# Proof, Completeness, Transcendentals, and Sampling

PHILIP J. DAVIS

*Brown University, Providence, Rhode Island*

ABSTRACT. This paper considers informally the relationship between computer aided mathematical proof, formal algebraic languages, computation with transcendental numbers, and proof by sampling.

KEY WORDS AND PHRASES: algebraic languages, completeness, transcendentality, sampling

CR CATEGORIES: 5 7

This is the fourth in a series of papers devoted to certain aspects of proof in elementary mathematics as it has been affected by the computing machine. The three previous papers are Davis and Cerrutti [7] and Davis [5, 6]. The point of view taken in the present paper is that of a person who is interested both in interpolatory function theory and in computer science.

## 1. *A Geometrical Theorem from Antiquity*

In [7], a computer proof of the following classical theorem of Pappus was discussed. Take any two straight lines in the plane and select three points arbitrarily on each. Connect the points in a crisscross fashion as indicated. The three points of the intersections of the crisscrosses are called the *Pappus points* for the original configuration, and the theorem states that the three Pappus points are collinear. The attack on this problem was through a brute application of coordinate geometry.

Deliberately, no attempt was made to achieve simplifying reductions. The two lines were given symbolic parametric form. Six symbolic points were selected thereon. The coordinates of the three Pappus points were obtained as functions of the parameters by solving the three $2\times2$ systems. The Pappus points were then shown to be collinear. (See Figure 1.)

This project was undertaken as an exercise in programming in FORMAC, an algebraic symbol manipulation language, and to explore the possibility of theorem discovery in elementary geometry by this means. This particular theorem was selected because it is simple to state, lies at the foundation of projective geometry (there are geometries in which the theorem is not true), has an interesting generalization (Pascal's theorem), but is difficult to prove with only high school geometry.

## 2. *Details of the Brute Force Analytics*

The coordinate geometry of the Pappus configuration is perhaps best programmed in the following way. Let a prototype crisscross be formed from the points with coordinates $(a, b), (c, d), (e, f), (g, h)$ as indicated in Figure 2. Set $[1] j = h - b$, $[2] k = g - a$, $[3] p = e - c$, $[4] q = f - d$, $[5] r = aj - bk$, $[6] s = cq - dp$, $[7] m = sk - pr$, $[8] n = js - qr$, $[9] w = qk - pj$. Then the $x$ and $y$ coordinates of the point of intersection of the crisscross are given, respectively, by $m/w$ and $n/w$. Now let the three arbitrary points on the two



Author's address Department of Applied Mathematics, Brown University, Providence, RI 02912

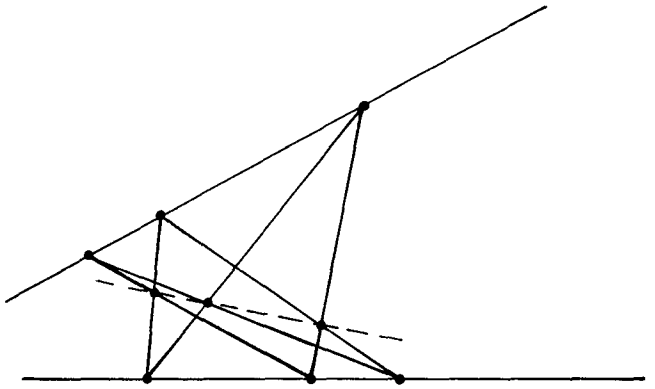

FIG 1

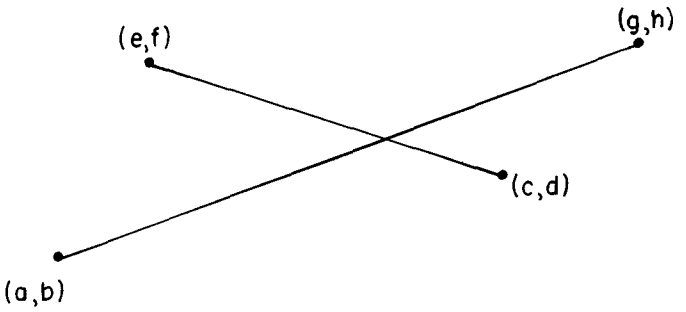


FIG 2

arbitrary lines be given coordinates $(x_i, y_i)$, $i = 1, 2, \ldots, 6$, as in Figure 3, and let the respective points of intersection of the three crisscrosses be $(M1/W1, N1/W1)$, $(M2/W2, N2/W2)$, $(M3/W3, N3/W3)$. Here, $M1, N1, W1$, etc., equal $m, n, w$ after an appropriate replacement of $a, b, \ldots, h$ by $x_1, \ldots, y_6$. The condition for the collinearity of the three Pappus points is now

$$\text{DET} = \begin{vmatrix} M1 & N1 & W1 \\ M2 & N2 & W2 \\ M3 & N3 & W3 \end{vmatrix} = 0$$

It should be clear that DET is a *polynomial* $P$ in the variables $x_i, y_i$. This polynomial is the sum of *several thousand monomials* in the $x_i$ and $y_i$; so the proof of Pappus's theorem by this means consists in the construction of DET and the verification that it is identically zero. In [7] this construction and verification was carried out in the FORMAC language.

## 3. *Was a Formal Algebra Language Necessary?*

The reason a formal algebra language (such as FORMAC) was selected is as follows. What is desired is a proof that is *valid generally*. Therefore it would not do merely to substitute *specific numerical values* for the coordinates and to verify numerically that DET = 0. This would establish Pappus only in the specific numerical case selected. Now in the usual programming languages such as Fortran, BASIC, and APL, normally operated, all variables must ultimately link back to numerical values. To perform formal algebra in them would require special programming. Therefore the computation was carried out within a language in which formal algebra is routinely available. (See Supplementary Note 1.[1])

Our demand for generality is similar to that frequently encountered by mathematics

[1] Supplementary Notes appear at the end of the paper

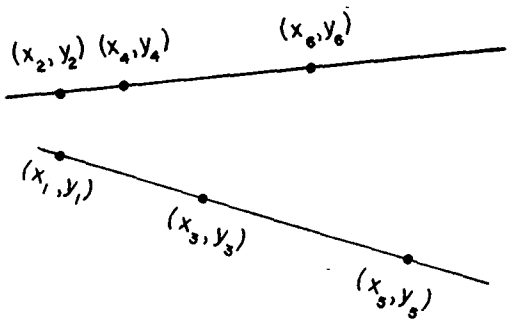


FIG. 3

teachers in elementary courses. A problem has been assigned in which something quite general is to be established. The teacher reads the student's proof and then writes, "You have proved the statement only for $n = 2$ (or for $(x, y) = (0, 0)$, or when $T$ is a *right* triangle, etc.). Prove in general!"

