OBITUARY

Alan Baker, FRS, 1939–2018

1.   *Life and career*

Alan's paternal grandparents were known as Marks and Mathilda Backer; they were married in 1902 in Lithuania. The surname presumably changed when children arrived. His parents were Barnet and Bessie (née Sohn). Alan was born in London on 19 August 1939 into this Jewish family. His earliest memories of wartime England were of evacuation to Camberley, Surrey. After the war, the family moved to Forest Gate in East London, where he spent most of his early life. From a very early age, he showed signs of mathematical brilliance (see the comments later about 'brainbox') and was encouraged by his parents. Already his father (who had been at school with Jacob Bronowski) was very gifted in this direction, but did not have the opportunity to develop, and became a tailor instead. This may have explained Alan's clothes sense; he was always well-turned-out with quality suits and tasteful ties (which were, however, not always entirely appropriate to his later travelling, for example, on the beach at Nice after being awarded the Fields Medal, or climbing a hill in the Australian Bush among snakes — such episodes may have contributed to his later investing in a distinctive yellow sun hat).

Alan's first education was at a Franciscan Convent (London E7) 1945–1949, followed by Godwin County Primary School 1949–1950 and Stratford Grammar School 1950–1958. He then went with a state scholarship to University College, London (he recalls that the staff included J. W. Archbold, L. S. Bosanquet, P. Du Val, T. Estermann, C. A. Rogers, and K. Roth), where he studied mathematics 1958–1961, obtaining a first-class honours BSc (Special) degree.

He then moved to Trinity College, Cambridge (where he would soon be based for the rest of his life) to study from 1961 to 1964 with Harold Davenport, at the time one of the leading number theorists with many international connections, of whom Alan writes:

> 'An excellent mathematician from whom I learnt a great deal; but I tended to follow my own lines of research.'

Still, Chris Morley recalls a loud mathematical conversation between them in the Trinity Parlour, and they did write a paper together (see below). It seems that in later life he enjoyed imitating Davenport's Lancashire (Accrington) accent (and the writer has carried on this worthy tradition of imitating one's supervisor). He obtained his PhD in 1965 and MA in 1966, by which time he had already been awarded a Prize Fellowship for 1964–1968, also at Trinity. During this period, which included a year 1964–1965 at University College, he took on John Coates as a PhD student, and recalls also sharing responsibilities with Davenport in supervising T. W. Cusick, M. N. Huxley, H. L. Montgomery, and R. W. K. Odoni.

In 1966, he was appointed as Assistant Lecturer and in 1968 as Lecturer; in 1972 he was promoted to Reader in the Theory of Numbers, and finally in 1974 he was elected to a personal chair for Pure Mathematics, all at Cambridge. Apparently he liked to draw attention to the unusual chronological order of Fields Medal, then Fellow of the Royal Society, then Professor. During that period he supervised for a PhD also, in chronological order, the writer, Cameron Stewart, Yuval Flicker, Roger Heath-Brown, Richard Mason, Mark Coleman, and Ellyn Lee.

At the time of writing, the Mathematical Genealogy Project lists 506 descendants.

The writer recalls meeting Alan every two weeks or so in his college rooms in Whewell's Court opposite the Great Gate. Actually these rooms (fitting C. P. Snow's description 'not specially agreeable') had been occupied previously by G. H. Hardy (about which the great German mathematician David Hilbert — see later — indignantly wrote to the Master that Hardy was the best mathematician, not only in Trinity, but in England, and should therefore have the best rooms). One had to ascend a slightly low and narrow spiral staircase, and then knock on the thick wooden door. Often in College *une porte peut en cacher une autre* and probably he did indeed have a second door immediately behind; at any rate, he was sometimes a long time coming, and I supposed that he did not hear me through the two doors (also against the heavy traffic then in Trinity Street). On each visit, I felt obliged to knock or bang harder and yet harder, sometimes bringing with me a heavy book to protect my hands.

As mentioned, Alan Baker was firmly based in Cambridge; it seems that college life there suited him — especially in the style of Trinity, whose society he enriched for many years in an unspectacular way, for example by being an interesting conversationalist (contrary to the impression given by some newspaper obituaries). He also threw after-seminar parties fuelled by particularly strong beer (on one occasion Frank Adams challenged others to climb around a table, underneath it, without the feet touching the ground, after he had himself demonstrated that it could be done). In the Combination Room, there was a bottle of Madeira kept especially for him. He was a reasonable ballroom dancer, enthusiastically participating in the College May Ball. He played regularly on the Trinity Bowling Green, using an unconventional throw which delivered the ball from waist level instead of lower down.

Outside the College, he enjoyed playing the slot machines in pubs, or playing table-tennis and snooker at the Graduate Centre. Eva McLean (née Gordon) writes of one occasion:

> 'In 1975 a vicious rapist was terrorising the female population of Cambridge. One evening when I was leaving the Graduate Centre one of the porters who knew me well expressed great concern that I would be walking home alone. Not to worry, I assured him brightly, pointing to Alan who was going to escort me. The burly ex-policeman looked him up and down, all five foot six of him, grabbed his coat and joined us on our way.'

She also draws attention to existential dangers of a different kind:

> 'Alan was once barricaded in his flat for several days and would not answer the door as a collector, all the way from America, kept coming back pleading, in vain, for a contribution to his sperm bank.'

He had learnt to drive in America, and in Cambridge he bought a Rover car, apparently as an investment ('you understand', as he insisted), although Eva McLean soon explained to him that there were much better ways of investing (and he seems to have acted very well on this advice). He drove with enthusiasm despite colliding with a stationary fire engine on his maiden voyage, and despite being stopped for speeding, which caused him to arrive late in chairing a session in London. William Chen recalls that it was actually the inaugural lecture of a colleague, and that Alan was very eager to explain that he had not exceeded 90 miles per hour, so he was only speeding rather than driving recklessly. On one occasion, the car got badly stuck in a driveway; Alan, however, remained cool and concentrated and succeeded in extricating it without damage. Later on the enthusiasm waned, and a rusting hulk with flat tyres had to be removed from New Court. After that, he reverted to his old mode of transport, which was not by bicycle, as practised by the majority of academics, but on foot. He always walked very briskly and, without having to unlock and lock a bike, arrived within the city no later than the others. Keeping up with him could prove a challenge whether walking around in Cambridge, along the river, to surrounding villages, or, as he liked to do, visiting the nearby stately homes of Anglesey Abbey, Wimpole Hall, and Audley End House.

He had a flat in Hendon and enjoyed life there in London too, for example, up-market restaurants, or the theatre at which he always bought the best seats (as at the Footlights Theatre in Cambridge). The latter interest may be due to his cousin Heather Rechtman, who writes:

> 'We lived near Stratford East (London) where the fledgling company Theatre Workshop was just beginning to make a name for itself, and as a stage-struck teenager I dragged him along to many of their plays.'

That was of course edgy stuff under Joan Littlewood, but he also enjoyed musicals. About one such trip to the theatre, Eva McLean narrates:

> '... we stopped over at his Hendon flat. While there, he decided to show me his mother's mink coat. When he opened the wardrobe door, hundreds of moths awoke to greet us. His response to this crisis? He shut the door firmly.'

as well as:

> 'On another eventful outing to the West End, we arrived in good time and right there by the theatre came across a perfect parking place. Alan, however, decided it was too good to be true and so the search continued. When we next passed the spot, the parking space was gone — as was the first act of the play by the time we were inside.'

Her funeral tribute (23 February 2018) sums up such anecdotes:

> 'So, in many ways, he was what Americans term "just a regular guy".'

In his private life, he was quite relaxed, never going beyond sometimes stamping his foot when frustrated. However, in his professional life, he was often reserved, even insecure in some ways, and could on occasion be difficult. There were episodes, not just in Cambridge, which perhaps still have not been entirely forgotten. But it is now impossible to give any balanced accounts.

Eva McLean also writes:

'His Jewish identity was also something which he strictly confined to the private sphere but was happy enough to discuss at length with co-religionists.

I recollect an occasion when I told him about a particularly nasty anti-Semitic outpouring which I witnessed at a social event, adding how atypical that was these days. He, however, was convinced that anti-Semitism was never going away and was perhaps too sensitive to it. It must be remembered that when he first arrived in Cambridge (in 1961), church and chapel were as yet prominent in university life — a pillar of the academic establishment. He was worried about fitting in, contrasting his humble East End origins with the professional and even aristocratic backgrounds of many of the members of his college.

Nevertheless, he still abided by the dietary rules forbidding the consumption of pork and shellfish and requiring meat to be kosher — ancient rules that even vast numbers of practising Jews regularly attending synagogue services conveniently overlook. Whenever we ate out he ordered fish, preferably smoked, but he found culinary life more restricted in College until, to his great relief, vegetarianism finally took hold. When in Hendon, all such requirements were, of course, easily met, and that is also where he would have shown up for services from time to time — foremost while his mother was still alive.

All this should help solve a mystery that so baffled Trinity that his obituary in The Times *opened with it — namely as to why he arrived for dinner every evening late by just the few minutes necessary to miss grace, especially when as the most senior Fellow he would have been obliged to read it. He just felt that he was not the right person for the task.'*

See ⟨**15**⟩ for the above-mentioned obituary. And Paula Tretkoff writes about a meeting, also with her mother:

'Alan asked to see our Israeli passports that had come up in conversation. He wanted to see if he could still read the Hebrew writing in them. He managed OK, though he no doubt only learned Hebrew as a child without pursuing it further.'

Alan was enthusiastic about travel, which in America started already in 1969 with visiting professorships in Ann Arbor, Michigan, and Boulder, Colorado, and in 1970 membership of the Institute for Advanced Study in Princeton, New Jersey. He had at least three offers of chairs. One of these he turned down because the ivy-clad walls were too much like those in Cambridge, as was the weather. He preferred both modernity and sunshine. It was also thought that he wanted to look after his mother in London. (This writer recalls around 1980 seeing two people slowly crossing a Maryland street, and recognizing one, but not the other, and he followed them into a restaurant, where Alan introduced his mother.) He made many visits to America, and Eva McLean has shown me a lot of his postcards and other correspondence from abroad, for example, from Texas (1984), New York (1987), and California (1989). Of course, as his reputation grew he was able to widen his horizons, for example, to Australia, China and Hong Kong (missive 1988), India, Japan (postcard 1983), Russia, and many parts of Europe.

In later life, he made regular trips to Switzerland (postcards 1989 and 1990) to work with Gisbert Wüstholz at ETH (a 1988 missive opens with 'greetings from a gnome of Zürich'). There were social events too, and he thought nothing of occasionally bringing presents such as a frying pan to dinner invitations — unconventional to be sure, but with its own unassailable logic. It was there, during a conference in honour of his 60th birthday, that he gave an entertaining and surprisingly candid speech about his life, starting with his recollections of wartime London, also mentioning that he was regarded as the 'brainbox' of the family, including one more rendering of Davenport's accent, and ending with his regrets about never marrying. Some people already

knew that these regrets were not abstract but concerned specific ladies. I quote again from the funeral tribute:

> 'I first met Alan in 1975 playing table tennis in the Graduate Centre. I was immediately taken by his generosity of spirit. Unlike some of the others, he never minded being beaten. He proved to be equally generous and attentive as a suitor, and later as an old friend.'

Alan's last years were made more difficult by increasing deafness and a series of falls (he did not pay serious attention to medical advice on these). He had long since advanced from Whewell's Court to Great Court and indeed had been proud to have the very best set of rooms there, like a maisonette on two floors. It overlooked the Bowling Green where he now could no longer play (see the cover of ⟨**13**⟩, the proceedings of the Zürich conference).

Alan Baker died on 4 February 2018 in Cambridge after a severe stroke a few days earlier.

During a Feast at Trinity given in his memory, he was described as 'idiosyncratic' in place of words like 'eccentric' or 'enigmatic'.

Other biographical articles about Alan Baker can be found in the *Hardy–Ramanujan Journal* ⟨**1**⟩. See also ⟨**14**⟩ for a scientific appraisal by Wüstholz, which also contains a complete list of his publications.

2. *Mathematics: a preview*

For a more structured narrative, we shall divide Alan Baker's work into eight categories (and in scientific detachment usually drop the first name).

