since $S(L(n),u) = S(R(n),u)$, it must be the case that

$$P_{L(n)}(x, t_1, \cdots, t_n) = P_{R(n)}(x, t_1, \cdots, t_n),$$

so that $x \in D_n$.

Since $D_1, D_2, D_3, \cdots$, gives an enumeration of all Diophantine sets, it is easy to construct a set different from all of them and hence non-Diophantine. That is, define:

$$V = \{n \mid n \notin D_n\}.$$

THEOREM 7.2. *V is not Diophantine.*

Proof. This is a simple application of Cantor's diagonal method. If V were Diophantine, then for some fixed i, $V = D_i$. Does $i \in V$? We have:

$$i \in V \Leftrightarrow i \in D_i; \qquad i \in V \Leftrightarrow i \notin D_i.$$

This is a contradiction.

THEOREM 7.3. *The function $g(n,x)$ defined by:*

$$g(n,x) = 1 \quad \text{if} \quad x \notin D_n,$$
$$g(n,x) = 2 \quad \text{if} \quad x \in D_n,$$

is not recursive.

Proof. If g were recursive then it would be Diophantine (Theorem 6.1), say:

$$y = g(n,x) \Leftrightarrow (\exists y_1, \cdots, y_m)\,[P(n,x,y,y_1,\cdots,y_m) = 0].$$

But then, it would follow that

$$V = \{x \mid (\exists y_1, \cdots, y_m)\,[P(x,x,1,y_1,\cdots,y_m) = 0]\}$$

which contradicts Theorem 7.2.

Using Theorem 7.1, write:

$$x \in D_n \Leftrightarrow (\exists z_1, \cdots, z_k)\,[P(n,x,z_1,\cdots,z_k) = 0].$$

where P is some definite (though complicated) polynomial. Suppose there were an algorithm for testing Diophantine equations for possession of positive integer solutions; i.e., *an algorithm for Hilbert's tenth problem*! Then for given n, x this algorithm could be used to test whether or not the equation

$$P(n,x,z_1,\cdots,z_k) = 0$$

has a solution, i.e., whether or not $x \in D_n$. Thus the algorithm could be used to compute the function $g(n,x)$. Since the recursive functions are just those for which a computing algorithm exists, g would have to be recursive. This would contradict Theorem 7.3, and this contradiction proves:

THEOREM 7.4. *Hilbert's tenth problem is unsolvable!*

Naturally this result gives no information about the existence of solutions for any *specific* Diophantine equation; it merely guarantees that there is no single algorithm for testing the class of all Diophantine equations. Also note that:

$$\begin{aligned} x \in V &\Leftrightarrow \sim(\exists z_1, \cdots, z_k)\,[P(x,x,z_1,\cdots,z_k) = 0] \\ &\Leftrightarrow \{(\exists z_1, \cdots, z_k)\,[P(x,x,z_1,\cdots,z_k) = 0] \to 1 = 0\} \\ &\Leftrightarrow (\forall z_1, \cdots, z_k)\,[P(x,x,z_1,\cdots,z_k) > 0 \\ &\qquad\qquad \vee\ P(x,x,z_1,\cdots,z_k) < 0] \end{aligned}$$

which shows that if either $\sim$ or unbounded universal quantifiers $(\forall z)$ or implication $(\to)$ are permitted in the language of Diophantine predicates, then non-Diophantine sets will be produced.

It is natural to associate with each Diophantine set a *dimension* and a *degree*; i.e., the *dimension of* S is the least n for which a polynomial P exists for which:

$$(*) \qquad S = \{x \mid (\exists y_1, \cdots, y_n)\,[P(x,y_1,\cdots,y_n) = 0]\},$$

and the *degree of* S is the least degree of a polynomial P satisfying (*) (permitting n to be as large as one likes). Now it is easy to see:

THEOREM 7.5. *Every Diophantine set has degree* ≤ 4.