Now it turns out that, subject to some reservations developed below, it is not necessary to work entirely symbolically; the students' idea of limited numerical verification can provide mathematically valid proofs. This is true not merely for the Pappus theorem but for all theorems of elementary algebraic analytic geometry which are equivalent to theorems of polynomial algebra.

The link which enables us to reduce the problem to numerical computation is the so-called *uniqueness theorem for polynomials*. It will be stated first in one variable.

Let $p(x) = a_0x^n + a_1x^{n-1} + \cdots + a_n$ be a polynomial of degree less than or equal to $n$ in the real (or complex) variable $x$. It is assumed that $a_0, \ldots, a_n$ are real (or complex) numbers. If now $x_1, x_2, \ldots, x_{n+1}$ are $n + 1$ distinct real (or complex) numbers and if $p(x_1) = 0, p(x_2) = 0, \ldots, p(x_{n+1}) = 0$, then $p(x)$ is identically 0.

The uniqueness theorem therefore operates as a *reduction principle* or a *sampling principle,* enabling us to reduce the ostensibly infinite task of verifying that $p(x) = 0$ for all $x$ to that of verification for a *finite number of values* of $x$

A similar theorem is valid for polynomials in several variables. Suppose for example that for fixed $y$, $p(x, y)$ is a polynomial in $x$ of degree less than or equal to $m$, and for fixed $x$ of degree less than or equal to $n$ in $y$. In other words, let

$$p(x, y) = \sum_{\substack{j=1, \ldots, m \\ p=1, \ldots, n}} a_{jk}x^jy^k.$$

We may rewrite it in the form

$$p(x, y) = a_0(x) + a_1(x)\cdot y + a_2(x)\cdot y^2 + \cdots + a_n(x)y^n,$$

where the functions $a_i(x)$, $i = 0, \ldots, n$, are all polynomials in $x$ of degree less than or equal to $m$. Suppose now that $x_1, \ldots, x_{m+1}$ are distinct and $y_1, \ldots, y_{n+1}$ are distinct. Then, if $p(x_i, y_j) = 0$ for $i = 1, \ldots, m + 1$ and $j = 1, \ldots, n + 1$, it follows that $p(x, y)$ is identically 0 from the following argument. Let $i$ be fixed and consider $\tilde{p}_i(y) = p(x_i, y)$. Now assume that

$$0 = p(x_i, y_j) = \tilde{p}_i(y_j) = a_0(x_i) + a_1(x_i)y_j + \cdots + a_n(x_i)y_j^n = 0,$$
$$j = 1, 2, \ldots, n + 1, \quad i = 1, 2, \ldots, m + 1.$$

By the uniqueness theorem in one variable, since $\tilde{p}_i(y)$ is of degree less than or equal to $n$, all the coefficients $a_0(x_i), \ldots, a_n(x_i)$ must vanish. This must be true for $i = 1, 2, \ldots, m + 1$. Since they, in turn, are polynomials of degree less than or equal to $m$, each of them vanishes for $m + 1$ points and hence vanishes identically. Therefore $p(x, y)$ vanishes identically.

Note that the number of verifications that must be carried out is now $(m + 1)(n + 1)$. In general let $x_1, x_2, \ldots, x_s$ be $s$ independent variables and let $p(x_1, x_2, \ldots, x_s)$ be a polynomial of degree less than or equal to $d_i$ in $x_i$, $i = 1, \ldots, s$. Suppose that $x_{i1}$, $x_{i2}, \ldots, x_{id_i+1}$, $i = 1, 2, \ldots, s$, are $s$ sets of values that are distinct (as far as the second subscript is concerned); then

$$p(x_{1j_1}, x_{2j_2}, \ldots, x_{sj_s}) = 0, \quad \begin{cases} j_1 = 1, \ldots, d_1 + 1, \\ \vdots \\ j_s = 1, \ldots, d_s + 1 \end{cases}$$

implies $p \equiv 0$. Note that $(d_1 + 1)(d_2 + 1) \cdots (d_s + 1)$ individual verifications are sufficient to reduce the problem. (See Supplementary Note 2.) The points that must be substituted into $p$ constitute the product set of the points

$$\begin{aligned} &x_{11}, \ldots, x_{1d_1+1}, \\ &x_{21}, \ldots, x_{2d_2+1}, \\ &\vdots \\ &x_{s1}, \ldots, x_{sd_s+1}. \end{aligned}$$

*Example* 1. Consider the famous algebraic identity of Euler which plays a key role in the "four-square" problem.

$$\begin{aligned}(a_1^2 + a_2^2 + a_3^2 + a_4^2)(b_1^2 + b_2^2 + b_3^2 + b_4^2) = {} & (a_1b_1 - a_2b_2 - a_3b_3 - a_4b_4)^2 + (a_1b_2 + a_2b_1 \\ & + a_3b_4 - a_4b_3)^2 + (a_1b_3 - a_2b_4 + a_3b_1 + a_4b_2)^2 + (a_1b_4 + a_2b_3 - a_3b_2 + a_4b_1)^2.\end{aligned}$$

The difference between the left-hand and the right-hand sides is a polynomial in $s = 8$ variables and of degree at most two in each variable. Hence this formal identity may be proved by at most $3^8 = 6561$ numerical verifications on, say, the 8-fold product set of $(-1, 0, 1)$.

Several remarks are in order. Although this identity is crucial to the "four-square problem," books on number theory never prove it. After all, it is a "mere" formal computation to show that a certain polynomial consisting of 80 monomials is identically 0.[2] There is a similar identity with 16 variables. (See Supplementary Note 3.)

Nor does the verification of the identity provide the slightest insight into deeper meanings which have been found for the identity. (The four-square identity is equivalent to $\|Q_1\|^2 \|Q_2\|^2 = \|Q_1Q_2\|^2$, where $Q_i$ are quaternions and where $\|Q\|^2 = \|a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}\|^2 = a^2 + b^2 + c^2 + d^2$. See, e.g. Curtiss [3].) Nor does it provide any insight into how further identities of this kind may be constructed. It is a purely post hoc affair.

