Hilbert's Proof of his Irreducibility Theorem

Mark B. Villarino, William Gasarch and Kenneth W. Regan

Abstract. Hilbert's Irreducibility Theorem is a cornerstone that joins areas of analysis and number theory. Both the genesis and genius of its proof involved combining real analysis and combinatorics. We try to expose the motivations that led Hilbert to this synthesis. His famous Cube Lemma anchored the proof but without the analytical foundation and framework it would have had no purpose. We also assess this lemma as a precursor of Ramsey Theory.

1. INTRODUCTION. In 1892, David Hilbert published what is known today as *Hilbert's irreducibility theorem*. We give his statement, using *integral polynomial* to mean a polynomial in any number of variables whose coefficients are integers.

Theorem 1. *If $F(x, y, \ldots, w; t, r, \ldots, q)$ is an irreducible polynomial with integral coefficients in the variables $x, y, \ldots, w$ and the parameters $t, r, \ldots, q$, then it is always possible, and indeed in infinitely many ways, to substitute integers for the parameters $t, r, \ldots, q$ such that the polynomial $F(x, y, \ldots, w; t, r, \ldots, q)$ becomes an irreducible polynomial in the variables $x, y, \ldots, w$ alone.*

This is a direct translation from [**14**]. The statement falls short of modern standards. For example, the irreducibility of F concerns the polynomial in the whole set of variables—parameters included—but the statement is technically false if there are no variables but only parameters. Nor is it clear whether one needs a lot of variables and parameters or whether proving this for one or two of each suffices; this is clarified below. One of our purposes is to lead readers to appreciate modern rigor and clarity compared to 19th-century standards.

To Hilbert, this theorem was not an end in itself but rather a tool to use for some remarkable applications. A simple one is that if a polynomial $f(x)$ over $\mathbb{Z}$ has values that are perfect squares for all sufficiently large x, then $f(x)$ must be the square of some other polynomial over $\mathbb{Z}$. One of his most striking is:

Corollary 2. *For every integer n there exist infinitely many polynomials p in $\mathbb{Z}[x]$ such that p has the symmetric group S_n as its Galois group.*

He began his paper [**14**] with a statement and proof of the two-variable case, which is the fundamental step in the proof of the general theorem. Again in Hilbert's own words, the statement is:

Theorem 3. *If $f(x, t)$ is an irreducible polynomial in the two variables x and t with integral coefficients*

$$f(x, t) = Tx^n + T_1x^{n-1} + \cdots + T_n, \tag{1}$$

where $T, T_1, \ldots, T_n$ are integral polynomials in t, it is always possible, indeed in infinitely many ways, to substitute an integer for t in $f(x, t)$ such that the polynomial $f(x, t)$ becomes an irreducible polynomial of the single variable x.

A more modern formulation is:

Theorem 4. *Let $f(x, y) \in \mathbb{Z}[x, y]$ be irreducible. For an infinite number of $t \in \mathbb{Z}$, $f(x, t)$, as an element of $\mathbb{Z}[x]$, is irreducible.*

In fact, Hilbert proved the *contrapositive*, which can be formulated as follows.

Theorem 5. *Let $f(x, y) \in \mathbb{Z}[x, y]$. If there exists t_0 such that for every integer $t \geq t_0$, $f(x, t)$ is reducible over $\mathbb{Z}$, then $f(x, y)$ is reducible over $\mathbb{Z}$.*

Hilbert proved this by formulating what is today called *Hilbert's Cube Lemma*. It can be viewed not only as an enhanced form of Dirichlet's pigeonhole principle but also as the first statement of a Ramsey-type theorem.

In Section 2 we discuss Ramsey Theory to illustrate why Hilbert's cube lemma is regarded as belonging to that field. In Section 3 we state and give a simple modern proof of the Hilbert's cube lemma (we discuss optimizations and Hilbert's original proof in Section 13). It is easy to appraise the Hilbert cube lemma as a gem in an isolated setting and forget the quest that led to it, which was was *to find a polynomial factor, $\varphi(x, t) \in \mathbb{Z}[x, t]$, of the polynomial $f(x, t)$*.

In Sections 4 through 11 we provide a motivated account of Hilbert's beautiful proof of (ir)reducibility by putting ourselves in his shoes and following the trail of ideas that we find in his 1892 paper. We have tried to make it as self-contained and elementary as possible.

After Hilbert, many mathematicians offered other proofs of the irreducibility theorem. Many of these proofs use so-called "density" arguments, a standard technique in today's Diophantine approximation theory, but a far cry from the natural idea of Hilbert to find a factor of a reducible polynomial. We will say more about modern proofs in Section 12.

Hilbert remains one of the greatest mathematicians of all time. His original proof still contains insights and arguments that are well worth study even today. We offer the reader a detailed exposition of this proof in hope of saving it from the oblivion of history.

2. RAMSEY THEORY.

Theorems in Ramsey Theory almost always follow this informally-stated pattern:

> *For any coloring of a large enough object there is a nice monochromatic sub-object.*

We give three examples of such theorems along with some of the history. See [**10, 16, 21**] for more on these theorems and also Alexander Soifer's book [**27**] for more of the history.

In 1916, Issai Schur [**24**] proved the following:

Lemma 6. *For all c there exists $S = S(c)$ such that for all c-colorings of $\{1, \ldots, S\}$ there exists a monochromatic triple x, y, z such that $x + y = z$.*

Schur viewed his lemma as a means to an end and so did not launch what is now called Ramsey Theory. He used it to prove the following theorem in number theory.

Theorem 7. *Let $n \geq 1$. There exists q such that, for all primes $p \geq q$, there exists $x, y, z \in \{1, \ldots, p-1\}$ such that $x^n + y^n \equiv z^n \pmod{p}$.*

Proof. Given n let $q = S(n)$. Let p be a prime such that $p \geq S(n)$. Then $\mathbb{Z}_p^*$, which denotes the numbers $\{1, \ldots, p-1\}$ together with the operation of modular multiplication, forms a group. All arithmetic henceforth is in $\mathbb{Z}_p^*$.

Let $H = \{x^n \mid x \in \mathbb{Z}_p^*\}$. Clearly H is a subgroup of $\mathbb{Z}_p^*$. It is known that $|H| = \frac{p-1}{\gcd(n,p-1)}$ so the number of cosets is $c = \frac{p-1}{|H|} = \gcd(n, p-1) \le n$. We denote the cosets by $d_1 H, \ldots, d_c H$.

Consider the following c-coloring of $\{1, \ldots, p-1\}$: color x by i such that $x \in d_i H$. Since $c \le n$ and $p - 1 \ge S(n)$, by Schur's Lemma, there exists a monochromatic x_1, y_1, z_1 such that $x_1 + y_1 = z_1$. Since they are all in he same coset there exists d such that $x_1, y_1, z_1 \in dH$. Hence $x_1 = dx^n$, $y_1 = dy^n$, $z_1 = dz^n$. Since $dx^n + dy^n = dz^n$ we get $x^n + y^n = z^n$. ■

Theorem 7 refuted the idea of proving Fermat's last theorem by showing that for all $n \ge 3$ there are arbitrarily large p such that $x^n + y^n \equiv z^n$ has no solution modulo p.