(a) Diophantine approximation.
(b) Linear forms in logarithms.
(c) Diophantine equations.
(d) Elliptic functions.
(e) Class numbers.
(f) *Abc*ology.
(g) Miscellaneous.
(h) Books.

But before starting, we would like to mention perhaps the most easily stated of all his deep results.

When we make a list of the perfect squares

$$1, 4, 9, 16, 25, 36, 49, 64, 81, 100, 121, 144, \ldots, 143384152921, \ldots,$$

we see that the gaps between consecutive members get larger and larger (and in a regular way). Similarly for the perfect cubes

$$1, 8, 27, 64, 125, 216, 343, 512, 729, 1000, 1331, 1728, \ldots, 143384152904, \ldots.$$

But, if we mesh the two lists together to obtain the 'squbes'

$$1, 4, 8, 9, 16, 25, 27, 36, 49, 64, 81, \ldots, 143384152904, 143384152921, \ldots,$$

then it is not so clear that the gaps become large. Indeed this was proved by Mordell only in 1922; thus, for example, the gap 17 (twice above) or the gap 1621 (to take a year apparently at random) occurs at most finitely often. Unfortunately Mordell's proof gave no way of determining all the occurrences of a given gap.

To do this amounts to specifying a non-zero k in the set $\mathbf{Z}$ of rational integers and finding all x, y in $\mathbf{Z}$ with

$$y^2 = x^3 + k. \tag{2.1}$$

Baker achieved this in 1968 by showing that they all satisfy

$$\max\{|x|, |y|\} \leqslant \exp\{(10^{10}|k|)^{10000}\}. \tag{2.2}$$

Despite (2.1) being around since at least the year 1621 (and the cases $k = -2, -4$ were set by Fermat in 1657 as a challenge to 'you English'), there were no estimates at all for x, y until (2.2) nearly 350 years later.

Thus, for example, to find all gaps 1621, one just has to examine all y^2 and x^3 with y, x between 1 and say

$$10^{10^{132098}}.$$

This looks hopelessly impractical; yet we shall see later that Baker (with Davenport) found exceedingly efficient ways to do such things.

There is an attractive single-sentence reformulation: for all positive integers x, y with $x^3 \neq y^2$, we have

$$|x^3 - y^2| > 10^{-10}(\log x)^{1/10000}.$$

3. *Diophantine approximation*

It is classical that

$$\pi = 3.1415926\ldots \quad \text{and} \quad \frac{355}{113} = 3.1415929\ldots$$

are suspiciously close. This raises the natural question: given a real number ξ, how well can we approximate it by a rational number p/q? An answer in convenient form was found by Dirichlet: provided ξ is not already rational, we can find infinitely many p/q with

$$\left|\xi - \frac{p}{q}\right| < \frac{1}{q^2}. \tag{3.1}$$

We pause to give the simple proof, which involves the Box Principle or Pigeonhole Principle. Pick any $Q \geqslant 1$ in $\mathbf{Z}$. For $i = 0, 1, \ldots, Q$ we can find p_i in $\mathbf{Z}$ with the $Q + 1$ pigeons $\theta_i = i\xi - p_i$ in the interval from 0 to 1. We divide this interval into Q holes of length $1/Q$. There are more pigeons than holes, so at least one hole must contain at least two pigeons. With say θ_j and θ_k ($j < k$), this leads at once to

$$|q\xi - p| \leqslant \frac{1}{Q} \tag{3.2}$$

for $p = p_k - p_j$ and $q = k - j$ satisfying $1 \leqslant q \leqslant Q$. And now (3.1) follows; it is not too hard to see that we get infinitely many p/q as Q varies.

But the answer to the next natural question of whether we can beat (3.1) depends critically on the number ξ.

For a class of ξ particularly interesting to number-theorists, this topic can be said to have begun in earnest with Liouville in 1844, although it underlies the older concept of continued fractions such as

$$3 + \frac{1}{7 + \frac{1}{16}} = \frac{355}{113},$$

or infinite ones like

$$3 + \cfrac{1}{7 + \cfrac{1}{15 + \cfrac{1}{1 + \cfrac{1}{292 + \cdots}}}} = \pi.$$

More generally,

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{a_4 + \cdots}}}} \tag{3.3}$$

is usually linearized to just

$$x = [a_0; a_1, a_2, a_3, a_4, \ldots]. \tag{3.4}$$

For example,

$$\alpha = [1; 1, 1, 1, 1, \ldots] = \frac{1 + \sqrt{5}}{2}, \tag{3.5}$$

and this is actually an algebraic number in that $\alpha^2 - \alpha - 1 = 0$. On the other hand, Liouville's work (see later) shows that, if the positive integers $a_0, a_1, a_2, a_3, a_4, \ldots$ increase very rapidly, then x in (3.4) does not satisfy any equation

$$b_0 x^d + b_1 x^{d-1} + \cdots + b_d = 0 \tag{3.6}$$

for $b_0, b_1, \ldots, b_d$ in **Z** not all zero. Thus by definition x is a transcendental number.

In 1906, Maillet had given a different sort of transcendental continued fraction. A typical example is to take (3.5) and replace 1 by 2 in the positions $k_1, k_2, \ldots$, where now these $k_1, k_2, \ldots$ increase very rapidly.

In his very first paper [**1**] from 1962, Baker simplified and improved that work and also made the estimates more explicit. A consequence here is that it suffices to take k_n as small as 4^n for transcendence. For the proof, one notes that something like

$$x = [1; 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 1, \ldots]$$

is rather close to α above, with similar approximations further along. In this case, an easy generalization of Liouville is applicable, but for other examples it is necessary to use Klaus Roth's 'revolutionary improvement' (Cassels) of Liouville's result, or more precisely a consequence by Davenport and Roth, and even LeVeque's generalization of Roth.

Acknowledgements such as Baker's:

> 'I should like to thank Professor Davenport for his valuable suggestions and help in preparing the manuscript.'

can be seen quite often around this time in the journal *Mathematika* (which Davenport founded).

The paper [**4**] from 1964 can be considered as a sort of continuation of [**1**]. To describe some of its results, we must recall that what Liouville proved is that, for any algebraic number α of degree $d \geqslant 2$ (the smallest integer such that α satisfies an equation (3.6) above), there is $c > 0$ such that

$$\left| \alpha - \frac{p}{q} \right| \geqslant \frac{c}{q^d} \tag{3.7}$$

for all p and $q \geqslant 1$ in **Z**. When $d \geqslant 3$, Roth improved this to

$$\left| \alpha - \frac{p}{q} \right| \geqslant \frac{c}{q^\kappa} \tag{3.8}$$

for any $\kappa > 2$, where now c is allowed to depend on κ. The Box Principle as in (3.1) shows that this is essentially best possible.

Equivalently, if, for some ξ and $\kappa > 2$, there is an infinite sequence of p_i and $q_i \geqslant 1$ in $\mathbf{Z}$ with p_i/q_i different and

$$\left| \xi - \frac{p_i}{q_i} \right| < \frac{1}{q_i^{\kappa}} \qquad (i = 1, 2, \ldots), \tag{3.9}$$

then ξ must be transcendental.

Baker then shows under an additional condition that ξ cannot be too close to an algebraic number in the sense of what is since called a 'transcendence measure'. The condition is that there should exist λ with $q_{i+1} \leqslant q_i^{\lambda}$ for all i. In that case, he shows that, for each n, there is μ_n (possibly depending on ξ) such that

$$|h_0 \xi^n + h_1 \xi^{n-1} + \cdots + h_n| > H^{-\mu_n} \tag{3.10}$$

for all $h_0, h_1, \ldots, h_n$ in $\mathbf{Z}$ not all zero, where $H = \max\{2, |h_0|, |h_1|, \ldots, |h_n|\}$.

The significance of this is the following. From (3.10) by definition (due to Kurt Mahler), ξ cannot be a so-called U-number, even though (3.9) implies that the 'partial quotients' $a_0, a_1, \ldots$ in (3.4) for $x = \xi$ are unbounded. On the other hand, the partial quotients for sums of the form $\sum_{k=1}^{\infty} 2^{-k!}$ are also unbounded and it is a U-number (for example, (3.10) fails already for $n = 1$ and $h_0 = 2^{k!}$ with k large). Furthermore, Baker shows that the method of [1] produces U-numbers with bounded partial quotients, and that a suitable generalization of the above result to quadratic fields produces ξ that are not U-numbers, but still have bounded partial quotients.

More succinctly, there is no correlation between the properties of having bounded partial quotients and being U-numbers. In fact, Baker is able to sharpen this to involve (also in Mahler's classification) T-numbers and S-numbers (see just below), and that was his main motivation. A spin-off is that either T-numbers exist or S-numbers 'of type exceeding 1' exist. This foreshadows Wolfgang Schmidt's breakthrough four years later showing that T-numbers exist.

The proofs are rather formidable; indeed (3.9) and (3.10) lie close to a strengthening of Roth's Theorem, and accordingly Baker has to ramp up Roth's entire machinery, even in the situation of LeVeque's generalization.

This time he writes only:

> 'I am indebted to Professor Davenport for valuable suggestions in connection with the present work.'

In [15] from 1967, Baker returned to these themes, with generalizations to several numbers, at least in the case of bounded partial quotients. This property for a single ξ is equivalent to $|\xi - p/q| \geqslant c/q^2$ analogous to (3.7), and the natural extension to a pair (ξ_1, ξ_2) is

$$\max\left\{ \left| \xi_1 - \frac{p_1}{q} \right|, \left| \xi_2 - \frac{p_2}{q} \right| \right\} \geqslant \frac{c}{q^{3/2}}$$

for all p_1, p_2 and $q \geqslant 1$ in $\mathbf{Z}$. He shows, for example, that there are U-numbers ξ_1, ξ_2 satisfying this, and also ξ_1, ξ_2 that are not U-numbers. This time the proofs need no Roth-type considerations but follow a still intricate 'interval nesting' technique of Cassels and Davenport.

In 1966, Baker [10] investigated certain 'metrical' properties of S-numbers. These are defined by refining (3.10) as follows. There is an ω such that for each n and each $\kappa > n\omega$, we can find $c > 0$ (possibly depending on ξ, n, κ) such that

$$|h_0 \xi^n + h_1 \xi^{n-1} + \cdots + h_n| > cH^{-\kappa}, \tag{3.11}$$

as in (3.10). For simplicity, we consider only real ξ. Much as before, the Box Principle shows that $\omega \geqslant 1$. It had been conjectured by Mahler in 1932 that we can in fact take $\omega = 1$ ('of type exactly 1') for almost all ξ in the sense of Lebesgue measure. This was proved by Sprindzhuk

in 1965. Having seen only preliminary announcements of that result, Baker was able to refine (3.11) even further to things like

$$|h_0\xi^n + h_1\xi^{n-1} + \cdots + h_n| > cH^{-n}(\log H)^{-\lambda} \tag{3.12}$$

for any $\lambda > n$.

Later on in 1970, he considered with Schmidt [29] further refinements in terms of Hausdorff dimension; these are closer to Koksma's classification into S^*-, T^*-, and U^*-numbers based instead on the distance from ξ to algebraic numbers.

Many of the results above can be expressed more concisely through the notation

$$\|x\| = \min_{m \in \mathbf{Z}} |x - m|$$

for x in the field $\mathbf{R}$ of real numbers, meaning the distance to the nearest integer. For example, (3.7) says that $\|q\alpha\| \geqslant c/q^{d-1}$ for all $q \geqslant 1$.

This notation can be used in other contexts. Returning to 1964, we may cite Baker's nice note [3], which shows that

$$|Q|\|Q\Theta_1\|\|Q\Theta_2\| \geqslant e^{-5} \tag{3.13}$$

for any non-zero polynomial Q in say $\mathbf{R}[t]$, where $\Theta_1 = e^{1/t}$ and $\Theta_2 = e^{2/t}$ (same e, but that is irrelevant) are interpreted as formal power series in the topological completion with respect to the valuation $|t| = e$, and

$$\|X\| = \min_{M \in \mathbf{R}[t]} |X - M|$$

measures the distance to the nearest polynomial. The significance here is that the analogue of (3.13) with $\mathbf{R}[t]$ replaced by $\mathbf{Z}$ (see just below) is thought to be false, by a famous conjecture of Littlewood (still unsolved). The note makes explicit an earlier result of Davenport and Lewis and provides a simpler proof, essentially by differentiating and using a Padé approximation (see later in this section). In fact, the Padé element can be eliminated to give a yet simpler proof as follows.