*Example* 2. The Pappus theorem. As observed in the second section, the proof of Pappus is equivalent to verifying that DET $\equiv 0$, where DET is a polynomial in the variables $x_1, \ldots, y_6$ and where $(x_i, y_i)$ lie by threes on two arbitrary lines. Parametrize the two lines as $l_1 : x = t, y = 0$; $l_2 : x = \gamma t, y = \alpha t + \beta$. Then select $(x_{2i+1}, y_{2i+1}) = (t_{2i+1}, 0)$, $i = 0, 1, 2$; $(x_{2i}, y_{2i}) = (\gamma t_{2i}, \alpha t_{2i} + \beta)$, $i = 1, 2, 3$. Thus DET is a polynomial in the nine independent variables $\alpha, \beta, \gamma, t_i$, $i = 1, 2, \ldots, 6$. Hence DET $\equiv 0$ may be proved by an appropriately selected finite sample of values. See Figure 3.

*Example* 3. Another proposition in elementary geometry proved via "finite sampling." Occasionally, special selection of the configuration may reduce the proposition to trivialities. Consider the following theorem: The midpoints of the sides of a triangle and the feet of altitudes lie on a common circle. (This circle is called the "nine-point circle" for the triangle. The circle contains numerous other special points of interest and the relevant mathematical theory dates from the synthetic geometry of the early 1800s.) Place the triangle as indicated and think of $a, b, d$ as fixed while $c$ is variable. (See Figure 4.) Since the general equation of a circle is $A(x^2 + y^2) + Bx + Cy + D = 0$, it follows that

[2] The decision problem for elementary polynomial algebra can be answered in the affirmative. It is one of the ironies of mathematical exposition that if a proof is mere routine, then intellectually it isn't worth going through

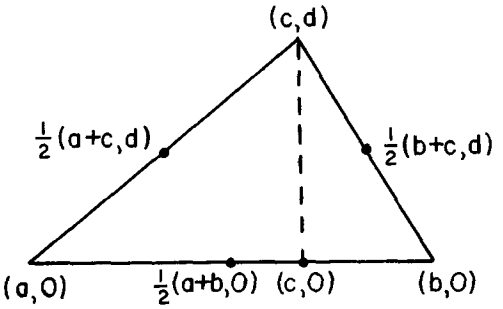


FIG 4

the necessary and sufficient condition for four points $(x_i, y_i)$, $i = 1, 2, 3, 4$, to lie on a common circle is that

$$D = \det(x_i^2 + y_i^2, x_i, y_i, 1) = 0.$$

Inserting $(x_1, y_1) = (c, 0)$, $(x_2, y_2) = ((a + c)/2, d/2)$, $(x_3, y_3) = ((b + c)/2, d/2)$, $(x_4, y_4) = ((a + b)/2, 0)$, we find that $D = D(c)$ is at most a cubic in $c$. Hence, four verifications in $c$ suffice to establish the proposition. Select $c = a$, $c = (a + b)/2$, $c = b$, $c = 2b - a$. The proposition in each of the four special cases pictured in Figure 5 is visually apparent.

4. *Completeness as a Principle of Reduction*

The uniqueness theorem for polynomials which we have used as a principle of reduction or of sampling can be written as $p(x_i) = 0$, $i = 1, 2, \ldots, N$, implies $p \equiv 0$. Within the context of the theory of linear spaces, this is a so-called *completeness* property, see e.g. Davis [4]. Let $X$ be a linear space and let $X^*$ be its conjugate space (i.e. the space of linear functionals defined over $X$). Then a set of elements $\{\phi_i\}$, $\phi_i \in X^*$, is called complete in $X^*$ if, when $x \in X$, $\phi_i(x) = 0$ for all $i$ implies $x = 0$. (F. Deutsch has expressed this condition picturesquely: He says that if $x$ is dead when tested by a complete set $\phi_i$, then it is really dead.) Thus a complete set of functionals serves as a reducer for the elements of $X$. It is a *test set.*

If $X$ has finite dimension $N$, then there is a complete set of $N$ elements. If $X$ is a normed linear space, completeness of $\{\phi_i\}$ is related to *closure*, which asserts the possibility of approximating elements of $X^*$ by finite combinations of elements of $\{\phi_i\}$.

*Example* 1 Let $X$ be the space of all polynomials $p(x) = \sum_{i=0}^{n} a_i x^i$ of degree less than or equal to $n$. This is of dimension $n + 1$. If $x_i$ are distinct, the point evaluations $\phi_i(p) = p(x_i)$, $i = 1, 2, \ldots, n + 1$, are complete.

*Example* 2. With the same space as in Example 1, the derivatives $\phi_i(p) = (1/i!)\, p^{(i)}(0)$, $i = 0, 1, \ldots, n$, form a complete set.

Similar examples hold in several variables. Note that whereas the functionals of Example 2 correspond in the context of this paper to proof by formal manipulation of coefficients $(\phi_i(a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n) = a_i)$, the functionals of Example 1 correspond to proof by numerical computation.

*Example* 3. Let $X$ consist of all trigonometric polynomials of the form $t(x) = a_0 + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx)$. Then the $2n + 1$ functionals

$$\phi_k(t) = \int_0^{2\pi} \begin{matrix} \cos kx \\ \sin kx \end{matrix} t(x)dx, \quad k = 0, 1, \ldots, n$$

constitute a complete set and similarly for algebraic polynomials using moments or coefficients in orthogonal expansions.

*Example* 4. Numerous examples can be based on Tschebyscheff systems. These include generalized polynomials $\sum_{i=1}^{n} a_i x^{\lambda_i}$, $\lambda_1 < \lambda_2 < \ldots$ and exponential polynomials $\sum_{i=1}^{n} a_i e^{\lambda_i x}$, $\lambda_1 < \lambda_2 < \ldots$ (see Karlin and Studden [11, p. 9]).

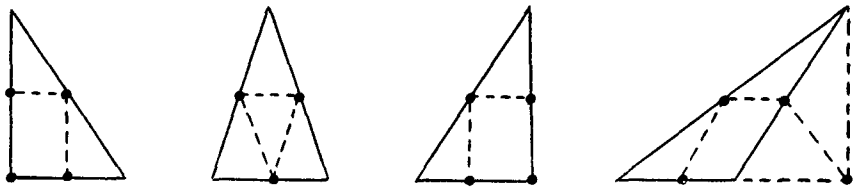

FIG. 5

*Example* 5. Let $X$ designate all the functions of a complex variable $f(z) = \sum_{n=0}^{\infty} a_n z^n$ that are all analytic in $|z| < r$ for some fixed $r$. Then, if $|z_i| < r$ and $\lim_{i\to\infty} z_i = 0$, $\phi_i(f) = f(z_i)$ form a complete set of functionals. This is the *uniqueness theorem* for analytic functions. A similar set is complete for analytic functions of several complex variables. The "coefficient" functionals $\phi_i(f) = f^{(i)}(0)/i!$, $i = 0, 1, \ldots$, are of course complete. The space $X$ is of course infinite dimensional, and verification of $f \equiv 0$ can be reduced to a countable set of individual verifications.