In 1927, Bartel van der Waerden [**31**] proved the following theorem which now bears his name:

Theorem 8. *For all k, c there exists a number $W = W(k, c)$. such that for all c-colorings $\{1, \ldots, W\}$ there there exists a monochromatic arithmetic sequence of length k.*

The title of [**31**] credits Pierre Baudet with having conjectured this, but Soifer [**27**] gives evidence that Schur had also done so. Even though van der Waerden did not have another goal in mind, he did not pursue this line of research and so did not launch what is now called Ramsey Theory.

Frank Ramsey [**22**] proved the following theorem which now bears his name. As others often do, we state only the case for graphs, not hypergraphs. A *graph* consists of a set V of *vertices* and a set $E \subseteq V \times V$ of *edges*. We consider only *undirected* graphs, in which edges are unordered pairs (x, y), without *self-loops*, meaning always $x \ne y$. The graph is *complete* if E includes all such pairs and is then denoted by K_n, $n = |V|$. A *c-coloring* of the edges is a mapping f from E to $\{1, \ldots, c\}$. A *monochromatic* K_m means a subset $V' \subseteq V$ of size m and $c' \leqslant c$ such that for all distinct $u, v \in V'$, (u, v) is an edge and $f(u, v) = c'$.

Theorem 9. *For all c, m there exists a number $R = R(c, m)$ such that for all c-colorings of the edges of K_R there exists a monochromatic K_m.*

The folkloric example of this theorem is that in any group of six people, *at least three know each other or at least three are complete strangers*. If the six are the vertices of a K_6 and each edge is colored green or blue (friends or strangers), then the theorem says there is at least one monochromatic triangle. In fact, there are at least *two* such triangles, whereas K_5 has none when a green five-pointed star is inscribed in a blue pentagon, so that $R(2, 3) = 6$.

Ramsey applied his lemma to problems in mathematical logic. He viewed it as a means to an end and so did not launch what is now called Ramsey Theory.

In 1892, before all of the results above, Hilbert [**14**] proved the lemma featured in the next section. Like the three statements, it applies to any c-coloring and yields a monochromatic nice substructure. Hilbert viewed his lemma as a means to an end and so did not launch what is now called Ramsey Theory. He used it to prove the Hilbert Reducibility Theorem, which is our main topic here.

Who did launch Ramsey Theory? Speaking later about his joint paper in 1935 with George Szekeres [**5**], Paul Erdös related that it was Szekeres who rediscovered the statement and proof of Ramsey's theorem. They used it as a means to the following end:

Theorem 10. *For all $n \ge 3$ there exists $m > n$ such that for any m points in the plane in general position there exists n points that form a convex hull.*

But, they also attracted a clique of mostly Hungarians who developed the ideas, conjectures, and results that grew into Ramsey Theory as we know it.

3. THE CUBE LEMMA. Hilbert's first paragraph crisply framed the *problem* of irreducibility under substitutions represented by the statement of Theorem 1. Then he continued right away, "Our developments rest on the following lemma." We reproduce his words but change his a, μ to c, β and compact his displayed formulas using 0–1 variables $b_1, \ldots, b_m$:

> "Given an infinite integer sequence $a_1, a_2, a_3, \ldots$ in which generally each a_s denotes one of the c-many positive integers $1, 2, \ldots, c$, let m be any positive whole number. Then there are always m-many positive whole numbers $\mu^{(1)}, \mu^{(2)}, \ldots, \mu^{(m)}$ such that the 2^m elements
>
> $$a_{\beta+\sum_{i=1}^{m} b_i \mu^{(i)}}$$
>
> for infinitely many whole numbers β are collectively the same number G, where G is one of the numbers $1, 2, \ldots, c$."

Call those elements collectively the m-*cube*, which we can denote by $C(\beta; \mu_1, \ldots, \mu_m)$. The sequence $a_1, a_2, a_3, \ldots$ can be called a *coloring* of $\mathbb{N}^+$ using c colors. Thus the conclusion is that every coloring gives rise to *increments* $\mu^{(1)}, \mu^{(2)}, \ldots, \mu^{(m)}$ that yield a monochromatic m-cube for infinitely many starting points β. This is implied by the following finitistic statement, which we regard as **Hilbert's Cube Lemma** in the modern sense:

Lemma 11. *For all m, c there is a number H such that, for all c-colorings of $\mathbb{N}^+$ and all intervals of length H in $\mathbb{N}^+$, there is a monochromatic m-cube within the interval.*

Proof. The proof is by induction on m. For the base case $m = 1$, we can take $H_1 = c + 1$. This just says that for any c-coloring of an interval of length $c + 1$ there will be two elements that are the same color. Taking β to be the smaller one and $\beta + \mu_1$ the larger one, $C(\beta; \mu_1)$ is a monochromatic 1-cube.

For the induction, assume that $h = H_{m-1}$ exists. We show that, for any c-coloring of an interval of length $H_m = h\cdot(1 + c^h)$, there is a monochromatic m-cube. Let COL be a c-coloring of an interval of length H_m. Partition the interval into $1 + c^h$ blocks of size h. By the pigeonhole principle, some two of those blocks have the same sequence of h colors. By the induction hypothesis, the former has a monochromatic $(m-1)$-cube $C(\beta; \mu_1, \ldots, \mu_{m-1})$, and since the color sequence of the latter is the same, it has $C(\beta'; \mu_1, \ldots, \mu_{m-1})$ with the same color and increments but $\beta' > \beta$. Take $\mu_m = \beta' - \beta$. Then $C(\beta; \mu_1, \ldots, \mu_m)$ is the required monochromatic m-cube. ■

4. MONIC POLYNOMIALS. Hilbert begins by reducing his general problem to the case of *monic polynomials in one variable* x with *rational* coefficients. That is, he shows that the following statement suffices to prove Theorem 5.

Theorem 12. *Let $g(y, t) \in \mathbb{Z}[y, t]$. If there exists t_0 such that for all $t > t_0$, $g(y, t)$ is monic and reducible in $\mathbb{Z}[y]$, then $g(y, t)$ is reducible in $\mathbb{Q}[y, t]$.*

Proving Theorem 12 will occupy all the sections to follow, but here we show:

Proposition 13. *Theorem 12 implies Theorem 5.*

For the proof we must preface the following lemma:

Lemma 14. *Let $f(x,y) \in \mathbb{Z}[x,y]$ and $g(x,y) \in \mathbb{Z}[x,y]$ and suppose that they are arranged according to powers of x:*

$$f(x,y) = a_0(y)x^n + a_1(y)x^{n-1} + \cdots + a_{n-1}(y)x + a_n(y),$$
$$g(x,y) = b_0(y)x^n + b_1(y)x^{n-1} + \cdots + b_{n-1}(y)x + b_n(y),$$

where each a_i and b_i belongs to $\mathbb{Z}[y]$. Then:

(a) *A necessary and sufficient condition that a polynomial $\psi(y)$ be a factor of $f(x,y)$ is that it is a factor of all the polynomials $a_i(y)$.*
(b) *If $\psi(y)$ is irreducible and divides the product $f \cdot g$ then either it is a factor of all $a_i(y)$ or a factor of all $b_i(y)$.*
(c) *If $f(x,y)$ can be factored into the product of two polynomials in x whose coefficients are rational functions of y with integral coefficients, i.e., in $\mathbb{Q}(y)[x]$, then it can be factored into the product of two polynomials in $\mathbb{Z}[x,y]$.*