Proof. The degree of P satisfying (*) may be reduced by introducing additional

variables z_j satisfying equations of the form

$$z_j = y_i y_k$$
$$z_j = y_i^2$$
$$z_j = x y_i$$
$$z_j = x^2.$$

By successive substitutions of the z_j's into P its degree can be brought down to 2. Hence the equation is equivalent to a system of simultaneous equations each of degree 2. Summing the squares gives an equation of degree 4.

A less trivial (and more surprising) fact is:

THEOREM 7.6. *There is an integer m such that every Diophantine set has dimension $\leqq m$.*

Proof. Write

$$D_n = \{x \mid (\exists y_1, \cdots, y_m)\,[P(x, n, y_1, \cdots, y_m) = 0]\},$$

which is possible by the universality theorem. Then the dimension of D_n is $\leqq m$ for all n.

An interesting example is given by the sequence of Diophantine sets:

$$S_q = \{x \mid (\exists y_1, \cdots, y_q)\,[x = (y_1 + 1)\cdots(y_q + 1)]\}.$$

Here S_2 is the set of composite numbers; S_q is the set of "q-fold" composite numbers. It is surely surprising that it is possible to give a Diophantine definition of S_q (for large q) requiring fewer than q parameters (cf. [19]).

How large is m, the number of parameters in the universal Diophantine set? A direct calculation using the arguments given here would yield a number around 50. Actually Matiyacevič and Julia Robinson have very recently shown that $m = 14$ will suffice!

The unsolvability of Hilbert's tenth problem can be used to obtain a strengthened form of Gödel's famous incompleteness theorem:

THEOREM 7.7. *Corresponding to any given axiomatization of number theory, there is a Diophantine equation which has no positive integer solutions, but such that this fact cannot be proved within the given axiomatization.*

A rigorous proof would involve a precise definition of "axiomatization of number theory" which is outside the scope of this article. An informal heuristic argument follows:

One uses the given axiomatization to systematically generate all of the theorems (i.e., consequences of the axioms). Among these theorems will be some asserting

that some Diophantine equation has no solution. Whenever such is encountered it is placed on a special list called LISTA. At the same time a list, LIST B, is made of Diophantine equations which have solutions. LIST B is constructed by a search procedure, e.g., at the nth stage of the search look at the first n Diophantine equations (in a suitable list) and test for solutions in which each argument is $\leqq n$. Thus every Diophantine equation which has positive integer solutions will eventually be placed in LIST B. *If* likewise each Diophantine equation with no solutions would eventually appear in LIST A, then one would have an algorithm for Hilbert's tenth problem. Namely, to test a given equation for possession of a solution simply begin generating LIST A and LIST B until the given equation appears in one list or the other. Since Hilbert's tenth problem is unsolvable, some equation with no solution must be omitted from LIST A. But this is just the assertion of the theorem.

8. Recursively enumerable sets. It is now time to settle the question raised at the beginning: which sets are Diophantine?

DEFINITION. 8.1. *A set S of n-tuples of positive integers is called recursively enumerable if there are recursive functions $f(x, x_1, \cdots, x_n)$, $g(x, x_1, \cdots, x_n)$ such that:*

$$S = \{\langle x_1, \cdots, x_n\rangle \mid (\exists x)\,[f(x, x_1, \cdots, x_n) = g(x, x_1, \cdots, x_n)]\}.$$

THEOREM 8.1. *A set S is Diophantine if and only if it is recursively enumerable.*

Proof. If S is Diophantine there are polynomials P, Q with positive coefficients such that:

$$\langle x_1, \cdots, x_n\rangle \in S \Leftrightarrow (\exists y_1, \cdots, y_m)\,[P(x_1, \cdots, x_n, y_1, \cdots, y_m) = Q(x_1, \cdots, x_n, y_1, \cdots, y_m)]$$

$$\Leftrightarrow (\exists u)\,[P(x_1, \cdots, x_n, S(1,u), \cdots, S(m,u)) = Q(x_1, \cdots, x_n, S(1,u), \cdots, S(m,u))],$$

so that S is recursively enumerable.