## 5. *Reduction in the Presence of Noise*; $\{\epsilon_n\}$-*Completeness*

Let us suppose that the tests $\{\phi_i\}$ cannot be carried out with absolute fidelity but always take place in the presence of noise. We are therefore led to consider the possibility of $|\phi_i(x)| \leq \epsilon_i$, all $i$, $\epsilon_i$ small, implying that $x = 0$. What can functional analysis tell us about this?

If the number of tests is finite, say $\phi_1, \ldots, \phi_n$, then, strictly speaking, the conditions $|\phi_i(x)| \leq \epsilon_i$ can never imply that $x = 0$. Consider for example polynomials $p$ in one real variable of degree less than or equal to $n$, and $n + 1$ tests $\phi_i(p) = p(x_i)$ where $x_i$, $i = 0, 1, \ldots, n$, are $n + 1$ distinct points. If $l_i(x)$ designate the fundamental Lagrange polynomials corresponding to $x_0, \ldots, x_n$, i.e. if $l_i(x)$ are of degree $n$ and $\phi_i(l_j) = \delta_{ij}$, $i, j = 0, 1, \ldots, n$, then we may write $p(x) = \sum_{i=0}^{n} \phi_i(p) l_i(x)$; so if $\phi_i(p)$ are prescribed in advance as $\epsilon_i \neq 0$, one has here a nontrivial polynomial satisfying the conditions.

On the other hand, letting $\| \ \|$ designate any norm, we have for any such $p$, $\|p\| \leq \sum_{i=0}^{n} |\phi_i(p)| \|l_i\| \leq \sum_{i=0}^{n} \epsilon_i \|l_i\|$. Now the norms $\|l_i\|$ are dependent of $\epsilon_i$ (they depend only on $x_0, \ldots, x_n$). Therefore if the noise level on a complete set is small, this compels $\|p\|$ to be correspondingly small. In this way one obtains an interpretation for noisy verifications.

If, however, the number of functionals in the test set $\{\phi_i\}$ is infinite, then the conditions $|\phi_i(x)| \leq \epsilon_i$ may very well imply that $x = 0$. This is known as $\epsilon_n$-*completeness*. The theory was introduced and developed by Davis and Fan [8]. An instance of this phenomenon will be shown in Section 8.

## 6. *Can the Problem Be Reduced to One Numerical Verification?*

We have seen that if the geometric (or algebraic) problem has been reduced to the question whether $p(x_1, \ldots, x_s) \equiv 0$, the function $p$ being a polynomial, this can be verified by a finite number of numerical evaluations of $p$. Can it be reduced further to *one* numerical verification? The answer is yes, providing we extend our concept of evaluation to mean something more than the usual numerical evaluation on a finite computing machine.

To see this, let us suppose that we deal initially with a polynomial of one variable $x$ of degree less than or equal to $n$ whose coefficients are *integers* (or *rational* numbers). Suppose that such a $p(x)$ is not identically 0 and that $p(\xi) = 0$. Then by definition $\xi$ is the root of a polynomial of degree less than or equal to $n$ and hence is an algebraic number of degree less than or equal to $n$. Suppose now that $\xi$ is selected to be an algebraic number of degree greater than $n$ or even as a transcendental number such as $e$ or $\pi$. Then the single equation $p(\xi) = 0$ implies $p \equiv 0$. Thus, reduction can take merely one "numerical"

verification. Now let $p(x, y)$ be a polynomial of degree less than or equal to $m$ in $x$ and less than or equal to $n$ in $y$ and suppose that $p(x, y)$ has integer (or rational) coefficients. We may write

$$p(x, y) = a_0(x) + a_1(x)y + \cdots + a_n(x)y^n,$$

where each coefficient $a_i(x)$ is a polynomial of degree less than or equal to $m$ in $x$ with integer or rational coefficients. Select $\xi$ to be a number which is not algebraic of degree less than or equal to $m$. Since the set of polynomials with rational coefficients is countable, it follows that the set of all possible values of $a_0(\xi), a_1(\xi), \ldots, a_n(\xi)$ as $a_i$ runs over the polynomials with rational coefficients is countable. For $\xi$ fixed, there are only a finite number of values $y$ such that $a_0(\xi) + a_1(\xi)y + \cdots + a_n(\xi)y^n = 0$. Hence, as $a_i$ runs over the rational polynomials, there are only a countable number of values $y$ satisfying this equation. It is clearly possible to select a number $\eta$ which is not one of these. Therefore, $p(\xi, \eta) = 0$ implies $a_0(\xi) = 0, a_1(\xi) = 0, \ldots, a_n(\xi) = 0$, and each of these implies in turn that $a_0(x) \equiv 0, a_1(x) \equiv 0, \ldots, a_n(x) \equiv 0$. Thus $p(\xi, \eta) = 0$ implies that $p \equiv 0$.

This argument is valid for polynomials in $s$ variables, and we can sum it up by saying that it is possible to find *algebraically independent* numbers and these numbers can be utilized as a reduction principle. (See Supplementary Note 4.)

We postpone to a later section our discussion of the meaning of this result as far as computation is concerned.

## 7. *Transcendentality*

We shall next sketch a theory which allows complete proof "in one test."

Let $\mathcal{P}$ be a set of functions, programs, or processes $P$ with the following features:

(a) The processes $P$ all have a common domain $D$ of inputs $x$.

(b) To each input $x$ in $D$ there is determined a unique output $P(x)$ lying in a range $R$ of outputs.

(c) A "zero" output in $R$ has been distinguished and is designated by 0.

The process $P$ is said to be identically 0 ($P = 0$) if $P(x) = 0$ for all $x$ in $D$.

(d) $\mathcal{P}$ contains the zero process.

Let $\phi$ designate a "testing functional." The functional $\phi$ will map elements of $\mathcal{P}$ into $R$. Very often, but not always, $\phi$ will be a *point evaluation*. That is, $\phi(P) = P(\alpha)$ for some fixed $\alpha$ in $D$ and all $P$ in $\mathcal{P}$.