Proof. The intuition for statements (a) and (b) is that since x occurs nowhere else it cannot help $\psi(y)$ divide f or $f \cdot g$ any other way than stated. Bôcher [**1**, pp. 203–204] has a formal proof. To prove (c), we write the given factorization in the form

$$f(x,y) = \frac{f_1(x,y)}{\varphi_1(y)} \cdot \frac{f_2(x,y)}{\varphi_2(y)},$$

where $f_1(x,y)$, $f_2(x,y)$, $\varphi_1(y)$ and $\varphi_2(y)$ are integral polynomials such that f_1 is not divisible by any factor of $\varphi_1(y)$ and f_2 is not divisible by any factor of $\varphi_2(y)$. By part (b), since $f_1 \cdot f_2$ is divisible by $\varphi_1 \cdot \varphi_2$, f_1 has the complete polynomial φ_2 as a factor and f_2 has the complete polynomial φ_1 as a factor. By (a) we can cancel φ_2 from the coefficients of f_1 and we can cancel φ_1 from the coefficients of f_2. This gives us our factorization into two polynomials in $\mathbb{Z}[x,y]$. ■

Proof of Proposition 13. Let $f(x,y) \in \mathbb{Z}[x,y]$ and suppose we have t_0 such that for all $t > t_0$, $f(x,t)$ is reducible in $\mathbb{Z}[x]$. Recall from (1) the integral polynomials $T, T_1, \ldots, T_n$ in t such that

$$f(x,t) = Tx^n + T_1x^{n-1} + \cdots T_{n-1}x + T_n.$$

Note that T and the T_j become integer constants for any fixed value of t. Define

$$g(y,t) = y^n + S_1y^{n-1} + \cdots + S_{n-1}y + S_n,$$

where for each j, $1 \leqslant j \leqslant n$, $S_j = T_jT^{n-j}$. Then

$$g(y,t) = f(\frac{y}{T}, t).$$

Since $f(x,t)$ factors in $\mathbb{Z}[x]$, $g(y,t)$ factors in $\mathbb{Q}[y]$. But since $g(y,t)$ is an integral polynomial and is monic in the one variable y, it factors in $\mathbb{Z}[y]$ by a famous and simple lemma of Gauss. Often called *Gauss's polynomial lemma*, it states that any product of two monic polynomials over the rationals with at least one rational non-integral coefficient is itself a monic polynomial over the rationals with at least one rational non-integral coefficient.

Thus we have satisfied the hypothesis of Theorem 12—and with the same t_0 as in Theorem 5. Assuming its conclusion gives us

$$g(y,t) = \Psi(y,t)\Psi'(y,t),$$

where $\Psi(y,t)$ and $\Psi'(y,t)$ belong to $\mathbb{Q}[y,t]$. Substituting back $y = xT$ yields the following equation for our original polynomial:

$$f(x,t) = \frac{\Phi(x,t)\Phi'(x,t)}{AT^{n-1}},$$

where $\Phi(x,t)$ and $\Phi'(x,t)$ both belong to $\mathbb{Z}[x,t]$, $A \in \mathbb{Z}$, and $T \in \mathbb{Z}[t]$. Now part (c) of our lemma completes the proof. ■

Gauss used his lemma, which appeared on page 42 of his *Disquisitiones* [**6**], to give the first proof of the irreducibility of the cyclotomic polynomial of prime degree over the rationals. As we've seen, Hilbert used it to reduce the Irreducibility Theorem to the case of monic polynomials with rational coefficients. But to go further and prove Theorem 12, a new tool is needed.

5. PUISEUX SERIES. The fundamental theorem of algebra shows us that the equation $g(y,t) = 0$ has n complex roots for each value of t. Thus, informally, there are n functions of t, say $y_1(t), \ldots, y_n(t)$, which satisfy the equation. Hilbert uses a refined form of the *implicit function theorem*, which we refer to as *Puiseux's theorem* (see discussion of origins below). It says that the n root functions $y_1(t), \ldots, y_n(t)$ can be expressed in a concrete way by means of fractional power series in decreasing powers of the variable. These are the so-called *Puiseux series at infinity*, whose definition we now recall.

Definition 1. Let $m \in \mathbb{N}$. A Puiseux series at infinity is an expression of the form

$$u(x^{1/k}) + \sum_{i=1}^{\infty} \frac{\mathrm{B}_i}{x^{i/k}},$$

where $u(x) \in \mathbb{C}[x]$ is of degree m, $k \in \mathbb{N}^+$, and $\mathrm{B}_1, \mathrm{B}_2, \cdots \in \mathbb{C}$.

We adopt the following theorem statement from [**30**, pp. 80–81] with slight alterations in notation and formatting. The power series in (2) are called *Puiseux expansions at infinity*.

Theorem 15 (Puiseux's Theorem). *Given $g(y,t)$ as above, there are n distinct power series*

$$\begin{aligned}
y_1(t) &= \mathrm{A}_{11}\tau^h + \mathrm{A}_{12}\tau^{h-1} + \cdots + \mathrm{A}_{1,h+1} + \frac{\mathrm{B}_{11}}{\tau} + \frac{\mathrm{B}_{12}}{\tau^2} + \frac{\mathrm{B}_{13}}{\tau^3} + \cdots \\
y_2(t) &= \mathrm{A}_{21}\tau^h + \mathrm{A}_{22}\tau^{h-1} + \cdots + \mathrm{A}_{2,h+1} + \frac{\mathrm{B}_{21}}{\tau} + \frac{\mathrm{B}_{22}}{\tau^2} + \frac{\mathrm{B}_{23}}{\tau^3} + \cdots \\
\vdots \quad & \quad \vdots \\
y_n(t) &= \mathrm{A}_{n1}\tau^h + \mathrm{A}_{n2}\tau^{h-1} + \cdots + \mathrm{A}_{n,h+1} + \frac{\mathrm{B}_{n1}}{\tau} + \frac{\mathrm{B}_{n2}}{\tau^2} + \frac{\mathrm{B}_{n3}}{\tau^3} + \cdots
\end{aligned} \tag{2}$$

which are all convergent for t greater than some constant, where the following hold:

(a) *For a certain positive integer* k, $\tau = t^{1/k}$, *where the positive real value of the root is meant;*

(b) *The given number* h *is the highest positive exponent of* τ *that occurs.*

(c) *All coefficients* $A_{i,j}$ *and* $B_{i,j}$ *are well-defined uniquely determined complex numbers;*

(d) *Any formal power series* $y(t)$ *satisfying the formal identity* $g\{y(t), t\} \equiv 0$ *and having properties analogous to those of the series* (2) *necessarily coincides with one of the above* n *series.*

(e) *The following formal identity holds:*

$$g(y,t) \equiv \prod_{i=1}^{n} \{y - y_i(t)\}.$$

■

Hilbert ascribed the idea he used to Runge in a work that had appeared three years earlier, and cited it in a footnote exactly like this.[1] He then notes—and this is the reason to reduce the problem to monic polynomials—the relation between the elementary symmetric functions of the roots $y_1, y_2, \ldots, y_n$ and the coefficient polynomials $S_1, S_2, \cdots, S_n$, namely:

$$\begin{aligned} S_1 &= -(y_1 + y_2 + \cdots + y_n) \\ S_2 &= (-1)^2(y_1y_2 + y_1y_3 + \cdots + y_{n-1}y_n) \\ \vdots & \quad \vdots \\ S_n &= (-1)^n(y_1y_2y_3 \cdots y_{n-1}y_n). \end{aligned} \tag{3}$$

The insight is that by Puiseux's theorem, when we plug the expansions (2) into the symmetric functions (3), *the resulting fractional power series for the coefficients all collapse down to the integral polynomials* S_k *in* t. For later referral, it will be convenient to formulate this simple observation as a theorem, calling the part of the expansion with positive exponents the *polynomial part*.