It can be checked that

$$\Delta = \begin{vmatrix} Q & P_1 & P_2 \\ t^2Q' & t^2P_1' + P_1 & t^2P_2' + 2P_2 \\ t^4Q'' + 2t^3Q' & t^4P_1'' + (2t^3 + 2t^2)P_1' + P_1 & t^4P_2'' + (2t^3 + 4t^2)P_2' + 4P_2 \end{vmatrix}$$

remains unchanged on replacing P_1, P_2 by $\mathcal{E}_1 = P_1 - Qe^{1/t}$ and $\mathcal{E}_2 = P_2 - Qe^{2/t}$, respectively.

Now suppose that Q, P_1, P_2, even in $\mathbf{C}[t]$, are all non-zero. Inspecting the coefficients of smallest powers of t in each entry shows that $\Delta \neq 0$. Thus $|\Delta| \geqslant 1$.

Next, for $Q \neq 0$, choose P_1, P_2 so that $|\mathcal{E}_1| = \|Qe^{1/t}\| < 1$ and $|\mathcal{E}_2| = \|Qe^{2/t}\| < 1$. Clearly, $P_1 \neq 0, P_2 \neq 0$. Using $|X'| \leqslant e^{-1}|X|$, we easily obtain that $1 \leqslant |\Delta| \leqslant e^3|Q||\mathcal{E}_1||\mathcal{E}_2|$ giving a slight improvement of the result.

He evidently had this note in mind when writing [7] in 1965, back in $\mathbf{R}$. A special result there is that, for any $\varepsilon > 0$, there is $c > 0$ such that

$$q\|q\theta_1\|\|q\theta_2\| \geqslant \frac{c}{q^\varepsilon} \tag{3.14}$$

for any $q \geqslant 1$ in $\mathbf{Z}$, where $\theta_1 = e, \theta_2 = e^2$.

More generally Baker treats a product of several terms (as indeed in [3]) involving various numbers $\theta = e^\varphi$ with (different) rational φ (that then implies quite easily the transcendence of e with quite a good measure); and he is even able to refine the q^ε. Of course one can no longer do things just by differentiation; but that does play a role, and the much more elaborate proof uses techniques introduced by Carl Ludwig Siegel in his work on the so-called E-functions, together with so-called transference principles.

The paper [**16**] from 1967 is Baker's only excursion into the Hardy–Littlewood Circle Method. Here one wants to know that certain approximations actually exist, and this time with prime values of the variables. For example, he shows that, given any κ, there are infinitely many primes p_1, p_2, p_3 with $|p_1 - \sqrt{2}p_2 - \sqrt{3}p_3| < (\log P)^{-\kappa}$, where $P = \max\{p_1, p_2, p_3\}$ (with extension to arbitrary coefficients). This was the first such result with an explicit function of P on the right-hand side. As Bob Vaughan (who later greatly improved the result, to Baker's surprise and dismay) has pointed out, he was also aware that a 'localized' version in the style of (3.2), with P given in advance and just $P \geqslant \max\{p_1, p_2, p_3\}$, is in principle impossible. We may remark that the result in the enigmatic footnote on the second page of this paper [**16**] had previously been proved by Vinogradov.

Returning again to 1964, we examine [**5**] in some detail. The starting point is again (3.7) and (3.8). There is a fundamental difference between these two results. Namely, given α in (3.7), one can calculate an explicit value of c (and in fact rather easily). One says informally that (3.7) is 'effective'.

In fact not much attention had been paid to this concept, which dramatically increased in importance thanks to the works of Alan Baker.

On the other hand, given α and κ in (3.8), no-one knows how to calculate any value of c. This is true even for the simple-looking

$$\left|2^{1/3} - \frac{p}{q}\right| \geqslant \frac{c}{q^{2.1}}. \tag{3.15}$$

The obstacle is a logical twist already present in Thue's improvement on (3.7) in 1909. Namely, his argument works very well if there happens to be already a very good rational approximation $p_?/q_?$ to α. This sounds paradoxical, but it has the effect of 'repelling' other $p/q \neq p_?/q_?$ through the triangle inequality

$$\left|\alpha - \frac{p}{q}\right| \geqslant \left|\frac{p_?}{q_?} - \frac{p}{q}\right| - \left|\alpha - \frac{p_?}{q_?}\right| \geqslant \frac{1}{q_? q} - \left|\alpha - \frac{p_?}{q_?}\right|. \tag{3.16}$$

And, if there are no such good approximations $p_?/q_?$, that is also good news. However, we have no way of resolving this dichotomy; and in the first case, c will depend on the uncontrollable $p_?/q_?$.

In [**5**], Baker treats some numbers of this special form $\alpha = (a/b)^{1/d}$ for positive a, b in $\mathbf{Z}$, as in (3.15). He proves (3.8), but with different quantifiers. Namely, for each such α, he finds an explicit κ such that (3.8) holds for some effective $c > 0$. Of course, this already holds for $\kappa = d$ by (3.7); but the new κ can sometimes be less than d. Furthermore, and this is Baker's key observation, one can obtain κ arbitrarily close to 2 by choosing α carefully, usually with a, b approximately equal. This amounts to imposing $p_?/q_? = 1$ in advance. For example, if $b = 3^{316}$, then

$$\left|\left(\frac{b+1}{b}\right)^{1/3} - \frac{p}{q}\right| \geqslant \frac{10^{-1024522}}{q^{2.1}}, \tag{3.17}$$

which is effective 'but only just' (as Estermann's Scotsman ⟨**5**⟩ said on p. 33); and not as elegant as (3.15).

Since Thue's work it is known that any $\kappa < d$ leads to consequences for diophantine equations, and Baker presents some of these. More about that just below. The proofs follow the so-called hypergeometric method used by Thue in his earlier paper of 1908 (also see just below).

Actually Thue in a later paper from 1918 had implicitly (in terms of diophantine equations) obtained results like (3.17). His work implies, for example, an effective $|\alpha - p/q| \geqslant cq^{-6.668}$ for $\alpha = 17^{1/7} = 1.4989\ldots$ thanks to $p_?/q_? = 3/2 = 1.5$; this a little more elegant than (3.17).

But the ultimate in elegance was then achieved by Baker [**6**], also in 1964: namely

$$\left|2^{1/3} - \frac{p}{q}\right| > \frac{10^{-6}}{q^{2.955}}. \tag{3.18}$$

The consequence for diophantine equations is that all solutions x, y in $\mathbf{Z}$ of the equation

$$x^3 - 2y^3 = m \tag{3.19}$$

satisfy

$$\max\{|x|, |y|\} \leqslant (300000|m|)^{23}. \tag{3.20}$$

Even though equations like (3.19) have been around since at least Thue's work, this was the first time that (3.19) itself could be solved even in principle. It is similar to the situation around (2.1) and (2.2) above.

Thus to solve completely, for example, the equation

$$x^3 - 2y^3 = 1621, \tag{3.21}$$

we have to check only all x, y with absolute value at most say 10^{200}. And with modern computers this is entirely feasible, using a trick found by Baker and Davenport a few years later (see Section 5).

Actually to prove these results, one needs infinitely many elements like $p_?/q_?$ in the set $\mathbf{Q}$ of all rationals. An old idea, going back essentially to Hermite, for obtaining many approximations p/q in $\mathbf{Q}$ to a number ξ in $\mathbf{R}$ is to construct many approximations A/B in $\mathbf{Q}(x)$ (preferably of so-called Padé type) to a suitable function $f(x)$. Then one specializes x.

The choice $f(x) = (1-x)^{1/3}$ will serve for (3.18). Thue had already used this, and found that hypergeometric functions turn up. Baker succeeded with (3.18) by means of a nice new twist.

By simple linear algebra, we see that, for any $r \geqslant 0$ in $\mathbf{Z}$, there are polynomials $A_r(x), B_r(x)$, not both zero, such that

$$\varphi_r(x) = A_r(x) - (1-x)^{1/3}B_r(x) \tag{3.22}$$

has a zero at $x = 0$ of order at least $2r + 1$ (these are in fact Padé approximations). They are unique if we normalize $B_r(0) = 1$, and then

$$B_r(x) = F\left(\tfrac{1}{3} - r, -r, -2r, x\right) = \sum_{j=0}^{r} b_{rj} x^j$$

for the Gauss function F (now a polynomial). Here the coefficients

$$b_{rj} = \prod_{k=0}^{j-1} \frac{((1/3) - r + k)(-r + k)}{(1 + k)(-2r + k)}$$

are all rational, and seem to involve 3^j in their denominators.

As $2^{1/3} = 1.2599\ldots$ the choice $p_?/q_? = 5/4$ is tempting in the discussion above; and this translates into putting $x = 3/128$ in (3.22), when $(1-x)^{1/3} = \frac{5}{4}/2^{1/3}$. As $\varphi_r(x) = cx^{2r+1} + \cdots$ is small, it seems plausible that we obtain a sequence of good rational approximations

$$\frac{5}{4} \frac{B_r(3/128)}{A_r(3/128)} \quad (r = 0, 1, 2, \ldots) \tag{3.23}$$

to $2^{1/3}$. For example, with $r = 0, 1, 2, 3, 4$, we obtain

$$\frac{5}{4}, \frac{635}{504}, \frac{96389}{76504}, \frac{15240955}{12096754}, \frac{26990767415}{21422586294}$$

(the last being accurate to 19 decimal places). The hope is that these will repel all others as in (3.16). But to check this we have to estimate denominators in (3.23). For example,

$$B_r\left(\frac{3}{128}\right) = \sum_{j=0}^{r} b_{rj}\left(\frac{3}{128}\right)^j;$$

and now we see that the 3^j in the denominators of b_{rj} are cancelled out. It is this '3-adic' feature that is Baker's new twist (which led later to a minor industry). It would fail for the slightly more tempting $p_?/q_? = 63/50 = 1.26$.

In the same paper, he did similar things with $a^{1/3}$ for $a = 17, 19, 20, 37$ and 43. But, at least before 1988, the method fails for $a = 5$, and so does not lead to a complete solution of, say, $x^3 - 5y^3 = 1621$ (see Section 4).

In all, Baker published no fewer than five papers in 1964, each highly non-trivial (he did even better with six in 1967). We have already discussed four of them. The fifth is [**2**], which follows the broad principles of [**5, 6**] above. Now he is dealing with numbers $\xi = \log \alpha$ whose irrationality (or transcendence) is by itself somewhat deep. In fact, for algebraic $\alpha \neq 0, 1$, this was proved by Lindemann in 1882 (see just below in Section 4). We indicate just one consequence in the style of (3.17): if $b = 10^{12}$, then

$$\left| \log \left(\frac{b+1}{b} \right) - \frac{p}{q} \right| \geqslant \frac{10^{-121506}}{q^{2.1}}.$$

But he is also able to extend to transcendence measures in the style of (3.11), with κ arbitrarily near n in the spirit of (3.12). In the proofs, the analogue of (3.22) is a system of forms

$$\varphi(x) = A^{(0}(x) + A^{(1)}(x) \log(1 - x) + \cdots + A^{(n)}(x)(\log(1 - x))^n \qquad (3.24)$$

found by Mahler.

This category of Baker's work is rounded off by [**17**] in 1967, which extends [**5**] to linear forms in various $(a/b)^{d_0/d}, (a/b)^{d_1/d}, \ldots, (a/b)^{d_n/d}$, again close to best possible in the spirit of (3.12). Instead of quoting an approximation result, we mention an attractive application to diophantine equations. Namely, if $h = 10^{11} + 1$ and $l = h^5 - 1$, then all solutions x, y, z in $\mathbf{Z}$ of the equation

$$x^5 + ly^5 + l^2 z^5 + 5\,lxyz(xz - y^2) = m \qquad (3.25)$$

satisfy $\max\{|x|, |y|, |z|\} \leqslant l^{500} m^2$. This may not be quite so elegant as (3.19) and (3.20), but the exponent of m is much smaller, and crucially this seems to be the first such example ever for equations in three variables (where at the time of writing it is unknown whether $x^3 + y^3 + z^3 = 114$ (see $\langle \mathbf{16} \rangle$) has any solutions at all). Here another system like (3.24), also due to Mahler, plays a key role.