*Definition.* The functional $\phi$ is said to be *transcendental* over $\mathcal{P}$ if $\phi(P) = 0$ implies that $P = 0$. In other words, $\phi(P) = 0$ implies $P(x) = 0$ for all $x$ in $D$. In the case where $\phi$ is a point evaluation $\phi(P) = P(\alpha)$, this reduces to $P(\alpha) = 0$ implies $P(x) = 0$ for all $x$. The appropriateness of the word "transcendental" will be seen from Example 2 below.

*Example* 1. Let $D$ designate the real numbers. Let $\mathcal{P}$ designate the set of linear functions $P(x) = \sigma x$, $\sigma$ real. Then the functional $\phi(P) = P(\alpha)$ is transcendental over $\mathcal{P}$ if $\alpha$ is any nonzero real number since $\phi(P) = 0$ means that $P(\alpha) = 0$. But $P(\alpha) = \alpha\sigma$. Hence $\sigma = 0$. Hence $P = 0$.

*Example* 2. Let $D$ designate the set of real numbers. Let $\mathcal{P}$ be the set of all polynomials $P$ (of unlimited degree) with *rational* coefficients. Then, if $\alpha$ is a transcendental number in the usual sense (for example $\alpha = 2.718\ldots$ or $\alpha = 3.14159\ldots$), the functional $\sigma(P) = p(\alpha)$ is transcendental over $\mathcal{P}$ since $\phi(P) = 0$ implies $P(\alpha) = 0$. This would force $\alpha$ to be the root of a polynomial equation with rational coefficients. This contravenes the usual definition of transcendental numbers.

*Example* 3. Let $D$ designate the set of column vectors $x$ with two real components $x = \binom{x_1}{x_2}$. Let $\mathcal{P}$ designate the set of all $2\times 2$ matrices $P$ with *rational* entries: $P = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}$. Let $P(x)$ mean the matrix product $Px$. Then the functional $\phi(P) = P\binom{1}{\sqrt{2}}$ is transcendental over $\mathcal{P}$ since $\phi(P) = 0$ means that $P\binom{1}{\sqrt{2}} = \binom{0}{0}$, or $p_{11} + p_{12}\sqrt{2} = 0$, $p_{21} + p_{22}\sqrt{2} = 0$. Hence, since $p_{ij}$ are rational, $p_{ij} = 0$ and $P = 0$.

This of course may be extended to $n \times n$ matrices. Any vector whose components are linearly independent over the rationals will serve to define the transcendental functional.

*Example* 4. Let $D$ consist of the set of continuously differentiable real functions $x(t)$ defined on some $a \leq t \leq b$. Let $\mathcal{P}$ consist of all operators of the form $P(x) = \sigma(dx/dt)$, $\sigma$ real. Let $\alpha(t)$ be any function in $D$ that is not identically constant. Then $\phi(P) = P(\alpha) = \sigma(d\alpha/dt)$ is transcendental over $\mathcal{P}$ since $\phi(P) = 0$ implies $\phi(d\alpha/dt) \equiv 0$. Now if $\sigma \neq 0$, then $(d\alpha/dt) = 0$, so that $\alpha(t) =$ constant. This is impossible; so $\sigma = 0$.

*Example* 5. This may be extended. Let $D$ consist of all infinitely differentiable functions on $0 < a < t < b$. Let $\mathcal{P}$ consist of all differential operators of the form

$$P(x) = P(t, x, x', x'', \ldots, x^{(k)}),$$

$k = 1, 2, \ldots$, where $P$ is a nontrivial polynomial in the indicated variables. Let $\Gamma(t)$ be the classical Eulerian Gamma function. Then the functional $\phi(P) = P(\Gamma(t))$ is a transcendental over $\mathcal{P}$. This is true because it is known that $\Gamma(t)$ does not satisfy any such algebraic differential equation (Holder-Ostrowski theorem). Such functions have been called *transcendentally transcendental.* (Ordinary transcendental functions such as $e^t$, $\sin t$, etc., satisfy algebraic differential equations.)

*Example* 6. Here is an example which is not analytical. Let $V^*$ consist of all finite strings of symbols where the individual symbols have been selected from an alphabet of symbols $V$. Let $\mathcal{P}$ consist of the processes $P_0, P_1, \ldots, P_\infty$, where the $P_i$ operate as follows. For any $x$, $P_\infty x =$ the null string $= 0$. For $n$ finite, $P_n(x)$ is the string that results by deleting the first $n$ elements of $x$. If $n$ is greater than or equal to the number of elements in the string $x$, then $P_n(x) = 0$.

This system has no point evaluations that are transcendental. Suppose that $\phi(P) = P(x)$. Suppose further that $P_k(x) = 0$; then it follows only that $k$ is greater than or equal to the number of symbols in the string $x$. However, if $V^*$ is enlarged to contain infinite strings of symbols and if $\alpha$ is any infinite string, then $\phi(P) = P(\alpha)$ is transcendental. This is true because $\phi(P) = 0$ means that $P(\alpha)$ is the null string. But $P_0, P_1, \ldots$ can delete only a finite number of symbols. Therefore $\phi(P) = 0$ implies $P = P_\infty = 0$.

## 8. *Transcendentals Generated by Functional Norms*

One of the basic properties of a norm in a normed space $X$ with elements $x$ is that $\| x \| = 0$ implies $x = 0$. This leads us immediately to many examples of transcendental functionals.

Consider for example the normed linear space $\mathcal{P}$ consisting of all continuous functions of a real variable $x(t)$ defined on a fixed interval $D : a \leq t \leq b$. Define $\| x \|^2 = \int_a^b |x(t)|^2 dt$. Then $\phi(x) = \int_a^b |x(t)|^2 dt$ is transcendental over $\mathcal{P}$. There is no point evaluation over $\mathcal{P}$ which coincides with $\phi$. However, $\phi$ may be approximated by point functionals in many ways. For example let $t_1, t_2, \ldots$ be any sequence of points which is equidistributed in $[a, b]$ (see e.g. Davis and Rabinowitz [9, p. 298]). Then

$$\phi(x) = \int_a^b |x(t)|^2 dt = \lim_{N \to \infty} (1/N) \sum_{k=1}^{N} |x(t_k)|^2, \quad \text{for all } x \in \mathcal{P}.$$