Theorem 16. *For any* $g(x,y) \in \mathbb{C}[x,y]$ *the elementary symmetric functions of* n *Puiseux expansions* (2) *collapse down to polynomials in* $\mathbb{Z}[t]$ *if and only if:*

(a) *The coefficients of all the* negative *powers of* τ *in the resulting fractional power series for the coefficients are* all equal to zero.

(b) *The numerical coefficients of the* "polynomial part" *of the resulting fractional power series for the coefficients* are all integers.

(c) *The numerical coefficients of the* positive fractional powers *of* τ *in the resulting fractional power series for the coefficients are* all equal to zero. ■

These three conditions will, with appropriate changes, characterize the coefficients of any polynomial factor in $\mathbb{Z}[y,t]$ of $g(y,t)$. Lemma 14 above shows that we can write "rational numbers" instead of "integers" in condition (b).

[1]The original work of Puiseux can be found in *Liouville's* Journal, vols. 15, 16 (1850, 1851). These expansions have already been used by *C. Runge* to derive necessary conditions that an equation between two unknowns have infinitely many integral solutions. See this Journal, vol. 100, p. 425. [Hilbert's footnote]

6. THE FORMAL FACTORS. Any nontrivial *formal* polynomial factor of $g(y,t)$ is a polynomial of the form

$$\pi_A(y,t) := \prod_{y_j \in A} (y - y_j), \tag{4}$$

where A is a subset of the roots $\{y_1, y_2, \ldots, y_n\}$. As Hilbert points out, there are $\binom{n}{2}$ quadratic factors, $\binom{n}{3}$ cubic factors, $\binom{n}{4}$ quartic factors, $\binom{n}{5}$ quintic factors, and so on, and finally $\binom{n}{n-1}$ factors of degree $n-1$. Additionally, we count the n linear factors for a grand total of

$$\binom{n}{2} + \binom{n}{3} + \cdots + \binom{n}{n-1} + n = 2^n - 2$$

possible factors. So, if $g(y,t)$ is reducible, *some* $\pi_A(y,t)$*, must be an integral polynomial factor.* Sometimes we prefer to think of A as a single item rather than a set, so we assign it a unique index a where $a = 1, 2, \ldots, 2^n - 2$. Then $\pi_a(y,t)$ means the same as $\pi_A(y,t)$. These items will become the "colors" in the Cube Lemma.

Let's look at a simple example of a reducible integral polynomial:

$$g(y,t) := y^3 - t^3.$$

Then the roots of $g(y,t) = 0$ are $y_1 = t$, $y_2 = \omega t$, $y_3 = \omega^2 t$ where $\omega^3 = 1$, $\omega \neq 1$.[2] When $n = 3$ there are $2^3 - 2 = 6$ formal factors. Thus the sets A are

$$\{y_1\}, \quad \{y_2\}, \quad \{y_3\}, \quad \{y_1, y_2\}, \quad \{y_1, y_3\}, \quad \{y_2, y_3\},$$

and we (arbitrarily) assign the indices $a = 1, 2, 3, 4, 5, 6$ to them respectively. Then, by (4) these formal factors are:

$$\begin{aligned}
\pi_{\{y_1\}} &\equiv \pi_1(y,t) = y - y_1 = y - t,\\
\pi_{\{y_2\}} &\equiv \pi_2(y,t) = y - y_2 = y - \omega t,\\
\pi_{\{y_3\}} &\equiv \pi_3(y,t) = y - y_3 = y - \omega^2 t,\\
\pi_{\{y_1,y_2\}} &\equiv \pi_4(y,t) = (y - y_1)(y - y_2) = y^2 + \omega t y + \omega^2 t^2,\\
\pi_{\{y_1,y_3\}} &\equiv \pi_5(y,t) = (y - y_1)(y - y_3) = y^2 + \omega^2 t y + \omega t^2,\\
\pi_{\{y_2,y_3\}} &\equiv \pi_6(y,t) = (y - y_2)(y - y_3) = y^2 + t y + t^2.
\end{aligned}$$

We observe that $\pi_1(y,t)$ and $\pi_6(y,t)$ are *integral* polynomial factors whereas the other four are not.

7. USING THE PIGEONHOLE PRINCIPLE. Our problem now is to discover *at least one* formal factor π_α that is an *integral* polynomial. (We note that its complementary factor is also an integral polynomial.) We begin our search by applying our hypothesis.

Take t_0 from the hypothesis of Theorem 12 and τ_0 to be the positive kth root of t_0. Usually, τ_0 will be irrational. By hypothesis, if we substitute τ_0 into all of the coefficient series of the formal factors $\pi_a(y,t)$, *at least one of them will be an integral polynomial in* y.

[2] Moreover, $y_1 = t, y_2 = \omega t, y_3 = \omega^2 t$ where $\omega^3 = 1, \omega \neq 1$ are also the Puiseux expansions of the roots.

Now, along with Hilbert, we observe that if we substitute $2\tau_0$ into all of the coefficient series of the formal factors $\pi_a(y,t)$, then *at least one of them will be an integral polynomial in* y, by the assumption that $g(y, 2^k t_0)$ is reducible.

Again, if we substitute $3\tau_0$ into all of the coefficient series of the formal factors $\pi_a(y,t)$ *at least one of them will be an integral polynomial in* y, by the assumption that $g(y, 3^k t_0)$ is reducible. The same is true of $4\tau_0$, $5\tau_0$, and indeed, of $\sigma\tau_0$ for $\sigma = 1, 2, 3, \dots$.

Therefore, we obtain *an infinite sequence of integral polynomial factors* $\pi_a(y, \sigma^k t_0)$ *in* y. Each of them has a unique index a where $a = 1, 2, \dots, 2^n - 2$. Let these indices be $a_1, a_2, a_3, \dots, a_s, \dots$. Then, by the pigeonhole principle, *at least one index* a_s *occurs infinitely often*. In our example above, we can take $a_s = 1$ or $a_s = 6$ and our sequence of indices contains either 1 or 6 or both infinitely often. The point is that:

> The corresponding formal polynomial $\pi_{a_s}(y,t)$ is a natural candidate for our integral polynomial factor.

To prove that it *is* our integral polynomial factor, we must verify that its Puiseux series satisfy the three conditions of Theorem 16. The rest of Hilbert's paper (and ours) is the proof that *the candidate formal factor* $\pi_{a_s}(y,t)$ *satisfies these three conditions* .