Conversely if S is recursively enumerable there are recursive functions $f(x, x_1, \cdots, x_n)$, $g(x, x_1, \cdots, x_n)$ such that

$$\langle x_1, \cdots, x_n\rangle \in S \Leftrightarrow (\exists x)\,[f(x, x_1, \cdots, x_n) = g(x, x_1, \cdots, x_n)]$$

$$\Leftrightarrow (\exists x, z)\,[z = f(x, x_1, \cdots, x_n)\ \&\ z = g(x, x_1, \cdots, x_n)].$$

Thus by Theorem 6.1, S is Diophantine.

9. Historical appendix. The present exposition has ignored the chronological order in which the ideas were developed. The first contribution was by Gödel in his celebrated 1931 paper [16]. The main point of Gödel's investigation was the existence of undecidable statements in formal systems. The undecidable statements Gödel obtained involved recursive functions, and in order to exhibit the simple number-theoretic character of these statements, Gödel used the Chinese remainder theorem to reduce them to "arithmetic" form. The technique used is just what is

used here in proving Theorem 1.3 (the sequence number theorem) and Theorem 6.1 (in the direction: every recursive function is Diophantine). However without the techniques for dealing with bounded universal quantifiers as discussed in this paper, the best result yielded by Gödel's methods is that every recursive function (and indeed every recursively enumerable set) can be defined by a Diophantine equation preceded by a finite number of existential and bounded universal quantifiers[6]. In my doctoral dissertation (cf. [5], [6]), I showed that all but one of the bounded universal quantifiers could be eliminated, so that every recursively enumerable set S could be defined as

$$S = \{x \mid (\exists y)\,(\forall k)_{\leqq y}(\exists y_1, \cdots, y_m)\,[P(k, x, y, y_1, \cdots, y_m) = 0]\}.$$

This representation became known as the Davis normal form. (Later R. M. Robinson [31], [32] showed that in this normal form one could take $m = 4$. More recently Matiyacevič has shown that one can even take $m = 2$. It is known that one cannot always have $m = 0$; whether one can always get $m = 1$ is open.)

Independent of my work and at about the same time, Julia Robinson began her study [27] of Diophantine sets. Her investigations centered about the question: *Is the exponential function Diophantine?* The main result was that a certain hypothesis implied that the exponential function was Diophantine. The hypothesis, which became known as the Julia Robinson hypothesis, has played a key role in work on Hilbert's tenth problem. Its statement is simply:

There exists a Diophantine set D such that:

(1) *$\langle u, v\rangle \in D$ implies $v \leqq u^u$.*

(2) *For each k, there is $\langle u, v\rangle \in D$ such that $v > u^k$.*

The hypothesis remained an open question for about 2 decades. (Actually the set

$$D = \{\langle u, v\rangle \mid v = x_u(2) \,\&\, u > 3\}$$

satisfies (1) and (2) by Lemma 2.19 and is Diophantine by Corollary 3.2, so the truth of Julia Robinson's hypothesis follows at once from the results in this article.) Julia Robinson's proof that this hypothesis implies that the exponential function is Diophantine used the Pell equation. And, the proof that the exponential function is indeed Diophantine given here is closely related to a more recent proof [28] by her of this same implication.

In [27], Julia Robinson studied also sets and functions which were *exponential Diophantine* (or existentially definable in terms of exponentiation) that is which possess definitions of the form:

$$(\exists u_1, \cdots, u_n, v_1, \cdots, v_n, w_1, \cdots, w_n)\,[P(x_1, \cdots, x_m, u_1, \cdots, u_n, v_1, \cdots, v_n, w_1, \cdots, w_n) = 0$$
$$\&\, u_1 = v_1^{w_1} \,\&\, \cdots \,\&\, u_n = v_n^{w_n}].$$

In particular, the functions $\binom{n}{k}$ and $n!$ were shown by her to be exponential

Diophantine. This is really what is shown in proving (1) and (2) of Theorem 4.1. The present proof of (2) is just hers; the proof of (1) given here is a simplified variant of that in [27]. (It is due independently to Julia Robinson and Matiyasevič.)