The set of points $\{t_k\}$ must be everywhere dense in $[a, b]$; so the point functionals $\phi_k(x) = x(t_k)$, $k = 1, 2, \ldots$, form a complete set over $\mathcal{P}$. But there is more than this. Suppose that $\eta_i > 0$ and $\lim_{i \to \infty} \eta_i = 0$. Then $(1/N) \sum_{k=1}^{N} |x(t_k)|^2 < \eta_N$, $N = 1, 2, \ldots$, implies that $\int_a^b |x(t)|^2 dt = 0$ and hence $x(t) \equiv 0$. These inequalities can be written as $\sum_{k=1}^{N} |x(t_k)|^2 < \eta_N N$, $N = 1, 2, \ldots$. Hence, if constants $\epsilon_k$ are taken so small that

$$\sum_{k=1}^{N} \epsilon_k^2 < \eta_N N, \qquad N = 1, 2, \ldots$$

(for example $0 < \epsilon_k \leq 1/\sqrt{k}$), then $\phi_k(x) = x(t_k)$ provides an illustration of a sequence of functionals that are $\epsilon_k$-*complete*.

## 9. *Transcendentals and Computation*

Let us return to the question of the verification of a theorem by means of a single

numerical evaluation. In the Pappus theorem (and similar theorems) the polynomials in question have rational coefficients; we therefore know it is possible to prove the theorem by the following strategy: Verify the theorem in the case of a configuration with transcendental (algebraically independent) coordinates; this will then imply the truth of the theorem for all configurations

Brushing under the rug the difficult theoretical question of which explicit sets of numbers are algebraically independent,[3] we observe next that even when we are provided with such a set, we must then compute with them. Such numbers are infinite nonrecurring decimals, and we must compute with them to infinite precision. This is manifestly impossible on real-world computing machines with finite word lengths, finite memories, and finite running times. The best one can do is to plug in finite decimal approximations. Thus, if one employed the number $e$, one might insert $e_1 = 2.71$, $e_2 = 2.718, \ldots$, etc. The theorem must check out for each of these insertions. Hence, if the ultimate goal is to verify that $P \equiv 0$ where $P$ is a polynomial in its arguments, then we have overkill here because, as seen, this decision can already be reached with only a finite number of point evaluations. If $P$ is not a polynomial but is analytic in its arguments, then it makes sense, but there is no particular virtue in dealing with a transcendental configuration.

## 10. *Transcendentals and Random Numbers*

In various parts of Monte Carlo theory, for example in the theory of numerical integration by means of sampling (see e.g. Davis and Rabinowitz [9, p. 298]), there is a gray region of useful confusion between random numbers, pseudorandom numbers, algebraically independent numbers, equidistributed numbers, and finite decimal approximations to all of these.

In the present context let us agree to replace a transcendental (algebraically independent) configuration by a random configuration, interpreting the latter uncritically as a finite but jumbled mess of digits. Our strategy then becomes: Verify the theorem in the case of jumbled coordinates; this will imply (with high probability) the truth of the theorem in general. The more verifications, the higher the probability.

*Example.* To get a feeling for how this works, return to the Pappus theorem. We shall verify by computer the truth of Pappus for the configuration $(x_1, y_1) = (13, 240)$, $(x_3, x_4) = (2, 53)$, $(x_5, y_5) = (37, 648)$; $(x_2, y_2) = (51, 1202)$, $(x_4, y_4) = (7, 190)$, $(x_6, y_6) = (23, 558)$. The first three points lie on the line $y = 17x + 19$, while the last three lie on $y = 23x + 29$.

The program was laid out as in Section 2. The computer was asked to output DET (DET = 0 means verification) as well as the expansion of DET $= D_1 - D_2 + D_3$ in terms of the minors of the first row of DET. The output for the above configuration read

$$\begin{aligned} D_1 &= -497\,152\,518\,511\,616 \\ -D_2 &= \phantom{-}396\,537\,652\,008\,960 \\ D_3 &= \phantom{-}100\,614\,866\,502\,656 \\ \hline \text{DET} &= 0 \end{aligned}$$

These are exact 15-figure integers. Note the substantial build-up of the lengths of the integers starting from numbers of around 2–3 figures. An insertion of, say, 10-digit integers at the $(x_i, y_i)$ stage would compel us to go to multiple precision programming. Note also the "random" jumble of digits in $D_1$, $D_2$, $D_3$, leading all but the most hardbitten skeptic to believe that the remarkable identity $D_1 - D_2 + D_2 = 0$ is due to a general tendency and not to pure coincidence. (See Supplementary Note 5.)

We may speak of this as a *principle of scientific induction: If something is true for a random configuration of circumstances it is generally true.*

*Example.* Consider, for example, the so-called "Theorem of Napoleon," often

[3] The reader interested in pursuing this question should consult Baker [1] and Zhidlovsky [14]

ascribed to Napoleon Bonaparte (who was a good student of mathematics as a boy). Let equilateral triangles be erected on the sides of *any* triangle (see Figure 6). Then the centers of the equilateral triangles form a triangle which is also equilateral. One might very well imagine the young Napoleon fooling around with paper, pencil, ruler, and compass, and starting from a random inner triangle, discovering experimentally that $c_1$, $c_2$, $c_3$ is equilateral. This discovery would then suffice to assure the truth of the general theorem, with high probability. (See Supplementary Note 6.)

11. *The Transcendental as Pure Symbol*

We began this paper by inquiring whether symbolic algebra could be replaced by arithmetic computation. We decided that this was indeed possible. In order to push the point to extremes, we went on to conclude that if computation with transcendentals is allowed, then only one verification is needed. Thus infinite precision arithmetic with transcendentals is equivalent to finite precision arithmetic with pure symbols.

*Example.* Consider the algebraic number field consisting of numbers of the form $a + b\sqrt{2}$, where $a$ and $b$ are rational numbers. Infinite precision arithmetic can be carried out in this field by making the identification $a + b\sqrt{2} \leftrightarrow (a, b)$. Then $(a, b) + (c, d) = (a + c, b + d)$, $(a, b) \times (c, d) = (ac + 2bd, ad + bc)$ would be the appropriate rules for $+$ and $\times$. Thus $\sqrt{2}$ is functioning as the pure symbol $(0, 1)$. Within this arrangement, there is no possibility of making the identification $(0, 1) \rightarrow 1.414\ldots$.