8. FRAMING THE CUBE LEMMA. Let's consider the first condition: *The coefficients of all the* negative *powers of* τ_0 *in the resulting fractional power series for the coefficients of* $\pi_{a_s}(y, \sigma^k t_0)$ *are all equal to zero*. Suppose the following system of coefficient power series for $\pi_{a_s}(y, (\sigma^k t_0))$ has a_s as its index:

$$\begin{aligned} y_1 + y_2 + \cdots + y_\nu &= \mathrm{A}_{11}(\sigma\tau_0)^h + \mathrm{A}_{12}(\sigma\tau_0)^{h-1} + \cdots + \mathrm{A}_{1,h+1} \\ &\quad + \frac{\mathrm{B}_{11}}{\sigma\tau_0} + \frac{\mathrm{B}_{12}}{(\sigma\tau_0)^2} + \frac{\mathrm{B}_{13}}{(\sigma\tau_0)^3} + \cdots \\ \vdots \qquad & \qquad \vdots \\ y_1 y_2 \cdots y_\nu &= \mathrm{A}_{\nu 1}(\sigma\tau_0)^{h\nu} + \mathrm{A}_{\nu 2}(\sigma\tau_0)^{h\nu-1} + \cdots + \mathrm{A}_{\nu,h\nu+1} \\ &\quad + \frac{\mathrm{B}_{\nu 1}}{\sigma\tau_0} + \frac{\mathrm{B}_{\nu 2}}{(\sigma\tau_0)^2} + \frac{\mathrm{B}_{\nu 3}}{(\sigma\tau_0)^3} + \cdots . \end{aligned}$$

The coefficients A, B are all completely determinate rational or irrational, real or complex numbers; some of them have the value zero since the positive exponents of τ_0 in general will be smaller than $(n-1)h$.

The variable quantity here is the integer σ. This suggests that *we rewrite the above fractional series as series in* σ and then obtain:

$$\begin{aligned} y_1 + y_2 + \cdots + y_\nu &= A_{11}\sigma^h + A_{12}\sigma^{h-1} + \cdots + A_{1,h+1} \\ &\quad + \frac{B_{11}}{\sigma} + \frac{B_{12}}{\sigma^2} + \frac{B_{13}}{\sigma^3} + \cdots \\ \vdots \qquad & \qquad \vdots \\ y_1 y_2 \cdots y_\nu &= A_{\nu 1}\sigma^{h\nu} + A_{\nu 2}\sigma^{h\nu-1} + \cdots + A_{\nu,h\nu+1} \\ &\quad + \frac{B_{\nu 1}}{\sigma} + \frac{B_{\nu 2}}{\sigma^2} + \frac{B_{\nu 3}}{\sigma^3} + \cdots \end{aligned}$$

where the new coefficients A, B are again determinate numerical quantities. Suppose that the index of the first occurrence of our infinitely repeated polynomial factor is $s = \sigma = \mu$. Then every repetition of the index μ produces the same ν power series in σ, but with a *larger* value of σ. Thus, since there are an infinite number of such indices, *there are infinitely many larger and larger values of σ substituted into the power series.*

If we look at the series of *negative* powers for any particular coefficient, we see that for sufficiently large σ it become arbitrarily small in absolute value. Yet, the total power series takes integral values for all of these values of σ. That suggests that the total contribution of the negative powers, for large σ, *is an integer of arbitrarily small absolute value, i.e., zero*.

Thus we might try to argue by contradiction as follows: Assume that there *are* nonzero coefficients of negative powers and deduce an absurd conclusion. The possible hitch is that this inference ignores the "polynomial part" of the coefficient series—which could exactly compensate for a tiny nonzero contribution of the negative powers.

To show that this is *not* the case we would like to somehow "eliminate" the polynomial part of the coefficient series without losing the property of being an integer for infinitely many values of σ. This suggests *forming suitable linear combinations of the coefficient series which successively subtract off the principal terms of the polynomial parts, and leaving finally only linear combinations of integer-valued series with negative coefficients*.

To see how this would work, let's look at a typical coefficient series. Let us choose any of the ν power series in the system under consideration, say the power series

$$\mathcal{P}(\sigma) = A_{11}\sigma^{m-1} + A_{12}\sigma^{m-2} + \cdots + A_{1m} + \frac{B_{11}}{\sigma} + \frac{B_{12}}{\sigma^2} + \frac{B_{13}}{\sigma^3} + \cdots .$$

where we have written $m-1$ for the highest power of σ.

Now comes a new insight. *This is the insight that is key to the whole proof.* Its simplicity belies the brilliance it took to think of it. Professional mathematicians are aware of this phenomenon: the deepest ideas, in the end, are based on a simple observation. Here is Hilbert's:

> Suppose that the series $\mathcal{P}(\sigma)$ takes on integral values, not only for infinitely many values of σ, *but also for all the infinitely many values of $\sigma + \mu^{(1)}$, where $\mu^{(1)}$ is a fixed increment independent of σ.*

Now we form the linear combination

$$\mathcal{P}^{(1)}(\sigma) := \mathcal{P}(\sigma) - \mathcal{P}(\sigma + \mu^{(1)})$$

and put the polynomial part equal to

$$\varphi_{m-1}(\sigma) := A_{11}\sigma^{m-1} + A_{12}\sigma^{m-2} + \cdots + A_{1m}$$

for brevity. We now start the argument-by-contradiction.

Suppose now that the other coefficients $B_{11}, B_{12}, B_{13}, \ldots$ of the power series $\mathcal{P}(\sigma)$ *are not all zero* and let B_{1v}/σ^v be the first term whose coefficient B_{1v} does *not* vanish. Then

$$\mathcal{P}^{(1)}(\sigma) = \varphi_{m-1}(\sigma) - \varphi_{m-1}(\sigma + \mu^{(1)}) + B_{1v}\left[\frac{1}{\sigma^v} - \frac{1}{(\sigma + \mu^{(1)})^v}\right] + \cdots .$$

Here the first difference on the right-hand side *is a polynomial of degree* $m-2$ *in* σ; we put

$$\varphi_{m-2}(\sigma) = \varphi_{m-1}(\sigma) - \varphi_{m-1}(\sigma + \mu^{(1)}).$$

We expand the remaining terms on the right-hand side in decreasing powers of σ; then we obtain[3]

$$\mathcal{P}^{(1)}(\sigma) = \varphi_{m-2}(\sigma) + \mu^{(1)} v \frac{B_{1v}}{\sigma^{v+1}} + \cdots .$$

We have reduced the maximum degree of the polynomial part by one unit. Moreover, $\mathcal{P}^{(1)}(\sigma)$ *takes on integral values for infinitely many* σ.

We have *not* proved that such an increment $\mu^{(1)}$ exists, but if we could, then we could make a first step in reaching our goal of producing a series of negative powers of σ that is an integer for infinitely many values of σ.

To carry out a similar program to *reduce the maximum degree of the polynomial part to* $m-3$*, then to* $m-4$*, and so on until finally to zero*, we would have to have to prove the existence of m fixed increments $\mu^{(k)}$, $k = 1, 2, \cdots, m$ *whose values are all independent of* σ *and such that if we substitute any of the integers*:

$$\boxed{\begin{array}{cccccc}
\mu & & & & & \\
\mu+\mu^{(1)} & & & & & \\
\mu+\mu^{(2)} & \mu+\mu^{(1)}+\mu^{(2)} & & & & \\
\mu+\mu^{(3)} & \mu+\mu^{(1)}+\mu^{(3)} & \mu+\mu^{(2)}+\mu^{(3)} & \mu+\mu^{(1)}+\mu^{(2)}+\mu^{(3)} & & \\
\vdots & \vdots & \vdots & \vdots & \cdots & \ddots \\
\mu+\mu^{(m)} & \mu+\mu^{(1)}+\mu^{(m)} & \mu+\mu^{(2)}+\mu^{(m)} & \cdots & \cdots & \mu+\mu^{(1)}+\cdots+\mu^{(m)}
\end{array}}$$

for σ *in* $\mathcal{P}(\sigma)$*, the values will all be integers*.