This concludes the section on diophantine approximation. Baker did more very important work around (3.7), but that deserves an extra section, which now follows. In the sequence of papers just described, it seems that he is exuberantly flexing his muscles on several different methods, (with hindsight) limbering up for his big achievements. In the early stages, he was possibly solving every problem Davenport threw at him, but pretty soon his own motivation took over.

4. *Linear forms in logarithms*

This title hardly existed before Alan Baker.

The classical theorem of Hermite–Lindemann (1873–1882) is equivalent to the fact that, if α is a non-zero algebraic number, and $\log \alpha$ is any non-zero choice of its complex logarithm, then 1 and $\log \alpha$ are linearly independent over the field $\overline{\mathbf{Q}}$ of all complex algebraic numbers. This includes the transcendence of e, π and so of

$$\int_0^1 \frac{\mathrm{d}x}{x^2 + 1} = \frac{\pi}{4},$$

and of course the numbers $\log \alpha$ mentioned in Section 3, as well as the numbers e^β for algebraic $\beta \neq 0$.

Similarly the classical theorem (1929–1934) of Gelfond and Schneider is equivalent to the fact that, if α_1, α_2 are non-zero algebraic numbers, and $\log \alpha_1, \log \alpha_2$ are any choices of logarithms which are linearly independent over $\mathbf{Q}$, then they are also linearly independent over $\overline{\mathbf{Q}}$. This includes the transcendence of e^{π} and most notably the numbers $2^{\sqrt{2}}$ and α^{β} mentioned by Hilbert (Seventh Problem) in his famous 1900 address to the International Congress of Mathematicians in Paris, as well as numbers like $\log 3 / \log 2$.

Before Baker, nothing was known about the independence of $1, \log \alpha_1, \log \alpha_2$; and practically nothing about that of $\log \alpha_1, \log \alpha_2, \log \alpha_3$, despite the assertions of Gelfond and Linnik in 1948 (repeated in Gelfond's 1960 book) of the enormous importance of the latter problem (for which see later). There had been a paper about $b_1 \log \alpha_1 + b_2 \log \alpha_2 + b_3 \log \alpha_3$ for integers b_1, b_2, b_3 by Gelfond and Feldman in 1949; but these coefficients were subject to a restrictive condition which essentially reduced the problem to two logarithms.

Then, in a sequence of four papers in 1966–1968 (which we shall call the 'Linear Forms Quartet'), Baker went straight for any number of logarithms and proved (among much else) the following.

THEOREM (Baker). *If $\alpha_1, \ldots, \alpha_n$ are non-zero algebraic numbers, and $\log \alpha_1, \ldots, \log \alpha_n$ are any choices of logarithms which are linearly independent over $\mathbf{Q}$, then*

$$1, \log \alpha_1, \ldots, \log \alpha_n \tag{4.1}$$

are linearly independent over $\overline{\mathbf{Q}}$.

The reader may easily construct simple examples of transcendental numbers like $e^{\beta_0} \alpha^{\beta}$ or $\alpha_1^{\beta_1} \alpha_2^{\beta_2}$ not covered by Hermite–Lindemann or Gelfond–Schneider; less simple examples are

$$\int_0^1 \frac{\mathrm{d}x}{x^3 + 1} = \frac{\pi\sqrt{3}}{9} + \frac{\log 2}{3} \tag{4.2}$$

quoted by Siegel in his famous 1949 transcendence monograph, and

$$\int_0^1 \frac{\mathrm{d}x}{x^3 - 3x - 1} = \beta_1 \log \alpha_1 + \beta_2 \log \alpha_2 + \beta_3 \log \alpha_3 ,$$

where

$$\alpha_1 = 4 - \alpha^2, \ \alpha_2 = 2 + \alpha, \ \alpha_3 = -\alpha + \alpha^2,$$

$$\beta_1 = \frac{-4 - \alpha + 2\alpha^2}{9}, \ \beta_2 = \frac{2 - \alpha - \alpha^2}{9}, \ \beta_3 = \frac{2 + 2\alpha - \alpha^2}{9} ,$$

and $\alpha^3 - 3\alpha - 1 = 0$ (with say $\alpha > 0$ to provide unique real choices of logarithms).

We proceed to examine these four papers I, II, III, IV in detail.

Linear Forms I is [**11**] from 1966. It is one of the great number-theoretic papers of the twentieth century.

In fact, the main result is a 'linear independence measure' (analogous to transcendence measure in Section 3) for just $\log \alpha_1, \ldots, \log \alpha_n$ with a condition slightly stronger than that in the above theorem. We shall return to such measures later; it is these refinements that are needed for the applications suggested by Gelfond and Linnik.

In an exemplary display of modesty, clarity, and foresight, Baker writes:

> *'Finally, as regards the proof of the theorem, our method depends on the construction of an auxiliary function of several complex variables which would seem to be the natural generalisation of the function of a single variable used in Gelfond's original work. The subsequent treatment employed by Gelfond, however, is not applicable in the more general context and so it has been necessary to devise a new technique.*

> *Nevertheless it will be appreciated that the argument involves many familiar ideas. The method will probably be capable of considerable development for it applies in principle to many other auxiliary functions apart from the one constructed here.'*

This is immediately followed by:

> *'The author is grateful to Prof. H. Davenport, who read the original draft of this paper, for his helpful criticism.'*

which makes it clear that Davenport did not suggest the problem. Indeed the problem, despite its great importance, seems not to have been very well known; for example, there is no mention of it in Lang's book on transcendental numbers, which appeared in the same year 1966. It is not referred to in the section 'Einige offene Fragestellungen' in Schneider's book of 1957; this section contains his celebrated Eight Problems. And in Siegel's book from 1949 there is no mention of logarithms in connection with (4.2). Also in Baker's paper there is no serious reference to any paper published after 1948. He sought out the problem himself, despite all previous transcendence breakthroughs coming through French, Russian, and German sources; and he solved it in spectacular style.

Let us give an idea of his 'new technique'. We start with $n = 2$ and Gelfond–Schneider; slightly changing the notation, we have to deduce a contradiction from a relation

$$\beta \log \alpha = \log \alpha' \tag{4.3}$$

with α, α', β algebraic and β irrational. Gelfond constructs an auxiliary function $\Phi(z)$ which is a non-zero polynomial of large degree in

$$e^z, e^{\beta z}. \tag{4.4}$$

Note from (4.3) that the functions (4.4) take algebraic values at all points

$$z = s \log \alpha, \qquad s = 0, 1, 2, \dots, \tag{4.5}$$

and that this is true even of their derivatives, thanks to the differential equation for the exponential function. This enables Φ to be constructed, with algebraic coefficients not all zero, such that

$$\frac{\mathrm{d}^t}{\mathrm{d}z^t} \Phi(s \log \alpha) = 0 \tag{4.6}$$

for all non-negative integers s, t in some large range

$$s \leqslant S, \quad t \leqslant T. \tag{4.7}$$

Thus Φ has many zeros (with multiplicity) inside a large disc in $\mathbf{C}$ (say centred at the origin).

Now well-known analytic techniques such as the Schwarz Lemma (or Maximum Modulus Principle; see (6.1) below) show that Φ must be very small on the whole of this disc, and even on a slightly larger one; and this holds for the derivatives, and even slightly more of them. Then well-known arithmetic techniques show that (4.6) holds for, say,

$$s \leqslant 2S, \quad t \leqslant 2T. \tag{4.8}$$

This step could be then iterated, even indefinitely, to obtain infinitely many zeros. In fact, a single zero of infinite multiplicity, say at $z = 0$, suffices to prove that Φ must be identically zero, and this leads easily to the required contradiction.

How did Baker adapt Gelfond's proof? For simplicity, we take $n = 3$ and ignore the extra 1 in (4.1), so that now we have to deduce a contradiction from a relation

$$\beta_1 \log \alpha_1 + \beta_2 \log \alpha_2 = \log \alpha_3 \tag{4.9}$$

with β_1, β_2 algebraic (under a suitable irrationality condition). The auxiliary function $\Phi(z_1, z_2)$ is now a polynomial of large degree in

$$e^{z_1}, e^{z_2}, e^{\beta_1 z_1 + \beta_2 z_2}, \tag{4.10}$$

which indeed generalize in a fairly natural way (although no-one had previously written them down) Gelfond's $\Phi(z)$ and (4.4). This time (4.9) shows that the functions (4.10) take algebraic values at all points

$$(z_1, z_2) \;=\; (s \log \alpha_1, s \log \alpha_2), \quad s = 0, 1, 2, \dots, \tag{4.11}$$

and this is true also of their partial derivatives. This enables Φ to be constructed, with algebraic coefficients not all zero, such that

$$\frac{\partial^{t_1}}{\partial z_1^{t_1}} \frac{\partial^{t_2}}{\partial z_2^{t_2}} \Phi(s \log \alpha_1, s \log \alpha_2) \;=\; 0 \tag{4.12}$$

for all non-negative integers s, t_1, t_2 in some large range

$$s \leqslant S, \quad t_1 + t_2 \leqslant T. \tag{4.13}$$

Thus Φ has many zeros (still with a natural concept of multiplicity) inside a larger disc now in $\mathbf{C}^2$.

In 1941, Schneider had used several complex variables for transcendence purposes, and in $\mathbf{C}^2$ was able to go further because the set of zeros was part of a structure like $\mathbf{Z}^2$. Later Baker expressed it:

> '... this type of argument requires that the points in question form a Cartesian product, a condition that can apparently be satisfied only with respect to particular multiply-periodic functions.'

But in (4.11) the structure is only $\mathbf{Z}$; and indeed the points all lie on a line. This is too thin a set for the usual type of Schwarz Lemma. Indeed the function $(z_1 \log \alpha_2 - z_2 \log \alpha_1)^T$ has the same zeros as Φ, but it need not be small, even on a disc of radius 1 in $\mathbf{C}^2$. And, to this day, no-one knows how to increase T in (4.13) to $2T$ as in (4.8).

Here comes Baker's decisive innovation. It does seem fairly natural to restrict to the line. But Baker considered also the partial derivatives

$$\Phi_{t_1, t_2} = \Phi_{t_1, t_2}(z) \;=\; \frac{\partial^{t_1}}{\partial z_1^{t_1}} \frac{\partial^{t_2}}{\partial z_2^{t_2}} \Phi(z \log \alpha_1, z \log \alpha_2),$$

however keeping only $t_1 + t_2 \leqslant \frac{1}{2} T$.

Clearly $\mathrm{d}^t \Phi_{t_1, t_2} / \mathrm{d}z^t$ can be expressed as a linear combination of the Φ_{τ_1, τ_2} with

$$\tau_1 + \tau_2 \leqslant \frac{1}{2} T + t.$$

So as long as we keep also $t \leqslant \frac{1}{2} T$, then $\tau_1 + \tau_2 \leqslant T$, and we see from (4.12) that

$$\frac{\mathrm{d}^t}{\mathrm{d}z^t} \Phi_{t_1, t_2}(s) \;=\; 0$$

for

$$s \leqslant S, \quad t \leqslant \frac{1}{2} T.$$

These are very similar to (4.6) and (4.7), firmly back in $\mathbf{C}$.

Now the Schwarz Lemma shows that the Φ_{t_1, t_2} are very small on a slightly larger disc; and so the arithmetic gives us (4.12) for

$$s \leqslant 2S, \quad t_1 + t_2 \leqslant \frac{1}{2} T. \tag{4.14}$$

We can iterate but not indefinitely in any profitable way, due to the loss of multiplicity.

See $\langle \mathbf{8} \rangle$ for an appropriate limerick.

Already this limited iteration was a new sort of difficulty, which Baker overcame by getting just as many zeros as are needed for the contradiction (along the principle that a polynomial of degree D cannot have $D+1$ zeros). Here the non-vanishing of a certain Vandermonde determinant suffices.

Of course all this extends to any number of variables and

$$\frac{\partial^{t_1}}{\partial z_1^{t_1}} \cdots \frac{\partial^{t_n}}{\partial z_n^{t_n}} \Phi(s \log \alpha_1, \ldots, s \log \alpha_n) \;=\; 0. \tag{4.15}$$

So Baker's main achievement was to introduce several complex variables, reduce them to a single variable along a line, introduce the magic trick with multiplicities, and supply the missing estimates for zeros.