The parallelism between transcendentals and pure symbols is reflected, for example, in the portion of algebra dealing with *field extensions*. Witness this theorem:

THEOREM. *If $\alpha$ is transcendental over a field $F$, the subfield generated by $F$ and $\alpha$ is isomorphic to the field $F(x)$ of all rational forms in an indeterminate $x$ with coefficients in $x$. The isomorphism may be chosen so that $\alpha \leftrightarrow x$ and $c \leftrightarrow c$ for each $c$ in $F$.*

If we decided to work with transcendentals to "infinite precision" by the device of introducing a pure symbol, then if we make computations with $\pi$ or $e$ (3.14159... or 2.718...), keeping the former as pure symbols, we are precisely back in the original position of operating with formal algebra. Thus this policy brings us full circle back to where we started. The practical policy of working with approximations leads us, on the other hand, to heuristics, probabilities, and analogies between transcendental numbers and "random" numbers.

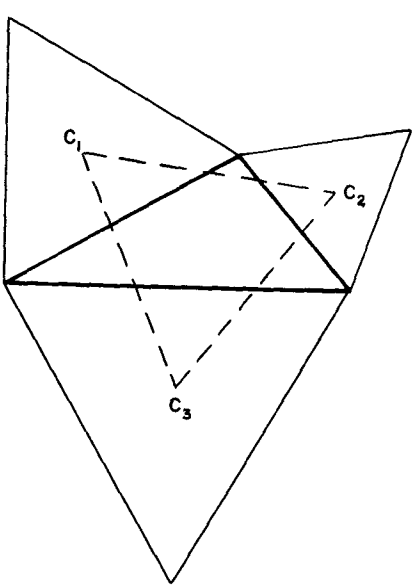


FIG 6

12. *Enormous Integers as Symbols, Programs, and Transcendentals*

Let us consider first some specific integers whose size is ridiculously large. As the numbers which appear in physics or cosmology appear to lie well within the range $10^{\pm 500}$, we shall seek integers far above this limit of physical significance. There is $A = 2^{4423} - 1 \approx 10^{1331}$ which at one time held the record as the largest known prime. The Gödel numbers which play such an important role in the theory of provability are a rich source of integers that are beyond the limit of physical significance. There is Skewes number $S = 10^{10^{10^{34}}}$ which once played a role in the theory of prime number distribution.

If one wants to generate large numbers rapidly, a fine device to employ is the iterative scheme stressed by Hugo Steinhaus and Leo Moser. Let $f_1(a) = a^a$. For $n = 1, 2, \ldots$, let $f_{n+1}(a) = f_n \circ \overset{(a)}{\cdots} \circ f_n(a)$, where the symbol on the right designates $a$ functional compositions. The integer $M = f_3(2)$ is known as the *mega* and $M^* = f_M(2)$ is known as the *moser*. The mind boggles at the size of the mega (the reader would do well to try to compute $f_3(2)$) and, as for the moser, it is out of sight.

Now there appears to be a vital difference between integers such as 7, 211, or even $A$ (which has only 1331 digits) and integers such as $S$, $M$, $M^*$, etc., in that one can physically perform arithmetic at the primitive level with the former, but one cannot perform it with the latter. Thus one can find $7 \times 211$ or $17 \times A$ and express the result as a decimal integer. However, if one asks for $211 + M$ or $53 \times (S + M + M^*)$, then this is not possible physically.[4] We must leave it at the symbolic level. Of course, we may replace $53 \times (S + M + M^*)$ by $(53 \times S) + (53 \times M) + (53 \times M^*)$, but then we are manipulating $S$, $M$, $M^*$ as mere symbols. We cannot verify this identity or even $M + M^* = M^* + M$, though we might assert their truth on the grounds of general properties shared by all integers. Let us say that integers of the former type are integers in esse; they can be written down. The integers of the latter type are integers in posse; they cannot be written down. One cannot draw a dividing line between the two types; perhaps the notion of fuzzy sets might be useful here. An integer in posse such as $S$ or $M$ or $M^*$ is really a program, and it is a program that can be stated in a small number of lines in some computer language.[5]

Suppose now, returning to the numerical verification of the Pappus theorem, that one were to specify six points on two lines whose coordinates are all integers in posse. Then we could not verify the critical equation DET = 0 by direct computation at the primitive level. What one would do is to observe that inasmuch as DET as a function of its atomic symbols is identically 0, then DET must also be zero when integers in posse are substituted into it.

Thus we arrive at the vague idea that certain statements about integers in posse can be true if and only if the statements are identically true. As we have seen, this is the characteristic behavior of the transcendental.

Some of this vagueness may be dispelled by constructing a mathematical model. Let $\Omega(+, \cdot, Z, \mathcal{P})$ be a structure built up in the following way. Let $Z$ designate a fixed subset of the set of all positive, negative, and zero integers. The elements of $Z$ will represent the integers in esse. Let $\mathcal{P}$ be a set of symbols (finite or infinite in number) which are to play the role of programs or integers in posse. The following rules are observed:

(1) $0 \in Z$. $0 \notin \mathcal{P}$.

(2) If $n \in Z$ and $P \in \mathcal{P}$ then $n + P$ and $P + n \in \mathcal{P}$.

(3) If $n \in Z$ and $n \neq 0$ and $P \in \mathcal{P}$, then $nP \in \mathcal{P}$.

(4) If $P \in \mathcal{P}$, then $0 \cdot P = P \cdot 0 = 0$.

(5) If $P_1$ and $P_2 \in \mathcal{P}$, then $P_1 + P_2$ and $P_1 P_2 \in \mathcal{P}$.

[4] Even the most hardbitten Platonist among mathematicians must at some point deal with the limitations imposed by the physical world For example, mathematical proofs and lectures are all required to be of finite length, and by finite one means in practice something that can be expressed by, say, ten volumes of symbols. One wonders what status therefore could be given a proof containing a mega of symbols

[5] To generate $f_n(2)$ in APL requires no more than 8 statements. One works by writing a program which calls itself. Of course, if one tried to run the program for $f_3(2)$, it would never output in the lifetime of the programmer.

Note that the behavior of this system imitates the behavior of the transfinite cardinal $\aleph_0$, insofar as $n + \aleph_0 = \aleph_0$, $\aleph_0 + n = \aleph_0$, $\aleph_0 + \aleph_0 = \aleph_0$, $\aleph_0 \cdot \aleph_0 = \aleph_0$.