Note that the proof of the existence of the increment $\mu^{(1)}$ amounts to proving that a_μ and $a_{\mu+\mu^{(1)}}$ are both *the same number in the set of indices* a_s. A similar property holds for the set of all the above sums of fixed increments $\mu^{(k)}$ (for $k = 1, 2, \ldots, m$) with μ, namely that they all are the subscripts of *the same number in the set of indices* a_s. In our running example they are all equal to 1 or they are all equal to 6.

The proof of the existence of these increments is the content of the *Hilbert cube lemma*. To serve the context of Hilbert's proof, we re-state it using his formulas much as he visualized them in his paper.

[3]The details, with simplified notation, are:

$$\begin{aligned}
B_{1v}\left[\frac{1}{\sigma^\nu} - \frac{1}{(\sigma+\mu)^\nu}\right] &= \frac{B_{1v}}{\sigma^\nu}\left[1 - \frac{1}{(1+\frac{\mu}{\sigma})^\nu}\right] \\
&= \frac{B_{1v}}{\sigma^\nu}\left[1 - \left\{1 + \binom{-\nu}{1}\frac{\mu}{\sigma} + \binom{-\nu}{2}\left(\frac{\mu}{\sigma}\right)^2 + \cdots\right\}\right] \\
&= \frac{B_{1v}}{\sigma^\nu}\left[\frac{\mu\nu}{\sigma} - \frac{\nu(\nu+1)}{2}\left(\frac{\mu}{\sigma}\right)^2 + \cdots\right] \\
&= \frac{\mu\nu B_{1v}}{\sigma^{\nu+1}}\left[1 - \frac{\nu+1}{2}\left(\frac{\mu}{\sigma}\right) + \cdots\right],
\end{aligned}$$

and this last expression on the right-hand side is equivalent to that in the main body of the paper.

Theorem 17. *Let $a_1, a_2, a_3, \ldots$ be an infinite sequence in which the general term, a_s, is one of the a positive numbers $1, 2, \ldots, a$. Moreover, let m be any positive integer. Then we can always find m positive integers $\mu^{(1)}, \mu^{(2)}, \ldots, \mu^{(m)}$ such that for infinitely many integers μ the 2^m elements*

$$\boxed{\begin{array}{llllll}
a_\mu & & & & & \\
a_{\mu+\mu^{(1)}} & & & & & \\
a_{\mu+\mu^{(2)}} & a_{\mu+\mu^{(1)}+\mu^{(2)}} & & & & \\
a_{\mu+\mu^{(3)}} & a_{\mu+\mu^{(1)}+\mu^{(3)}} & a_{\mu+\mu^{(2)}+\mu^{(3)}} & a_{\mu+\mu^{(1)}+\mu^{(2)}+\mu^{(3)}} & & \\
\vdots & \vdots & \vdots & \vdots & \cdots & \ddots \\
a_{\mu+\mu^{(m)}} & a_{\mu+\mu^{(1)}+\mu^{(m)}} & a_{\mu+\mu^{(2)}+\mu^{(m)}} & \cdots & \cdots & a_{\mu+\mu^{(1)}+\mu^{(2)}+\cdots+\mu^{(m)}}
\end{array}}$$

are all equal to the same number G, where G is one of the numbers $1, 2, \ldots, a$.

Thus we see that the statement *arises naturally from the necessity of proving that the coefficients of the negative powers of σ must be all equal to zero*. It is the strengthened form of the pigeonhole principle we mentioned earlier. It is stronger because it imposes a *structure* on the distribution of infinitely many common values a_s whereas the pigeonhole principle only implies their *existence*.

9. THE COEFFICIENTS OF THE NEGATIVE POWERS OF σ ARE ZERO. Employing the above idea, we form the following m linear combinations:

$$\begin{aligned}
\mathcal{P}^{(1)}(\sigma) &= \mathcal{P}(\sigma) - \mathcal{P}(\sigma + \mu^{(1)}), \\
\mathcal{P}^{(2)}(\sigma) &= \mathcal{P}^{(1)}(\sigma) - \mathcal{P}^{(1)}(\sigma + \mu^{(2)}), \\
\vdots \quad & \quad \vdots \\
\mathcal{P}^{(m)}(\sigma) &= \mathcal{P}^{(m-1)}(\sigma) - \mathcal{P}^{(m-1)}(\sigma + \mu^{(m)}).
\end{aligned}$$

It follows from what we proved earlier that each of these m power series also assumes integer values for infinitely many integral arguments $\sigma = \mu$.

As we indicated, assuming the cube lemma, we obtain

$$\mathcal{P}^{(2)}(\sigma) = \varphi_{m-3}(\sigma) + \mu^{(1)}\mu^{(2)} v(v+1)\frac{B_{1v}}{\sigma^{v+2}} + \cdots ,$$

where $\varphi_{m-3}(\sigma)$ is a polynomial in σ of degree $m-3$. *After m steps we arrive finally at the formula*

$$\mathcal{P}^{(m)}(\sigma) = \mu^{(1)}\mu^{(2)} \cdots \mu^{(m)} v(v+1)\cdots(v+m-1)\frac{B_{1v}}{\sigma^{v+m}} + \cdots .$$

Since this power series begins with negative powers of σ, we can find a positive number Γ such that for all values of σ that exceed Γ the absolute value of the power series *will be smaller than one*. On the other hand, the power series $\mathcal{P}^{(m)}(\sigma)$ is itself *equal to an integer for infinitely many arguments σ* and since an integer whose absolute value is less than one is necessarily equal to zero, it follows that *there are infinitely many integers σ for which the power series vanishes*.

But, our last formula shows us that

$$\lim_{\sigma\to\infty}\left[\sigma^{v+m}\mathcal{P}^{(m)}(\sigma)\right]=\mu^{(1)}\mu^{(2)}\cdots\mu^{(m)}v(v+1)\cdots(v+m-1)B_{1v},$$

where the expression on the right hand-side represents a quantity *different from zero*. This last result stands in *contradiction* with the conclusion above, and therefore *it is impossible that a nonzero coefficient* B_{1v} *occurs among the coefficients* $B_{11}, B_{12}, B_{13}, \ldots$. It follows in the same way that *also the coefficients* $B_{2i}, B_{3i}, B_{4i}, \ldots, B_{\nu i}$ *must all be equal to zero*.

This completes the proof of the first condition of Theorem 16 about the Puiseux expansions of the coefficients of $\pi_{a_s}(y,t)$. ■

This step was the heart of Hilbert's proof and his paper's most brilliant insight. The other parts are clever too, but in our opinion this best shows his penetrating originality.

10. THE COEFFICIENTS OF THE POLYNOMIAL PART ARE RATIONAL NUMBERS. The next condition of Theorem 16 to be verified is: *the numerical coefficients in the polynomial part of the Puiseux expansions of the coefficients of* $\pi_{a_s}(y,t)$ *are rational numbers*. Our expansion has collapsed to the polynomial part:

$$\mathcal{P}(\sigma)=A_{11}\sigma^{m-1}+A_{12}\sigma^{m-2}+\cdots+A_{1m}, \tag{5}$$

where the right-hand side assumes integer values for infinitely many values of σ. If we set the right-hand side equal to these integers for m values of σ we obtain m linear equations with m unknowns $A_{11}, A_{1,2}, \ldots, A_{1m}$ which have a *rational solution* by Cramer's rule. By Proposition 13, getting "rational" suffices to prove the condition. ■

11. ONLY INTEGRAL POWERS OF t. The final condition of Theorem 16 to be verified is: *the only nonzero terms in the polynomial part of the Puiseux expansions of the coefficients of* $\pi_{a_s}(y,t)$ *are those with integral powers of* t.