Actually he was helped in this last 'zero estimate' by Gelfond himself. In a 1935 paper, Gelfond had obtained a linear independence measure for two logarithms. This amounts to assuming that $\beta \log \alpha - \log \alpha'$ is very small rather than zero as in (4.3). That adds some technical complications, but more significantly it means that here too one cannot iterate (4.8) indefinitely; hence the need for some sort of zero estimate.

To state Gelfond's result, we have at last to introduce heights. Recall that the degree of an algebraic number α is the smallest d in any equation (3.6) satisfied by $x = \alpha$. We can then suppose that the coefficients have no common factor, and we define the height by

$$\mathcal{H}(\alpha) = \max\{|b_0|, |b_1|, \ldots, |b_d|\}$$

(we are using $\mathcal{H}$ to distinguish it from a more modern version that need not concern us here; see, however, the end of this section).

Thus Gelfond was able to show that, for any $\kappa > 5$ (and later for any $\kappa > 3$), there is an effective $c > 0$ depending only on α_1, α_2 and the degrees of β_1, β_2 such that

$$|\beta_1 \log \alpha_1 + \beta_2 \log \alpha_2| \geqslant c e^{-(\log B)^\kappa}, \tag{4.16}$$

where $B = \max\{2, \mathcal{H}(\beta_1), \mathcal{H}(\beta_2)\}$, provided that $\log \alpha_1, \log \alpha_2$ are linearly independent over $\mathbf{Q}$ and their principal values are chosen.

Accordingly, in Linear Forms I, Baker could prove an analogous

$$|\beta_1 \log \alpha_1 + \cdots + \beta_n \log \alpha_n| \geqslant c e^{-(\log B)^\kappa} \tag{4.17}$$

for any $\kappa > n+1$, where now $B = \max\{2, \mathcal{H}(\beta_1), \ldots, \mathcal{H}(\beta_n)\}$, except that now $\log \alpha_1, \ldots, \log \alpha_n$ and $2\pi i$ have to be linearly independent over $\mathbf{Q}$.

In the second paper, Linear Forms II [**12**] from 1967, Baker obtained (4.17) for any $\kappa > 2n+1$ when only $\log \alpha_1, \ldots, \log \alpha$ are linearly independent over $\mathbf{Q}$. This was crucial to an application to an effective improvement of Liouville's inequality (3.7) (see later).

In Linear Forms III [**13**], also from 1967, Baker improved the result of [**12**] for any $\kappa > n$. He also made a start on obtaining lower bounds for the 'inhomogeneous form' which is $\beta_0 + \beta_1 \log \alpha_1 + \cdots + \beta_n \log \alpha_n$. This finally yields the theorem quoted above.

In the last paper of the Quartet, Linear Forms IV [**20**] of 1968, Baker starts by noting that the various positive constants c in his lower bounds, although of course effective, are very small indeed, which could lead to practical computing difficulties in the applications. He gives the following version leading to explicit values of c which are not so small. Suppose that, for some δ with $0 < \delta \leqslant 1$, there exist $b_1, \ldots, b_n$ in $\mathbf{Z}$ with absolute values at most H such that

$$0 < |b_1 \log \alpha_1 + \cdots + b_n \log \alpha_n| < e^{-\delta H}. \tag{4.18}$$

Then

$$H < (4^{n^2} \delta^{-1} d^{2n} \log A)^{(2n+1)^2}, \tag{4.19}$$

where $d \geqslant 4$ is an upper bound for the degrees of $\alpha_1, \ldots, \alpha_n$ and $A \geqslant 4$ is an upper bound for their heights.

Implicit in this paper is the following associated result (see additionally [**42**, pp. 3, 8]), which is also the first of its kind. If $\log \alpha_1, \ldots, \log \alpha_n$ are this time linearly dependent over $\mathbf{Q}$, then there is a 'relatively bounded' relation

$$0 = b_1 \log \alpha_1 + \cdots + b_n \log \alpha_n \tag{4.20}$$

for $b_1, \ldots, b_n$ in $\mathbf{Z}$ with

$$0 < \max\{|b_1|, \ldots, |b_n|\} < (4^{n^2} d^{2n} \log A)^{(2n+1)^2}. \tag{4.21}$$

The idea here was greatly exploited in later work on isogenies between elliptic curves and between abelian varieties and consequences of that (see (10.1) below).

Shortly afterwards, Baker wrote two important papers [**18, 19**] which further improve the Quartet results, but since their main emphasis is on diophantine equations we shall postpone our description to the corresponding section below. It is also [**18**] that contains the first effective improvement on Liouville's inequality (3.7).

Similar remarks hold for his 1971 paper [**31**], published in the esteemed journal *Annals of Mathematics*, in connection with class numbers.

We turn now to the later developments of (4.17). These may seem rather technical, but they are specifically aimed at marvellous applications, to be described in Sections 5 and 7. We start with his paper [**34**] with Stark, also in the *Annals* and from 1971.

Baker had observed already in [**18**] that the term $(\log A)^{(2n+1)^2}$ in (4.19) could be replaced by $C(\log A_n)^\kappa$ for any $\kappa > n + 1$, where $A_n \geqslant 4$ is an upper bound for $\mathcal{H}(\alpha_n)$ alone, and C (effective) is here allowed to depend on δ, κ, d as well as $\alpha_1, \ldots, \alpha_{n-1}$. In [**34**], he and Stark replace (4.18) by

$$0 < |\beta_1 \log \alpha_1 + \cdots + \beta_n \log \alpha_n| < e^{-\delta H}, \tag{4.22}$$

where now $\beta_1, \ldots, \beta_n$ are algebraic numbers also of degree at most d. The analogue of (4.19) is

$$H < \max\{e^{\sqrt{\log B}}, C'(\log A_n)^\kappa\}, \tag{4.23}$$

now for any $\kappa > 1$, where $B \geqslant 4$ is an upper bound for the heights of $\beta_1, \ldots, \beta_n$ and C' has the same nature as C above.

This indeed technical result could be applied to class number problems (see Section 7), and also to general effective improvements of Liouville's inequality (3.7), as described later on. But actually much interest attaches to the proof. Already in [**27**] (see Section 6), Baker had the nice idea of taking s in (4.5) and (4.11) as rational not just integral, but now he and Stark combine this with another important idea, which for simplicity we illustrate with (4.5). Namely, the functions (4.4) evaluated at $z = (1/p) \log \alpha$ take the values $\alpha^{1/p}, \alpha'^{1/p}$, and the vanishing of Φ there implies an algebraic relation between these values. Such relations can be ruled out by Kummer theory (based on Galois theory) if p is, for example, a sufficiently large prime (thus $1, 2^{1/p}, 4^{1/p}$ are linearly independent over $\mathbf{Q}$ for $p > 2$); this gives another approach to the zero estimates.

It is not difficult to guess which author wrote the following in a footnote:

> 'However, the latter work involves an appeal to Galois theory which we prefer to avoid.'

We now come to the 'Sharpening Trio' I, II, III (changed at the last moment from 'strengthening').

In Sharpening I [**35**] (in memory of Davenport and Sierpiński) from 1972, Baker is able to improve (4.23) for $\delta = 1$ to $H < C'' \log B \log A_n$ provided that $\beta_1, \ldots, \beta_n$ are in $\mathbf{Z}$, where C''

is allowed to depend only on $d, \alpha_1, \ldots, \alpha_{n-1}$. Effective improvements on Liouville have now become the main motivation, and a little competition opens up between Baker and Feldman. Indeed, the much sharper dependence on B (now essentially best possible) comes from a 1968 paper ⟨**6**⟩ of Feldman (who realized how to avoid certain factorials — Thue himself had to do this as well). Baker concedes in a footnote that Feldman had recently obtained those effective improvements (for which having $\kappa = 1$ in (4.23) is essential), but immediately adds that he can do it himself with a 'slight generalization' of this paper [**35**].

That slight generalization is Sharpening II [**36**], from 1973 (dedicated to Siegel). It contains a partial improvement of (4.19) when $b_n = -1$ in (4.18), namely: $H < C''' \log A_n$, where the ever-effective C''' is allowed to depend only on $d, \delta, \alpha_1, \ldots, \alpha_{n-1}$. Indeed this does do the trick for Liouville, and it is customary to declare the competition a draw. Thus we have:

THEOREM (Baker–Feldman). *Given any algebraic number α of degree $d \geqslant 3$, there is effective $\kappa < d$ and effective $c > 0$ such that*

$$\left| \alpha - \frac{p}{q} \right| \geqslant \frac{c}{q^\kappa} \tag{4.24}$$

for all p and $q \geqslant 1$ in $\mathbf{Z}$.

Regarding this κ, we pause to mention the 1988 work [**44**] of Baker and Stewart. They gave explicit values of κ and c for every irrational $\alpha = a^{1/3}$ with $a \geqslant 2$ in $\mathbf{Z}$. We already mentioned a problem with $a = 5$ in connection with (3.18); this they solve with

$$\left| 5^{1/3} - \frac{p}{q} \right| > \frac{10^{-12900}}{q^{2.9999999999998}}$$

(which would have left Estermann's Scotsman without words). The proofs use cubic units (here $41 + 24\alpha + 14\alpha^2$) and an ad hoc linear form in three logarithms.

The Trio was interrupted with [**37**] in 1973. Here Baker introduces the now-familiar notation

$$\Lambda = \beta_0 + \beta_1 \log \alpha_1 + \cdots + \beta_n \log \alpha_n \tag{4.25}$$

(back to inhomogeneous) and shows that, if $\Lambda \neq 0$, then

$$|\Lambda| > (B \log A_n)^{-C'''' \log A_n} , \tag{4.26}$$

with A_n as above, and now $B \geqslant 4$ is an upper bound for the heights of $\beta_0, \beta_1, \ldots, \beta_n$, and C'''' is allowed to depend effectively on $d, \alpha_1, \ldots, \alpha_{n-1}$. This starts to look more like the modern versions. Baker gives the application

$$\left| e^\pi - \frac{p}{q} \right| \geqslant \frac{1}{q^{C \log \log q}}$$

for all p and $q \geqslant 4$ in $\mathbf{Z}$, with C absolute (that is, depending on no additional parameters); this has never been improved.

The Trio ends with Sharpening III [**39**] (in memory of Linnik and Mordell) from 1975. There is now a mini-competition with Stark. Baker shows that when $\beta_0 = 0$ and $\beta_1, \ldots, \beta_n$ are in $\mathbf{Z}$ in (4.25), still with $\Lambda \neq 0$, then

$$|\Lambda| > B^{-C\Omega \log \Omega}. \tag{4.27}$$

Here $\Omega = \log A_1 \cdots \log A_n$, where each $A_i \geqslant 4$ is an upper bound for the height $\mathcal{H}(\alpha_i)$. Things seem to be getting more complicated, but now the effective C depends only on n and d, and so we are heading back in the direction of totally explicit estimates like (4.19). And the temporary neglect of $\alpha_1, \ldots, \alpha_{n-1}$ has ended (this had been done by Stark two years earlier).

Baker gives no applications, but a footnote 'added in proof' mentions that the result had 'recently' been used by Tijdeman to deliver what is perhaps the most spectacular application of linear forms in logarithms to diophantine equations; more of that in the next section.

With [**42**] from 1977, we are finally back to totally explicit estimates. After a detailed account of the history (up to then), Baker shows first that for $\Lambda \neq 0$ as in (4.25), we have

$$|\Lambda| > (B\Omega)^{-C_0 \Omega \log \Omega'}, \tag{4.28}$$

with $\Omega' = \Omega / \log A_n$ and $C_0 = (16nD)^{200n}$; now $D = [\mathbf{Q}(\alpha_1, \ldots, \alpha_n, \beta_0, \beta_1, \ldots, \beta_n) : \mathbf{Q}]$ measures degrees in a different way. Second he shows that, when $\beta_0 = 0$ and $\beta_1, \ldots, \beta_n$ are in $\mathbf{Z}$, the bracketed factor Ω in (4.28) can be eliminated to yield

$$|\Lambda| > B^{-C_0 \Omega \log \Omega'}. \tag{4.29}$$

In the proof, his own formidable techniques are supplemented with ideas of Tijdeman and van der Poorten (about the new Ω') and Shorey (about C_0).

And there, Baker's extraordinary solo work on linear forms in logarithms ended. We have already mentioned a paper with Stewart in 1988, but it was not until 1993 that he did any more work on (4.29).