We next consider polynomials $\pi(x) = a_0x^n + a_1x^{n-1} + \cdots + a_n$, where $a_i$ and $x \in \Omega$. If now the coefficients are restricted to lie in $Z$, then any $P \in \mathcal{P}$ is a transcendental over this set of polynomials. For suppose $\pi(P) = a_0P^n + a_1P^{n-1} + \cdots + a_{n-1}P + a_n = 0$. Suppose that for some $k \neq n$, $a_k \neq 0$. Then $a_kP^{n-k} \in \mathcal{P}$. Then, by the above rules, $\pi(P) \in \mathcal{P}$ and hence cannot be 0. Therefore $a_0 = a_1 = \ldots = a_{n-1} = 0$. Hence also $a_n = 0$. Thus $\pi(P) = 0$ implies $\pi = 0$. (See Supplementary Note 7.)

*Supplementary Notes*

1 The distinction, if any, between numbers and symbols (variables) may well be disputed. Certainly in the usual languages this distinction is preserved at the compile level by the use of attributes by which the identifier is declared to be a number (in one of a variety of forms), a string, a label, a variable, a logical constant, etc

2. For a discussion of polynomial evaluation via the division algorithm and its relation to the fast Fourier transform, see Fiduccia [10]

3 The identity in 16 variables is ascribed variously to Cayley and to Brioschi It reads as follows

$$\sum_{i=0}^{7} X_i^2 \sum_{i=0}^{7} Y_i^2 = \sum_{i=0}^{7} Z_i^2,$$

where

$$Z_0 = \sum_{i=0}^{7} X_iY_i,$$
$$Z_i = X_0Y_i - Y_0X_i + X_{i+1}Y_{i+5} - X_{i+5}Y_{i+1} + X_{i+2}Y_{i+3} - X_{i+3}Y_{i+2} + X_{i+4}Y_{i+6} - X_{i+6}Y_{i+4}.$$

The indices are reduced mod 7 An interpretation for this identity can be found within the theory of Cayley algebras or within the theory of spinors (see, e g Cartan [2, p. 121]). One of the principal theorems in this area is that no further identities of this type are possible

4 There are other ways of arranging for a reduction to "one" numerical verification The one presented here fits in best as an analogue of sampling

The point here is not to find complicated proofs of simple theorems, but to focus attention on the properties and implications of a certain kind of methodology.

5 One can argue that the presence of an unusual numerical circumstance is indicative of an underlying theory which explains it Consider for example the amusing instance presented in Martin Gardner's column in the April 1975 issue of *Scientific American*.

The number of $G = e^{\pi\sqrt{163}}$ is conjectured to be an integer Now, a multiple precision computation shows that $G$ consists of 18 digits to the left of the decimal point, followed by 12 9s to the right of the decimal point. The 13th digit to the right of the decimal point is 2, spoiling the conjecture

An explanation of the coincidence of 12 9s can be given within the theory of class number of quadratic number fields and expansions in terms of the Klein absolute modular invariant Historically this theory dates from the late 1890s and *preceded* any high accuracy computation of $G$. The number $G$ is known to be transcendental as a result of the Gelfand-Schneider theorem (early 1930s).

Returning to the example of the text, one may very well ask. What really does this computation prove? Does it prove (a) the truth of Pappus's theorem (already known since 350 A D ) or only (b) that the hardware/software combination is doing what it is supposed to be doing (presumably the manufacturer has checked this out) and (c) that the Pappus program I wrote is correct (it can't really be proved correct; only that it is not demonstrated to be incorrect)?

According to the view which stresses an experimental or experiential criterion of mathematical truth (see Davis [5, 6]), it does all of these things simultaneously.

Diagnostic programming is frequently carried out by the insertion of randomized inputs

6 Here is a further example of "sampling" on a "transcendental" element, a way of predicting the presidental election: There is a certain congressional district in the country which has gone with the winner for about a century Merely sample this district On election night, reporters always report on this district. Of course, we might even go further and look for a "transcendental" household and accept its verdict. This type of sampling has been called "purposive" selection and is not generally recommended

7 The finiteness of enormous integers such as $S$, $M$, $M^*$, etc., has been questioned by numerous authors. The list includes Borel, Fréchet, Mannoury, Rieger, and van Dantzig (see Kino, Myhill, and Vesley [12, p. 4]). Yessenin-Volpin [12] uses the expression "feasible" number to express a meaning that comes close to our numbers "in esse "

REFERENCES

(Note Reference [13] is not cited in the text )

1. Baker, A Linear forms in the logs of algebraic numbers *Matematik 13* (1966), 204–216.

2. CARTAN, E *Theory of Spinors* M.I.T. Press, Cambridge, Mass., 1966.
3 CURTISS, C.W The four and eight square problem and division algebras. In *Studies in Modern Algebra, Vol. 2*, A.A. Albert, Ed , Math. Assn Amer , Washington, D.C., 1963.
4. DAVIS, P.J. *Interpolation and Approximation*. Blaisdell, Waltham, Mass., 1963; reprint, Dover, New York, 1975
5. DAVIS, P.J. Fidelity in mathematical discourse. *Amer Math. Monthly 79* (1972), 252–263
6 DAVIS, P.J. Visual geometry, computer graphics, and theorems of perceived type In *The Influence of Computing on Mathematical Research and Education*, Proc Symp Applied Math , Vol 20, Amer Math. Soc., Providence, R I , 1974, pp 113–127
7 DAVIS, P J., AND CERUTTI, E FORMAC meets Pappus *Amer Math Monthly 76* (1969), 895–905
8. DAVIS, P J , AND FAN, K Complete sequences and approximation in normed linear spaces *Duke Math J 24* (1957), 183–192.
9 DAVIS, P.J AND RABINOWITZ, P *Methods of Numerical Integration*. Academic Press, New York, 1975
10 FIDUCCIA, C.M Polynomial evaluation via the division algorithm· The fast Fourier transform revisited Proc. Fourth Annual ACM Symp on Theory of Comptg., 1972, pp. 88–93.
11 KARLIN, S , AND STUDDEN, W J *Tschebyscheff Systems*. Wiley-Interscience, New York, 1966.
12 KINO, A., MYHILL, J , AND VESLEY, R.E , Eds *Intuitionism and Proof Theory*. North-Holland Pub Co , Amsterdam, 1970
13. PÓLYA, G *Mathematics and Plausible Reasoning* Princeton U Press, Princeton, N J , 1954
14 ZHIDLOVSKY, A.B On the transcendence and algebraic independence of the values of certain *E*-functions. Math Mech , Vestnik Moscow U , Ser 1, 1961, pp 44–59

RECEIVED JULY 1975, REVISED JULY 1976