Take τ_0 to be a *prime* number p larger than C'' and recall $\sigma\tau_0=\tau$. We now determine 2^n-2 distinct prime numbers $p', p'', \ldots, p^{2^n-2}$ all greater than p. Then also for each of these prime numbers there exists at least one among the 2^n-2 formal factors whose coefficients have the above polynomial form (5). However, since the number of prime numbers $p, p', p'', \ldots, p^{2^n-2}$ is equal to 2^n-1 while the number formal factors only reaches 2^n-2, necessarily *there must exist at least one formal factor admitting a double representation by these polynomials* (5). That is to say, as above:

$$\begin{aligned}
y_1+y_2+\cdots+y_\nu &= A_{11}p^{-(m-1)/k}\tau^{m-1}+A_{12}p^{-(m-2)/k}\tau^{m-2}+\cdots+A_{1m}\\
&\ \ \vdots \\
y_1y_2\cdots y_\nu &= A_{\nu 1}p^{-(m-1)/k}\tau^{m-1}+A_{\nu 2}p^{-(m-2)/k}\tau^{m-2}+\cdots+A_{\nu m}
\end{aligned}$$

and simultaneously

$$\begin{aligned}
y_1+y_2+\cdots+y_\nu &= A'_{11}p'^{-(m-1)/k}\tau^{m-1}+A'_{12}p'^{-(m-2)/k}\tau^{m-2}+\cdots+A'_{1m}\\
&\ \ \vdots \\
y_1y_2\cdots y_\nu &= A'_{\nu 1}p'^{-(m-1)/k}\tau^{m-1}+A'_{\nu 2}p'^{-(m-2)/k}\tau^{m-2}+\cdots+A'_{\nu m}\,.
\end{aligned}$$

Since, by Puiseux's theorem, the coefficients of the powers of τ are unique, if we equate coefficients of equal powers of τ on the right-hand sides we obtain:

$$\begin{array}{ccc} A_{11}p^{-(m-1)/k} = A'_{11}p'^{-(m-1)/k} & \cdots = \cdots & A_{1m}p^{-(m-1)/k} = A'_{1m}p'^{-(m-1)/k} \\ A_{21}p^{-(m-1)/k} = A'_{21}p'^{-(m-1)/k} & \cdots = \cdots & A_{2m}p^{-(m-1)/k} = A'_{2m}p'^{-(m-1)/k} \\ \vdots \quad = \quad \vdots & \cdots = \cdots & \vdots \quad = \quad \vdots \\ A_{\nu 1} = A'_{\nu 1} & \cdots = \cdots & A_{\nu m} = A'_{\nu m} \end{array}$$

Since the coefficients A, B and A', B' are all rational numbers and p and p' are distinct prime numbers, the above equations show us that *the only coefficients that can be different from zero are those for which the corresponding exponent of τ must be an integer divisible by k.*

That is, the power series of our system *are polynomials in τ^k with rational coefficients,* and if we put $\tau^k = t$, we obtain

$$\begin{aligned} y_1 + y_2 + \cdots + y_\nu &= F_1(t) \\ \vdots \qquad & \quad \vdots \\ y_1 y_2 \cdots y_\nu &= F_\nu(t), \end{aligned}$$

where $F_1(t), \ldots, F_\nu(t)$ are *polynomials in t with rational coefficients.*

This completes the proof of the third condition of Theorem 16 of the Puiseux expansions of the coefficients of some formal factor $\pi_{a_s}(y,t)$. ■

We note that the final formal factor $\pi_{a_s}(y,t)$ is not necessarily the same one that we started with. All we needed was that it fulfills the three conditions of Theorem 16, and therefore the proof of Theorem 12 is complete. ■

12. LATER PROOFS OF THE IRREDUCIBILITY THEOREM. After Hilbert, many mathematicians offered other proofs of the irreducibility theorem.

Most of the modern proofs of the (two-variable) irreducibility theorem are based on that of Karl Dörge [**4**], which sharpened an idea of Thoralf Skolem [**26**]. Dörge proved it without using the cube lemma and obtained a stronger result. To begin contrasting his and Hilbert's results, recall Hilbert's statement that if $f \in \mathbb{Z}[x, t_1, \ldots, t_s]$ is irreducible, then for infinity many $a_1, \ldots, a_s \in \mathbb{Z}$, $f(x, a_1, \ldots, a_s)$ is irreducible.

Now let $|f|$ be the maximum of 8 and the absolute values of the coefficients of f (the reason for insisting $|f| \geq 8$ is technical). A simplified statement of Dörge's theorem is:

Theorem 18. *There is a function $c(d,s)$ such that the following holds. Let $f \in \mathbb{Z}[x, t_1, \ldots, t_s]$ be irreducible of degree d. Let $N > |f|^{c(d,s)}$. Then the number of $(a_1, \ldots, a_s) \in \{-N, \ldots, N\}^s$ such that $f(x, a_1, \ldots, a_s)$ is not irreducible is at most $|f|^{c(d,s)} N^{s-(1/2)} \log N$.*

Note that the number of such $(a_1, \ldots, a_s)$ has density 0. Dörge actually presented a generalization of this theorem where he replaces $\mathbb{Z}$ with the integers of a finite extension of a number field.

Dörge also showed (in fact this was his primary interest) that if f, viewed as an element of $\mathbb{Z}[t_1, \ldots, t_s][x]$, has Galois group G, then the number of $(a_1, \ldots, a_s) \in \{-N, \ldots, N\}^s$ such that $f(x, a_1, \ldots, a_s)$ does not have Galois group G is at most

$|f|^{c(d,s)} N^{s-(1/2)} \log N$. And again, he actually presented a generalization of this theorem that replaces $\mathbb{Z}$ with the integers of a finite extension of a number field.

Lang [**17, 18**] and Prasolov [**20**] have expositions of Dörge's proof. Franz [**8**] also gave a proof that does not use the cube lemma, and this is expounded further by Schinzel [**25**]. There is a another alternative proof by Fried [**9**]. Serre [**28**] recasts these results in geometric terms and presents results about which groups can be Galois groups.

13. HILBERT CUBE NUMBERS AND CONCLUSIONS. Define the "Hilbert Cube Number" $H(m,c)$ to be the *least* number H such that every c-coloring of $1,\ldots,H$ has a monochromatic m-cube. Our proof of the cube lemma in Section 3 showed a recursive upper bound $H(m,c) \leqslant H(m-1,c)(1+c^{H(m-1,c)})$, with basis $H(1,c)=1$ for all c. This is far from best possible. For one thing, when $2 \leqslant m \leqslant c$ one can improve the upper bound to $H(m,c) \leqslant h(1+c(m-1)^h)$, where $h = H(m-1,c)$, by a different counting argument. One can further tweak this with $\binom{m-1}{h}$ in place of $(m-1)^h$. These formulas are not bounded by any fixed tower of exponents in c and m.