This was [**45**] with Wüstholz. Its main result was expressed with a more modern version of height and also took into account non-principal values of the logarithms, but we here give only the consequence for principal values. Again we are in the so-called 'rational case' when $\beta_0 = 0$ and $\beta_1, \ldots, \beta_n$ are in $\mathbf{Z}$. They improve (4.29) to

$$|\Lambda| > B^{-C_1 \Omega}, \tag{4.30}$$

with $C_1 = (16nD)^{2n+4}$. The elimination of $\log \Omega'$ is due mainly to the zero estimates of Wüstholz arising from his more general work on group varieties.

Incredibly sharp as all these results are, we should note that there is much room for conjectural improvement; for example, one believes that the product of logarithms Ω can be replaced by the sum (see Section 8).

Much more recently, Dimitrov and Habegger ⟨**4**⟩ have made a conjecture which would have valuable consequences in the theory of dynamical systems of algebraic origin. In our present framework, it runs as follows. Given $\varepsilon > 0$ and $T \geqslant 1$, there exists $C = C(\varepsilon, T)$ such that

$$|b_1 \log \alpha + b_2 \log(-1)| > \exp(-CD(DB)^\varepsilon)$$

for all algebraic α, not a root of unity, with $|\alpha| = 1$ and $\mathcal{H}_M(\alpha) \leqslant T$. Here the notation is by now familiar, except for $\mathcal{H}_M$ which is the standard modern version of height (or Mahler measure) defined using the conjugates of α. The crux is the very good dependence on D.

The theory of Linear Forms in Logarithms (or logarithmic forms for short) continues to be one of the two main tools for solving diophantine problems in number theory (or indeed several other branches of mathematics), with a key emphasis on effectivity. The other is the Subspace Theorem due to Schmidt with important extensions by Schlickewei, which is partly non-effective.

5. *Diophantine equations*

We already had several examples of these, such as (3.21), and we even mentioned the year 1621; but the so-called Pell equation

$$x^2 - 410286423278424 y^2 = 1 \tag{5.1}$$

was encountered in the solution of the 'Cattle Problem' of Archimedes (see, for example, ⟨**12**⟩).

In general, it is a question of finding a solution with the variables in $\mathbf{Z}$ (which any idiot can do for Pell above) or more ambitiously finding all solutions (there are infinitely many for Pell above, all but two of them safely out of the reach of idiots).

To warm up, let us show that there are at most finitely many positive integers r, s with

$$3^r - 2^s = 1621 \tag{5.2}$$

(the equation $3^r - 2^s = 1$, with a musicological background, was already solved by Levi ben Gershon in 1343; see, for example, $\langle \mathbf{9} \rangle$). Suppose that $H = \max\{r, s\}$ is large, say $H \geqslant H_0$. Then 3^r is relatively close to 2^s and of course

$$3^r 2^{-s} = 1 + \frac{1621}{2^s}. \tag{5.3}$$

Taking logarithms shows that $r \log 3 - s \log 2$ is very small, and we obtain without trouble (4.18) with $n = 2, \alpha_1 = 3, \alpha_2 = 2, b_1 = r, b_2 = -s, \delta = 1/2$ for an easily computable H_0. So (4.19) with $d = 4, A = 4$ gives $H < 10^{132}$, and this completes the proof.

But for most of this section we shall consider only polynomial equations. It was Thue who, in 1909, first proved that equations like $x^3 - 2y^3 = 1621$ in (3.21) have at most finitely many solutions, as a consequence of his improvement on Liouville's inequality (3.7). But we already observed the ineffectivity in that his method did not allow all the solutions to be found. We also saw Baker overcome this problem with (3.20) for $x^3 - 2y^3 = m$ in (3.19). However, we noted that his method did not work for $x^3 - 5y^3 = m$.

It was in [**18**], part I of a Duo, from 1968 that Baker was able to treat the general equation considered by Thue. Namely, if f is a homogeneous polynomial in x, y of degree $d \geqslant 3$ (so Pell is out), irreducible over $\mathbf{Q}$ and with coefficients in $\mathbf{Z}$, and $m \neq 0$ is in $\mathbf{Z}$, then all solutions of $f(x, y) = m$ satisfy

$$\max\{|x|, |y|\} < Ce^{(\log |m|)^\kappa} \tag{5.4}$$

for any $\kappa > d + 1$, where C depends only on f and κ.

He deduces easily the very first effective improvement on (3.7), namely

$$\left| \alpha - \frac{p}{q} \right| \geqslant \frac{ce^{(\log q)^{1/\kappa}}}{q^d}, \tag{5.5}$$

so he is already on his way to (4.24).

The proof uses a new estimate for linear forms in logarithms (what else?), which we already commented on. We sketch the method, because it shows that arbitrarily many logarithms may be needed. For simplicity, we take the equation

$$x^d - 2y^d = 1 \tag{5.6}$$

with $d \geqslant 3$ an arbitrary prime. Pick any α with $\alpha^d = 2$. Then factorizing the left-hand side of (5.6) shows that $x - \alpha y$ is a unit in the field $\mathbf{Q}(\alpha, \zeta)$ with $\zeta = e^{2\pi i/d}$. Dirichlet's Theorem shows the unit group to have rank $t = d(d-1)/2 - 1$, generated by say $\eta_1, \ldots, \eta_t$ modulo torsion. Thus we may write

$$x - \alpha y = \theta \eta_1^{a_1} \cdots \eta_t^{a_t}$$

with integer exponents and a root of unity θ.

We have d choices for α at our disposal; but it suffices to take just two more, say β, γ. We deduce analogously

$$x - \beta y = \varphi \eta_1^{b_1} \cdots \eta_t^{b_t}, \quad x - \gamma y = \psi \eta_1^{c_1} \cdots \eta_t^{c_t}.$$

Now 'Siegel's Identity'

$$(\beta - \gamma)(x - \alpha y) + (\gamma - \alpha)(x - \beta y) + (\alpha - \beta)(x - \gamma y) = 0$$

leads to

$$\delta \eta_1^{b_1 - c_1} \cdots \eta_t^{b_t - c_t} = 1 + \varepsilon \tag{5.7}$$

(a little as in (5.3) above) with

$$\delta = -\frac{(\gamma - \alpha)\varphi}{(\alpha - \beta)\psi}, \quad \varepsilon = \frac{(\beta - \gamma)(x - \alpha y)}{(\alpha - \beta)(x - \gamma y)}.$$

Choosing α as the real $2^{1/d}$ it may be seen that, if x, y are very large, then $x - \alpha y$ and so ε is very small. Thus taking logarithms in (5.7) gives an integer a such that

$$\log \delta + (b_1 - c_1) \log \eta_1 + \cdots + (b_t - c_t) \log \eta_t + 2\pi i a$$

is also very small (note that $2\pi i$ is a logarithm of 1). This turns out to be more than enough to contradict (4.17).

A similar treatment of $x^d - 2y^d = m$ leads to (5.5) with $\alpha = 2^{1/d}$.

In [**19**] (part II, also from 1968), Baker solves the Mordell equation $y^2 = x^3 + k$ in (2.1), as we already described in Section 2. He also makes (5.4) completely explicit with

$$C = \exp(d^{\nu^2} H^{\nu d^3}), \tag{5.8}$$

where $\nu = 32\kappa^2 d/(\kappa - d - 1)$ and H is the maximum of the absolute values of the coefficients of f.

The proofs rely on an explicit version of the linear forms estimate in [**18**]. The Mordell equation leads to a Thue equation by the classical method of reduction theory of binary cubic forms (which he also has to make explicit along the way, itself of no small interest), and then the general Thue equation needs some simple estimates for units (which were immediately improved by Siegel himself).

And thick and fast they came at last. Mathematically, next comes [**21**] in 1967 (dedicated to Mordell on his 80th birthday). There Baker treats the 'elliptic' equation

$$y^2 = ax^3 + bx^2 + cx + d \tag{5.9}$$

for a, b, c, d in $\mathbf{Z}$ such that the cubic has distinct zeros. He proves that

$$\max\{|x|, |y|\} < \exp\{(10^6 H)^{1000000}\},$$

where $H = \max\{|a|, |b|, |c|, |d|\}$. This looks a little like (2.2), but cannot be deduced from it. Instead Baker now uses explicit versions of reduction theory of binary quartic forms and then applies (5.8).

Then in [**23**] from 1969, Baker generalizes to the 'hyperelliptic' equation

$$y^2 = a_0 x^n + \cdots + a_n \tag{5.10}$$

with $n \geqslant 3$ and the polynomial on the right having at least three simple zeros, and the 'superelliptic' equation

$$y^m = a_0 x^n + \cdots + a_n, \tag{5.11}$$

with $m \geqslant 3, n \geqslant 3$ and the polynomial on the right having at least two simple zeros; where the coefficients $a_0, \ldots, a_n$ are in $\mathbf{Z}$ again with H as the maximum of their absolute values.

For (5.11), he proves the doubly exponential

$$\max\{|x|, |y|\} < \exp\exp\{(5m)^{10}(n^{10n} H)^{n^2}\}$$

and for (5.10), the triply exponential 'finite but only just'

$$\max\{|x|, |y|\} < \exp\exp\exp\{(n^{10n} H)^{n^2}\} \tag{5.12}$$

(at which the lecture audience always laughs). We emphasize though that these were the first ever bounds of their kind. The proofs rely on the calculations of [**18, 19**] in suitably generalized form.

In [**28**] with Coates from 1970, the bounds become even bigger, but the problem is somewhat different. They take now a polynomial F in x, y with coefficients in $\mathbf{Z}$ which is irreducible over $\mathbf{C}$. Then $F(x, y) = 0$ defines a curve $\mathcal{C}$ in $\mathbf{C}^2$ which has a certain geometric genus $g \geqslant 0$ in $\mathbf{Z}$. They suppose in fact that $g = 1$; this does not restrict the total degree n of F. Then, for all points (x, y) in $\mathbf{Z}^2$ on $\mathcal{C}$, they prove that

$$\max\{|x|, |y|\} < \exp\exp\exp\{(2H)^{10^{n^{10}}}\}$$

(audience goes crazy). The proof reduces to (5.9) (but now with algebraic coefficients) using the Riemann–Roch Theorem in the classical way. Of course this step must be made effective and explicit; this is done by Coates in a separate paper of great independent value. They can then follow the arguments of [**23**] just above. They remark of their result:

> 'It does not seem to extend easily, however, to curves of genus > 1, and an effective proof in the general case remains an important quest.'

And today no-one has the slightest idea how to extend this even to genus $g = 2$, even though the finiteness of integral points is known since Siegel in 1929. Of course the standard forms are (5.10) for $n = 5, 6$, and so are already covered. But there are many other models; for example in conversation with Zannier, we found the simple family $x^4 - y^3 - axy = 0$, which does not seem to be effectively solvable.

We said back in Section 2 that bounds like the above can be efficiently dealt with. This was done for the first time in 1969 by Baker and Davenport [**25**].

The numbers 1,3,8,120 have the property that the product of any two, increased by 1, is a square. Van Lint asked if this can be extended to $1, 3, 8, 120, N$ for some integer $N > 120$. Plainly then the property holds just for $1, 3, 8, N$, and he was able to show that, if such an N exists, then

$$N \geqslant 10^{1700000}, \tag{5.13}$$

which makes its existence unlikely. In [**25**], Baker and Davenport prove that there is indeed no N. See $\langle \mathbf{8} \rangle$ for another appropriate limerick.

It seems worthwhile to give some details, as they also illustrate the raw power of linear forms in logarithms. Clearly, if N exists, it has the form $x^2 - 1$ for x in $\mathbf{Z}$; and now the other conditions amount to

$$y^2 - 3x^2 = -2, \quad z^2 - 8x^2 = -7 \tag{5.14}$$

for y, z in $\mathbf{Z}$. These imply that $(yz)^2 = 24x^4 - 37x^2 + 14$, of the form (5.10); but the bound (5.12) really may be too big.

Instead we use the Pell theory behind things like (5.1) to see that the first equation of (5.14) forces

$$y + x\sqrt{3} = (1 + \sqrt{3})(2 + \sqrt{3})^r \tag{5.15}$$

for some r in $\mathbf{Z}$; and the second

$$z + x\sqrt{8} = (1 + \sqrt{8})(3 + \sqrt{8})^s \tag{5.16}$$

for some s in $\mathbf{Z}$ (there is actually a similar second possibility here which we shall ignore).