The idea of using the Chinese remainder theorem to code the effect of a bounded universal quantifier first occurred in the work of myself and Putnam [7]. In [8], we refined our methods and were able to show, beginning with the Davis normal form, that *IF* there are arbitrarily long arithmetic progressions consisting entirely of primes (still an open question), then every recursively enumerable set is exponential Diophantine. In our proof we needed to establish that $h(a, b, y) = \prod_{k=1}^{y} (a + bk)$ is exponential Diophantine, which we did extending Julia Robinson's methods. (The proof given here of (3) of Theorem 4.1 is a much simplified argument found much later by Julia Robinson—cf. [29].) Julia Robinson then showed first how to eliminate the hypothesis about primes in arithmetic progression, and then how to greatly simplify the proof along the lines of Lemma 5.2 of this article. Thus we obtained the theorem of [9] that every recursively enumerable set is exponential Diophantine.

Attention was now focused on the Julia Robinson hypothesis since it was plain that it would imply that Hilbert's tenth problem was unsolvable.

Many interesting propositions were found to imply the Julia Robinson hypothesis.[7]. However the hypothesis seemed implausible to many, especially because it was realized that an immediate and surprising consequence would be the existence of an absolute upper bound for the dimensions of Diophantine sets (cf. Theorem 7.6). Thus in his review [19] Kreisel said concerning the results of [9]: "... it is likely the present result is not closely connected with Hilbert's tenth problem. Also it is not altogether plausible that all (ordinary) Diophantine problems are uniformly reducible to those in a fixed number of variables of fixed degree... ."

The Julia Robinson hypothesis was finally proved by Matiyasevič [23], [24]. Specifically he showed that if we define

$$a_1 = a_2 = 1, \quad a_{n+1} = a_n + a_{n-1}$$

so that a_n is the nth Fibonacci number, then the function a_{2n} is diophantine. Then since, for $n \geqq 3$, as is easily seen by induction,

$$\left(\frac{5}{4}\right)^n < a_n < 2^{n-1},$$

the set

$$D = \{\langle u, v\rangle \mid v = a_{2u} \,\&\, u \geqq 2\}$$

satisfies the Julia Robinson hypothesis. Subsequently, direct diophantine definitions of the exponential function were given by a number of investigators, several of them using the Pell equation as in this article (cf. [3], [4], [14], [18a]). The treatment in §2, 3 is based on Matiyasevič's methods, although the details are Julia Robinson's.

In particular, it was Matiyasevič who taught us how to use results like Lemmas 2.11, 2.12, and 2.22 of the present exposition. (Matiyasevič himself used analogous results for the Fibonacci numbers.)

It was soon noticed (by S. Kochen) that by a simple inductive argument the use of the Davis normal form could now be entirely avoided, as has been done in the present exposition.

Let $\#(P)$ be the number of solutions of the Diophantine equation $P = 0$. Thus $0 \leq \#(P) \leq \aleph_0$. Hilbert's tenth problem seeks an algorithm for deciding of a given P whether or not $\#(P) = 0$. But there are many related questions: Is there an algorithm for testing whether $\#(P) = \aleph_0$, or $\#(P) = 1$, or $\#(P)$ is even? I was able to show easily (beginning with the unsolvability of Hilbert's tenth problem) that all of these problems are unsolvable. In fact if

$$A = \{0,1,2,3,\cdots\aleph_0\}$$

and $B \subset A$, $B \neq \varnothing$, $B \neq A$, then one can readily show that there is no algorithm for determining whether or not $\#(P) \in B$ (cf. [15]).

The fact that no general algorithm such as Hilbert demanded will be forthcoming adds to the interest of algorithms for dealing with special classes of Diophantine equations. Alan Baker and his coworkers [1], [2] have in recent years made considerable progress in this direction.

Notes

1. These pairing functions (but of course not their being Diophantine) were used by Cantor in his proof of the countability of the rational numbers. J. Roberts and D. Siefkes each corrected an error in the definition of these functions. They, as well as W. Emerson, M. Hausner, Y. Matiyasevič, and Julia Robinson made helpful suggestions.

2. For example, cf. [25], pp. 175–180. Matiyasevič used instead the equations $x^2 - xy - y^2 = 1$, $u^2 - muv + v^2 = 1$.