As observed by Brown et al. [**2**], Hilbert's original proof yields bounds with $(c+1)$ rather than $(m-1)$ in the base and the Fibonacci number F_{2m} in the exponent. Namely, $H(m,c) \leqslant (c+1)^{F_{2m}}$, where $F_0=0, F_1=1, F_2=1, F_3=2,\ldots$. These bounds have double-exponential growth. Szemerédi [**29**] (see also [**10**]) improved both the bounds and the nature of the result, showing that any subset A of $[1,\ldots,H]$ of density $1/c$ (that is, $|A| \geqslant H/c$) contains an m-cube where $m \geqslant \log\log(H) - C$ and C depends only on c. The best known upper and lower bounds appear still to be those of Gunderson and Rödl [**11**]:

$$c^{(1-\epsilon_c)(2^m-1)/m} \leqslant H(m,c) \leqslant (2c)^{2^{m-1}},$$

where $\epsilon_c \to 0$ as $c \to \infty$. The same upper bound was recently ascribed to [**12**] by Conlon, Fox, and Sudakov [**3**] but see also Sándor [**23**] with different asymptotics. Erdős and Turán [**7**] proved that $H(2,c)$ is asymptotic to c^2, but the remarks in [**2**] that less is known about $H(m,c)$ for fixed $m \geqslant 3$ appear still in force, and [**3**] remarks that $H(m,2)$ depends on unknown properties of van der Waerden numbers.

We have shown the significance of the cube lemma in the context of Hilbert's original paper. The question of whether Hilbert might have expanded on it bids comparison with Ramsey's motivation in [**22**]. That was a problem in logic not number theory *per se*. But Hilbert was the world's master in the relationship between number theory and logic until Gödel emerged, so one may ask again why Hilbert didn't pursue this further or develop areas of extremal combinatorics *vis-à-vis* logic in the direction of Ramsey theory. We close with a speculative answer: The world in which Hilbert was immersed is as different from that of Ramsey theory as "doubly-exponential" is from "singly-exponential."

The years 1890–1893 saw the publication of Hilbert's great foundational works in commutative algebra, including his basis theorem and *Nullstellensatz* [**13, 15**]. A common thread through all this work is the notion of *regularity*: given a finitely-specified system of elements that may have arbitrarily large values of some parameter t (such as the degree of polynomials over a ring), there is some integer t_0 such that for all $t \geqslant t_0$ the system conforms to a simple description. Hilbert first proved his basis theorem nonconstructively. Later was it shown that the growth of the relevant t_0 (in terms of the degrees d of basis elements or the n-variable equations in the *Nullstellensatz*) is double-exponential, of order at most d^{2^n}. Our answer to the question we posed in the previous paragraph is that Hilbert was simply occupied with more-rarefied levels of algebra and analysis revolving around invariants. Irreducibility of polynomials plays into irreducible varieties and primary decompositions of polynomial ideals, which Hilbert's student Emanuel Lasker (the world chess champion) and colleague Emmy Noether built upon for some great work in the next two decades. Meanwhile, Hilbert swooped down to the utterly ground-level task of formalizing Euclid's geometry in the later 1890s, which presaged his work on formal systems of logic.

The divide in purpose and growth rate doesn't ward us off from appreciating the cube numbers and seeking other uses for them. That is why we have devoted this paper to expounding their original use and context. We have highlighted how the cube lemma completed an insight about estimates by infinite series. We hope that our exposition will foster a greater appreciation of combinatorial underpinnings of more "analytical" areas of mathematics.

REFERENCES

1. M. Bôcher. *Introduction to Higher Algebra*. Dover, New York, 2004.
2. T. C. Brown, F. R. K. Chung, P. Erdős, R. L. Graham. Quantitative forms of a theorem of Hilbert. *J. Combin. Theory Ser. A* **38** (1985) 210–216.
3. D. Conlon, J. Fox, B. Sudakov. Short proofs of some extremal results. *Combin. Probab. Comput.* **23** (2014) 8–28.
4. K. Dörge. Einfacher Beweis des Hilbertschen Irreduzibilitätssatzes. *Math. Ann.* **96** (1927) 176–182.
5. P. Erdős and G. Szekeres A combinatorial problem in geometry. *Compositio Math.* 2 (1935) 463-470.
6. C.F. Gausds *Disquisitiones Arithmeticae*. Translated by Arthur A. Clarke Yale University Press, New York, 1965.
7. P. Erdős and P. Turán. On a problem of Sidon in additive number theory, and on some related problems. *J. London Math. Soc.* 16 (1941) 212–215.
8. W. Franz. Untersuchungen zum Hilbertschen Irreduzibilitätssatz. *Math. Z.* **33** (1931) 275–293.
9. D. Fried. On Hilbert's irreducibility theorem. *J. Number Theory* **6** (1974) 211–231.
10. R. Graham, B. Rothschild, J. Spencer. *Ramsey Theory*. Wiley, New York, 1990.
11. D. S. Gunderson, V. Rödl. Extremal problems for affine cubes of integers. *Combin. Probab. Comput.* **7** (1998) 65–79.
12. D. S. Gunderson, V. Rödl, A. Sidorenko. Extremal problems for sets forming Boolean algebras and complete partite hypergraphs. *J. Combin. Theory Ser. A 88* (1999) 342–367.
13. D. Hilbert. Über die Theorie der algebraischen Formen. *Math. Ann.* **36** (1890) 473–534.
14. D. Hilbert. Über die Irreducibilität ganzer rationaler Functionen mit ganzzahligen Coefficienten. *J. reine angew. Math.* **110** (1892) 104–129.
15. D. Hilbert. Über die vollen Invariantensysteme. *Math. Ann.* 42 (1893) 313–373.
16. B. Landman, A. Robertson. *Ramsey Theory on the Integers*. AMS, Providence, 2004.
17. S. Lang. *Diophantine Geometry*. Wiley, New York, 1962.
18. S. Lang. *Fundamentals of Diophantine Geometry*. Springer, New York, 1983.
19. K. Nowak. Some elementary proofs of Puiseux's theorems. *Univ. Iagel. Acta Math.* **38** (2000) 279–282.
20. V. Prasolov. *Polynomials*. Springer, New York, 2004.
21. H. J. Promel. *Ramsey Theory for Discrete Structures*. Springer, New York, 2013.
22. F. P. Ramsey. On a problem in formal logic. *Proc. London Math. Soc.* **30** (1930) 264–286.
23. C. Sándor. An upper bound for Hilbert cubes. *J. Combin. Theory Ser. A* **114** (2007) 1157–1159.
24. I. Schur. Uber die Kongruenz of $x^m + y^m \equiv z^m \pmod p$ *Jahresbericht der Deutschen Mathematiker-Vereinigung* **24** (1916) 114-116.
25. A. Schinzel. *Selected Topics on Polynomials*. Univ. of Michigan Press, Ann Arbor, 1982.
26. T. Skolem. Untersuchungen über die moeglichen Verteilungen ganzzahliger Lösungen gewisser Gleichungen. *Kristiania Vid. Selsk. Skr.* **17**, 1921; 57 pp.
27. A. Soifer. *The Mathematical Coloring Book: Mathematics of coloring and the colorful life of its creators*. Springer, Berlin, 2009.
28. J.-P. Serre *Topics in Galois Theory*. A.K. Peters-CRC Press, New York, 2008.
29. E. Szemerédi. On sets of integers containing no four elements in arithmetic progression. *Acta. Math. Hungar.* **20** (1969) 89–104.
30. N. Tzanakos. *Elliptic Diophantine Equations*. De Gruyter, Berlin, 2013.
31. B. van der Waerden. Beweis einer Baudetschen Vermutung. *Nieuw Arch. Wisk.* **15** (1927) 212-216.