It is easy to see that (5.15) is very near $2x\sqrt{3}$ and (5.16) is very near $2x\sqrt{8}$. Eliminating x, we see that $\sqrt{8}(1 + \sqrt{3})(2 + \sqrt{3})^r$ is very near $\sqrt{3}(1 + \sqrt{8})(3 + \sqrt{8})^s$. This is also like (5.2).

Taking logarithms in the usual way, we see that $b_1 \log \alpha_1 + b_2 \log \alpha_2 + b_3 \log \alpha_3$ is very small for

$$\alpha_1 = 2 + \sqrt{3}, \quad \alpha_2 = 3 + \sqrt{8}, \quad \alpha_3 = \frac{\sqrt{3}(1 + \sqrt{8})}{\sqrt{8}(1 + \sqrt{3})}, \quad b_1 = r, \quad b_2 = -s, \quad b_3 = -1.$$

In fact, we have (4.18), and then (4.19) leads easily to

$$r < 10^{487}. \tag{5.17}$$

But as $N = x^2 - 1$ grows exponentially in r by (5.15), this far from contradicts (5.13).

More explicitly, (4.18) leads to (a stronger form of) say

$$|r\theta - s + \varphi| < 2^{-r} \tag{5.18}$$

for certain real irrational θ, φ whose exact form we can now ignore.

Here comes the new idea. Suppose that in place of (5.17) we have obtained

$$r < R \tag{5.19}$$

for some huge integer R. We pick a largish integer K, and then by Dirichlet's inequality (3.2), we find q with

$$\|q\theta\| \leqslant \frac{1}{KR}, \quad q \leqslant KR.$$

It follows easily from (5.18) and (5.19) that

$$\|q\varphi\| \leqslant \frac{q}{2^r} + \frac{1}{K}. \tag{5.20}$$

Now q was chosen specifically to make $q\theta$ very near an integer, and there seems no reason to think that $q\varphi$ is also near an integer (shades of Littlewood). For example, we have $\|q\varphi\| \geqslant 1/200$ with 'probability' 99/100. If we could have picked K such that $\|q\varphi\| \geqslant 2/K$, then by (5.20) we obtain $2^r \leqslant qK \leqslant K^2 R$. For fixed K, this gives a new bound for r that is logarithmic in R, and so it is plausible that (5.19) will be improved.

And indeed in the stronger form of (5.18) they pick K around 10^{33}, find q efficiently from the continued fraction of θ, and end up with $r < 500$ (due to the Atlas Computer Laboratory — and presumably not to the pocket calculator that Baker bought after seeing that Eva McLean had one). This does contradict (5.13) and the thing is done: there is indeed no N.

This section would be incomplete without a mention of the Catalan equation, namely

$$x^r - y^s = 1, \tag{5.21}$$

this time to be solved for x, y, r, s in **Z** all bigger than 1. Thus it is a combination of (3.21) and (5.2). It is looking for all gaps 1 when we throw in fourth powers, fifth powers, etc., with the squbes of Section 2.

In 1842, Catalan conjectured that the only solution is $3^2 - 2^3 = 1$. Progress was very slow until a result of Cassels from 1960, who also earlier in 1953 had made the weaker conjecture that there are only finitely many solutions. Despite all the work of Baker, it came as a sensation when Tijdeman in 1976 proved the conjecture of Cassels with an ingenious application of linear forms in logarithms. He used Sharpening III (4.27), together with his own sharpening (now the knives were really out) of Sharpening I. That was something like (4.27) with a slightly better dependence on $\log A_n$ and a worse dependence on the other factors of Ω (but not as good as (4.28) which appeared soon after).

At the time, it was considered as a pinnacle of 'Bakery' (van der Poorten) applied to diophantine equations. Since then Mihăilescu proved the original Catalan conjecture in 2004. Shortly afterwards, in 2006, he found another proof not using linear forms in logarithms (and no computers either). See, for example, ⟨**2**⟩.

6. *Elliptic functions*

We have seen that the Hermite–Lindemann Theorem essentially concerns the exponential function $\mathcal{E} = e^z$, and the simple differential equation $\mathcal{E}' = \mathcal{E}$ plays a role in the proofs. It was Siegel who, in 1921, first investigated what happens for a Weierstrass elliptic function $\wp = \wp(z)$. This too satisfies a differential equation, namely $\wp'^2 = 4\wp^3 - g_2\wp - g_3$, and, since we are working with algebraic numbers, it is customary to suppose that the 'invariants' g_2, g_3 are algebraic. Minor technical problems are caused by the non-linearity, but a much more serious difficulty comes from analytic growth. The Schwarz Lemma in its simplest form says that, if f is an entire function with $f(0) = 0$, then

$$|f(z_0)| \leqslant \frac{|z_0|}{R} \sup_{|z|=R} |f(z)| \tag{6.1}$$

for any z_0 with $|z_0| \leqslant R$. However, $\wp$ is far from entire. We can make it entire by multiplying by the square of the Weierstrass sigma function $\sigma(z)$. Unfortunately, $\sup_{|z|=R} |\sigma(z)|$ grows exponentially in R^2, while $\sup_{|z|=R} |e^z| = e^R$; and this restricts the iteration procedure in the proof (there is a similar problem with 'arithmetic growth').

Nevertheless, in 1937 Schneider could prove the elliptic analogues of Hermite–Lindemann and Gelfond–Schneider. Above we formulated those in terms of logarithms

$$\log \alpha = \int_1^\alpha \frac{\mathrm{d}x}{x}, \tag{6.2}$$

and the elliptic analogue is

$$\int_\infty^\alpha \frac{\mathrm{d}x}{\sqrt{4x^3 - g_2 x - g_3}} \tag{6.3}$$

for which there is no standard 'principal value' (and not even a standard notation).

Thus it was natural that Baker should start thinking about elliptic analogues of his theorem that was quoted in Section 4. If we take $\alpha = 1$ in (6.2), then there is a loop of integration which gives $2\pi i$, a 'fundamental period' of e^z. Similarly, if we take $\alpha = \infty$ in (6.3), then there are two loops which give 'fundamental periods' ω_1, ω_2 of $\wp(z)$, so that

$$\wp(z + \omega_1) = \wp(z), \qquad \wp(z + \omega_2) = \wp(z). \tag{6.4}$$

In [27] from 1970, also a major work with far-reaching consequences later on, Baker succeeded in proving that, if β_1, β_2 are algebraic, then the linear form in elliptic logarithms $\beta_1\omega_1 + \beta_2\omega_2$ is either zero or transcendental.

The analogues of the functions (4.10) are

$$\wp(z_1), \wp(z_2), \beta_1 z_1 + \beta_2 z_2, \tag{6.5}$$

and the analogues of the zeros (4.11) are

$$(z_1, z_2) = \left(\left(s + \tfrac{1}{2}\right)\omega_1, \left(s + \tfrac{1}{2}\right)\omega_2\right), \quad s = 0, 1, 2, \ldots \tag{6.6}$$

(the $1/2$ just to avoid poles).

However, the growth problems mentioned above mean that these zeros are not sufficiently numerous to do the iterative steps. As hinted in Section 4, Baker was forced to use more zeros, with

$$s + \frac{p}{q}$$

(for $p = 1, \ldots, q - 1$) in (6.6); such things had not occurred in any earlier transcendence proofs. Then not only must the range of s be increased step by step, as in (4.14); but simultaneously also the range of q. This brought new problems; for example, $\wp(\omega_1/q)$, although still an algebraic number, will have large degree growing like q^2 as q increases. Nevertheless, Baker was able to

get enough zeros for a contradiction, although the Vandermonde argument had to be ratcheted up a bit.

In fact, he proved something more general, with ω_1 as a period of a Weierstrass $\wp_1$ and ω_2 a period of a second Weierstrass $\wp_2$. Schneider had already determined when $\beta_1\omega_1 + \beta_2\omega_2$ can be zero.

In [**24**], published in 1969, Baker considered also not just periods of $\wp$, but also the corresponding quasi-periods η_1, η_2 of the associated Weierstrass zeta function $\zeta(z)$; these satisfy

$$\zeta(z + \omega_1) = \zeta(z) + \eta_1, \qquad \zeta(z + \omega_2) = \zeta(z) + \eta_2$$

similar to (6.4). He showed that, if $\beta_1, \beta_2, \gamma_1, \gamma_2$ are algebraic, then $\beta_1\omega_1 + \beta_2\omega_2 + \gamma_1\eta_1 + \gamma_2\eta_2$ is either zero or transcendental. Schneider had already in 1937 considered the numbers $\beta_1\omega_1 + \gamma_1\eta_1$ and $\beta_2\omega_2 + \gamma_2\eta_2$ separately. Now Baker had to deal with new algebraic numbers like $\zeta(\omega_1/q) - \eta_1/q$. He could treat $\wp_1, \wp_2$ also here. In the paper, he points out that this implies the transcendence of the sum of the circumferences of two ellipses with algebraic 'axes-lengths'.

Finally the note [**26**] from 1970 gives a lower bound for $|\beta_0 + \beta_1\omega_1 + \beta_2\omega_2|$ like (4.17) which has the amusing consequence $|\wp(n)| \leqslant Ce^{(\log n)^\kappa}$ for any $n \geqslant 1$ in $\mathbf{Z}$.

7. Class numbers

Now we come to more pinnacles of Bakery.

Inside the field $\mathbf{Q}$, we have the ring $\mathbf{Z}$, and inside $\mathbf{Z}$ sit the prime numbers, which it is now convenient to take as

$$\pm 2, \pm 3, \pm 5, \pm 7, \pm 11, \pm 13, \pm 17, \pm 19, \pm 23, \pm 29, \ldots, \pm 163, \ldots, \pm 1979339339, \ldots;$$

and in the present context one calls them 'irreducibles'. Euclid more-or-less proved that there is unique factorization, in the sense that:

(1) every $n \neq 0, \pm 1$ in $\mathbf{Z}$ is a product of irreducibles;
(2) if two such products $p_1 \cdots p_k, q_1 \cdots q_l$ are equal, then $k = l$;
(3) and in (2), possibly after a permutation, we have $p_i = \pm q_i$ for all i.

Inside a number field K, we have similarly a natural ring $\mathcal{O}$ together with irreducibles, and one can hope for the analogues of (1), (2), and (3), except that the '$\pm$' must be interpreted in terms of units as in the discussion around (5.6). And indeed, if this were true for all K, then 'Fermat's Last Theorem' would be relatively easy to prove. But this fails already for $\mathbf{Q}(\sqrt{-5})$.

The nearest approach to these analogues is through ideal classes. One works with ideals $\mathcal{A}$, which are just (non-zero) $\mathcal{O}$-modules in $\mathcal{O}$. Two ideals $\mathcal{A}, \mathcal{A}'$ are said to be equivalent if there are non-zero α, α' in $\mathcal{O}$ with $\alpha'\mathcal{A} = \alpha\mathcal{A}'$. The set of equivalence classes is finite, and its cardinality is the 'class number' h of K.

It is elementary that $h = 1$ is the same as the analogues of (1), (2), and (3).

When K is quadratic over $\mathbf{Q}$ and imaginary, all this was known to Gauss, who in his famous *Disquisitiones Arithmeticae* of 1801 calculated many K with small given class number (he actually worked with quadratic forms instead). He conjectured that his lists were complete, but that rigorous proofs seem to be very difficult: *Demonstrationes autem rigorosae harum observationum perdifficiles esse videntur.* Now $K = \mathbf{Q}(\sqrt{-d})$ for some unique positive square-free integer d, and we write $h(d)$ for the class number. His list of d for $h(d) = 1$ was

$$1, 2, 3, 7, 11, 19, 43, 67, 163,$$

so nine fields K. The last d here corresponds to Euler's discovery that the expression $x^2 + x + 41$ takes prime values for $x = 0, 1, 2, \ldots, 39$, and also to the fact that

$$e^{\pi\sqrt{163}} = 262537412640768743.99999999999925007259719818 5688\ldots$$

is so close to an integer.