3. The recursive functions are usually defined on the nonnegative integers. This creates a minor but annoying technical problem in comparing the present definition with one in the literature (e.g., cf. [6], p. 41; also Theorem 4.2 on p. 51). Thus one can simply note that $f(x_1,\cdots,x_n)$ is recursive in the present sense if and only if $f(t_1 + 1,\cdots,t_n + 1) - 1$ is recursive in the usual sense. From the point of view of the intuitive "computability" of the functions involved this doesn't matter at all; one is simply in the position of using the positive integers as a "code" for the nonnegative integers — using $n + 1$ to represent n.

4. Inclusion of $S(i, u)$ in this list is redundant. That is, $S(i, u)$ can be obtained using our three operations from the remaining initial functions.

5. The method of proof is Julia Robinson's, [28], [30]. If one were permitted to use the enumeration theorem in recursive function theory ([6], p. 67. Theorem 1.4), the Universality Theorem would follow at once from Theorem 6.1.

6. Actually the result which Gödel stated (as opposed to what can be obtained at once by use of his techniques) was somewhat weaker. Indeed, the very definition of the class of recursive functions and the perception of their significance came several years later in the work of Gödel, Church, and Turing. In particular the suggestion that recursiveness was a precise equivalent of the intuitive

notion of being computable by an explicit algorithm was made independently by Church and by Turing. And of course it is this identification which is essential in regarding the technical results discussed in this account as constituting a negative solution of Hilbert's tenth problem. (For further discussion and references, cf. [6].)

7. For example, I showed ([13]) that the Julia Robinson hypothesis would follow from the non-existence of nontrivial solutions of the equation

$$9(u^2 + 7v^2)^2 - 7(x^2 + 7y^2)^2 = 2.$$

The methods used readily show that the same conclusion follows if the equation has only finitely many solutions. Čudnovskii [4] claims to have proved that 2^x is diophantine (and hence the Julia Robinson hypothesis) using this equation. Apparently there is a possibility that some of Čudnovskii's work may have been done independently of Matiyasevič — but I have not been able to obtain definite information about this.

References

1. Alan Baker, Contributions to the theory of Diophantine equations: I. On the representation of integers by binary forms, II. The Diophantine equation $y^2 = x^3 + k$, Philos. Trans. Roy. Soc. London Ser. A, 263 (1968) 173–208.

2. Alan Baker, The Diophantine equation $y^2 = ax^3 + bx^2 + cx + d$, J. London Math. Soc., 43 (1968) 1–9.

3. G. V. Čudnovskii, Diophantine predicates (Russian), Uspehi Mat. Nauk, 25 (1970) no. 4 (154), pp. 185–186.

4. ———, Certain arithmetic problems (Russian), Ordena Lenina Akad. Ukrains. SSR, Preprint IM-71-3.

5. Martin Davis, Arithmetical problems and recursively enumerable predicates, J. Symbolic Logic, 18 (1953) 33–41.

6. ———, Computability and Unsolvability, McGraw Hill, New York, 1958.

7. Martin Davis and Hilary Putnam, Reduction of Hilbert's tenth problem, J. Symbolic Logic, 23 (1958) 183–187.

8. ——— and ———, On Hilbert's tenth problem, U. S. Air Force O. S. R. Report AFOSR TR 59–124 (1959), Part III.

9. Martin Davis, Hilary Putnam, and Julia Robinson, The decision problem for exponential Diophantine equations, Ann. Math., 74 (1961) 425–436.

10. Martin Davis, Applications of recursive function theory to number theory, Proc. Symp. Pure Math., 5 (1962) 135–138.

11. ———, Extensions and corollaries of recent work on Hilbert's tenth problem, Illinois J. Math., 7 (1963) 246–250.

12. Martin Davis and Hilary Putnam, Diophantine sets over polynomial rings, Illinois J. Math., 7 (1963) 251–256.

13. Martin Davis, One equation to rule them all, Trans. New York Acad. Sci., Series II, 30 (1968) 766–773.

14. ———, An explicit Diophantine definition of the exponential function, Comm. Pure Appl. Math. 24 (1971) 137–145.

15. ———, On the number of solutions of Diophantine equations, Proc. Amer. Math. Soc., 35 (1972) 552–554.

16. Kurt Gödel, Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I, Monatsh. Math. und Physik, 38 (1931) 173–198. English translations: (1) Kurt, Gödel, On Formally Undecidable Propositions of Principia Mathematica and Related Systems, Basic

Books, 1962. (2) Martin Davis (editor), The Undecidable, Raven Press, 1965, pp. 5–38. (3) Jean Van Heijenoort (editor), From Frege to Gödel, Harvard University Press, 1967, pp. 596–616.

17. G. H. Hardy and E. M. Wright, An Introduction to the Theory of Numbers, Fourth edition, Oxford University Press, 1960.

18. David Hilbert, Mathematichse Probleme, Vortrag, gehalten auf dem internationalen Mathematiker-Kongress zu Paris 1900. Nachrichten Akad. Wiss. Göttingen, Math. -Phys. Kl. (1900) 253–297. English translation: Bull. Amer. Math. Soc., 8 (1901–1902) 437–479.

18a. N. K. Kosovskii, On Diophantine representations of the solutions of Pell's equation (Russian), Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklova, 20 (1971) 49–59.

19. Georg Kreisel, Review of [9]. Mathematical Reviews, 24 (1962) Part A, p. 573 (review number A 3061).

20. Yuri Matiyasevič, The relation of systems of equations in words and their lengths to Hilbert's tenth problem (Russian). Issledovaniya po Konstruktivnoi Matematike i Matematiceskoi Logike II. Vol. 8, pp. 132–144.

21. ———, Two reductions of Hilbert's tenth problem (Russian), *Ibid.*, pp. 145–158.

22. ———, Arithmetic representation of exponentiation (Russian). *Ibid.*, pp. 159–165.

23. ———, Enumerable sets are Diophantine (Russian), Dokl. Akad. Nauk SSSR, 191 (1970) 279–282. Improved English translation: Soviet Math. Doklady, 11 (1970) 354–357.

23a. ———, Diophantine representation of the set of prime numbers (Russian). Dokl. Akad. Nauk SSSR, 196 (1971) 770–773. Improved English translation with Addendum: Soviet Math. Doklady, 12 (1971) 249–254.

23b. ———, Diophantine representation of recursively enumerable predicates, Proc. Second Scandinavian Logic Symp., editor, J. E. Fenstad, North-Holland, Amsterdam, 1971.

23c. ———, Diophantine representation of recursively enumerable predicates, Proc. 1970 Intern. Congress Math., pp. 234–238.

24. ———, Diophantine representation of enumerable predicates (Russian), Izv. Akad. Nauk SSSR, Ser. Mat. 35 (1971) 3–30.

24a. ———, Diophantine sets (Russian), Uspehi Mat. Nauk, 27(1972) 185–222.

25. Ivan Niven and Herbert Zuckerman, An Introduction to the Theory of Numbers, 2nd ed., Wiley, New York, 1966.

26. Hilary Putnam, An unsolvable problem in number theory, J. Symb. Logic, 25 (1960) 220–232.

27. Julia Robinson, Existential definability in arithmetic, Trans. Amer. Math. Soc., 72 (1952) 437–449.

28. ———, Diophantine decision problems, MMA Studies in Mathematics, 6 (1969) [Studies in Number Theory, edited by W. J. LeVeque, pp. 76–116].

29. ———, Unsolvable Diophantine problems, Proc. Amer. Math. Soc., 22 (1969) 534–538.

30. ———, Hilbert's tenth problem, Proc. Symp. Pure Math., 20 (1969) 191–194.

31. Raphael M. Robinson, Arithmetical representation of recursively enumerable sets, J. Symb. Logic, 21 (1956) 162–186.

32. ———, Some representations of Diophantine sets, J. Symb. Logic, forthcoming.

THE AMERICAN MATHEMATICAL MONTHLY, Vol. 80 No. 3, February 1973, pp. 117-232

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

THE OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA

VOLUME 80 NUMBER 3

CODEN: AMMYAE

CONTENTS

(*Continued on inside cover*)

MARCH 1973