In 1934, Heilbronn and Linfoot proved that there is at most one more value of d; and, if it exists, then

$$d > \exp(10^7). \tag{7.1}$$

But it was not until Baker's work in 1966 that this could in principle be ruled out. In his great paper [**11**], he remarked that Gelfond and Linnik in 1948 had reduced the problem using Kronecker's Limit Formula to a weaker form of his (4.17) with $n = 3$ logarithms. So that did it. On 22 December 1966, Serre ⟨**3**⟩ wrote to Tate:

> 'Et Davenport dit qu'un jeune Anglais a prouvé l'inexistence du 10-ième corps quadratique imaginaire principal. Pas mal d'un coup!'

to which Tate replied on 13 January 1967 in apparent contradiction:

> 'The guy who proved the inexistence of the 10th imaginary quadratic field with $h = 1$ is American, not English. A student of Lehmer, whose thesis proved something like $|D| \geqslant 10^{10^7}$ if $D < -163$ and $h_D = 1$. Now he has done better! He talked recently at MIT about it. I won't take time to try to give the idea now, although I think I could if I tried. Stark (that's his name) says he sees no way to generalize the thing, in order to determine all D with $h(D) = 2$, for example. Too bad.'

Then Serre on 24th June 1967:

> 'Je reviens à Baker: les anglais sont très excités par son travail, notamment par le fait qu'il donne des bornes calculables pour les problèmes d'approximation diophantienne qu'il considère.'

Actually in [**32**] from 1971, Baker improved the work of Gelfond and Linnik by showing that there is an effective C such that, if $d > C$ exists (he later calculated $C = 10^{20}$), then $21d$ is also square-free and

$$\left| h(21d) \log \left(\frac{5 + \sqrt{21}}{2} \right) - \frac{32}{21} \pi \sqrt{d} \right| < e^{-\pi \sqrt{d}/100}. \tag{7.2}$$

This involves just two logarithms (including $2\pi i$), and classical estimates for $h(21d)$ then show that the inequality can be dealt with by Gelfond's much older bound (4.16). A strange turn of events! Later, he produced a second inequality like (7.2), and then eliminated $\pi \sqrt{d}$ to obtain (4.18) for $n = 2$ for which (4.19) leads to $d < 10^{500}$. So by (7.1), 'the tenth man' d does not exist (except as a novel by Graham Greene set in Paris not Vienna).

As hinted above, events turn even stranger when it is realized that Stark ⟨**10**⟩ in 1967 had given an independent proof that the tenth man does not exist; that was based on work of Heegner from 1952 which was, at least in some circles, also accepted as a proof — and then more widely discounted until it was finally recognized as essentially correct.

Tate refers to difficulties with class number $h = 2$, which for general K is the same as saying that the analogue of (1) or (3) fails, but (2) survives (Carlitz). And indeed, doing $h(d) = 2$ was much tougher. Here Gauss's list was

$$5, 6, 10, 13, 15, 22, 35, 37, 51, 58, 91, 115, 123, 187, 235, 267, 403, 427.$$

We have already referred to Baker's paper [**31**] from 1971; here is where he showed how to settle this in principle.

Already in [**22**] from 1969, he had used two logarithms to prove that $d < 10^{500}$ if it is not 3 modulo 8. So he could assume d is 3 modulo 8. In that case Gauss' theory of genera shows that $d = pq$ for primes p congruent to 1 modulo 4 and q congruent to 3 modulo 4. If $q \leqslant d^{1/4}$, then [**22**] also suffices. So he could also assume that $q > d^{1/4}$.

He then proves an analogue of (7.2). Namely, there is effective C such that, if $d > C$ exists, then $21d$ and $21q$ are square-free and

$$\left| h(21d) \log\left(\frac{5 + \sqrt{21}}{2} \right) + h'h(21q) \log \eta - \frac{64}{21} \pi \sqrt{d} \right| < e^{-\pi\sqrt{d}/10}, \tag{7.3}$$

where h' is the class number of the real quadratic field $\mathbf{Q}(\sqrt{21p})$ and η is a fundamental unit there. An essential difference with (7.2) is that one of the logarithms, namely $\log \eta$, is not fixed. So nothing from the Quartet can be used, except possibly IV. But there the coefficients were rational; and anyway the exponent $(2n+1)^2 = 49$ in (4.19) is far too large. The troublesome η might have $\mathcal{H}(\eta)$ as large as $p^{\sqrt{p}}$, so that $\log \mathcal{H}(\eta)$ is as large as $d^{3/8}$. As H is around $d^{1/2}$, we would need an exponent $\kappa < 4/3$. Baker was able to obtain any $\kappa > 1$, also for algebraic coefficients $\beta_1, \beta_2, \beta_3$ (compare (4.23) above), thereby settling $h(d) = 2$. This was another motivation for the Sharpening Trio. It may be noted that Stark $\langle\mathbf{11}\rangle$ independently settled the problem, and their papers could not have appeared more simultaneously, with Baker's occupying pages 139 to 152 and Stark's 153 to 173 in volume 94 of the *Annals*.

To this day, no one has been able to treat $h(d) = 3$ by similar methods. However, the subsequent work of Goldfeld, Gross and Zagier (see, for example, $\langle\mathbf{7}\rangle$) enables in principle any class number to be treated effectively.

We already referred to the note [**22**]. This was used in a paper [**33**] with Schinzel (in memory of Davenport), also from 1971, where they make a contribution to the (still unsolved) problem of showing that Euler's list

$$1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 13, 15, 16, 18, 21, 22, 24, \dots, 408, 462, 520, 760, 840, 1320, 1365, 1848$$

of 65 'numeri idonei' is complete. These were defined in terms of binary quadratic forms but they are also the positive integers not of the form $xy + yz + zx$ for distinct positive integers x, y, z.

8. *Abcology*

This refers to the *abc* Conjecture, formulated in 1985 by Oesterlé and the writer. Namely, for any $\kappa > 1$, there is $C = C(\kappa)$ such that, for any non-zero a, b, c in $\mathbf{Z}$ with no common factor and $a + b + c = 0$, we have

$$\max\{|a|, |b|, |c|\} \leqslant CS^\kappa \tag{8.1}$$

for the 'square-free kernel'

$$S = \prod_{p|abc} p \geqslant 2,$$

the product being taken over all primes $p > 0$.

In [**46**] from 1998, Baker proposes a version

$$\max\{|a|, |b|, |c|\} \leqslant C_0 S \frac{(\log S)^s}{s!}, \tag{8.2}$$

where $s = \sum_{p|abc} 1$ and now C_0 is absolute. It is easily seen that $s \leqslant L = [C_0' \log S / \log \log S]$ (if $S \geqslant 3$) for $C_0' > 0$ absolute, and then Stirling's formula gives (for $L \leqslant \log S$)

$$\frac{(\log S)^s}{s!} \leqslant \frac{(\log S)^L}{L!} \leqslant \left(\frac{e \log S}{L} \right)^L;$$

so (8.2) strengthens (8.1) — but not so much as to contradict known lower bounds.

He suggested an even stronger version and showed that it is equivalent to a certain lower bound for

$$\min\{1, |\Lambda|\} \prod_p \min\{1, p|\Lambda|_p\}$$

for the linear form in logarithms $\Lambda = b_1 \log a_1 + \cdots + b_n \log a_n$ with $b_1, \ldots, b_n, a_1 \geq 1, \ldots, a_n \geq 1$ in $\mathbf{Z}$. Here $|\Lambda|_p$ is defined in a natural p-adic way. He remarks that this lower bound suggests that the expression $\Omega = \log A_1 \cdots \log A_n$ in (4.30) could be replaced by $\Gamma = \log A_1 + \cdots + \log A_n$. He even speculates that $|\Lambda| \geq e^{-C_1 \Gamma} B^{-C_1 \omega}$, where

$$\omega = \sum_{p|a_1\cdots a_n} 1 \leq C_1' \frac{\Gamma}{\log \Gamma}$$

for C_1, C_1' absolute.

But despite the extraordinarily precise works of Yu Kunrui on p-adic linear forms in logarithms, it seems that this approach to *abc* remains for future generations to explore.

Later on in [**47**] from 2004, Baker stuck his neck out by conjecturing (on extensive computer evidence) that (8.2) holds for $C_0 = 6/5$. See also his paper [**48**] from 2007 for related remarks.

9. *Miscellaneous*

Here we comment on five papers that do not so neatly fall into the above categories.

In [**8**] from 1965, Baker considers algebraic functions $f = f(x)$ satisfying $P(x, f) = 0$, where $P(x, y)$ has the special form

$$P(x, y) = 1 + \left(1 + \frac{p}{q}x\right)y + xQ(x, y)$$

for $Q(x, y)$ in $\mathbf{Z}[x, y]$ and coprime p, q in $\mathbf{Z}$ with $q \geq 2$. As $P(0, -1) = 0$ but $(\partial P/\partial y)(0, -1) \neq 0$, imposing $f(0) = -1$ gives a unique power series $f(x) = -1 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots$. He shows that $q^n a_n$ is in $\mathbf{Z}$ and also prime to q. This is in accordance with the well-known Eisenstein Theorem about denominators, but the point is that there is no $q_? \geq 1$ in $\mathbf{Z}$ with $q_?^{n-1} a_n$ in $\mathbf{Z}$ for all $n \geq 1$, so the theorem is best possible in this respect, no matter how we choose Q. The simplest choice $Q = 0$ gives of course

$$f(x) = -1 + \frac{p}{q}x - \frac{p^2}{q^2}x^2 + \frac{p^3}{q^3}x^3 + \cdots.$$

In [**9**] from 1966, Baker extends the Padé systems used in [**5, 6**] (see (3.22) above), [**2**] (see (3.24) above) and (later) [**17**]. Namely, let $\omega_0, \ldots, \omega_k$ be in $\mathbf{C}$ distinct modulo $\mathbf{Z}$, and then let ρ_{rs} $(r = 0, \ldots, k; s = 0, \ldots, l)$ be in $\mathbf{Z}$ with $1 \leq \rho_{rl} \leq \cdots \leq \rho_{r0}$ $(r = 0, \ldots, k)$. Then, if $P_{rs}(x)$ is in $\mathbf{C}[x]$ with degree at most $\rho_{rs} - 1$, the function

$$\sum_{r=0}^{k} \sum_{s=0}^{l} P_{rs}(x)(1-x)^{\omega_r} (\log(1-x))^s$$

cannot have a zero of order at least $\sum_{r=0}^{k} \sum_{s=0}^{l} \rho_{rs}$ at $x = 0$ unless all $P_{rs}(x) = 0$. He mentions (without proof) that such techniques can be used to prove the transcendence of e^π (see Section 4 above).

In [**14**] from 1967, Baker extends a famous result of Pólya to functions of several variables. As the former played an important part in the history of transcendence, and the latter a key role in linear forms in logarithms, we should say a few words.

Pólya shows that, if $f(z)$ is an entire function such that $f(0), f(1), f(2), \ldots$ are all in $\mathbf{Z}$, then a relatively slow growth rate

$$|f(z)| \leqslant Ce^{\theta|z|}$$

for some $\theta < \log 2$ and some C forces it to be a polynomial. This led about 20 years later to the Gelfond–Schneider Theorem (compare (4.4) and (4.5) above). The example $f(z) = 2^z$ shows that Polya's result would be false for $\theta = \log 2$.

Baker proves a similarly sharp result for $f(z_1, \ldots, z_n)$ with values in $\mathbf{Z}$ at the points $z_1, \ldots, z_n = 0, 1, 2, \ldots$. Here it suffices that

$$|f(z_1, \ldots, z_n)| \leqslant Ce^{\theta(|z_1| + \cdots + |z_n|)},$$

again for some $\theta < \log 2$, to force a polynomial. He refers to Schneider's 1941 paper that was mentioned in Section 4. He also gives generalizations that involve the partial derivatives $(\partial/\partial z_1)^{t_1} \cdots (\partial/\partial z_n)^{t_n} f(z_1, \ldots, z_n)$ as in (4.15), so this paper seems inextricably linked with Linear Forms I (which appeared just one year earlier).

The paper [**38**] from 1973, with Birch and Wirsing, uses Linear Forms II and III to prove several results about a problem of Chowla, for example, that the values of $L(1, \chi) = \sum_{n=1}^{\infty} \chi(n)/n$, as χ runs over all non-principal characters modulo q, are linearly independent over $\mathbf{Q}$ provided that q is prime to its image $\varphi(q)$ under the Euler function.

The paper [**41**] from 1975 with Coates is diophantine approximation of a different sort. Let p, q be coprime in $\mathbf{Z}$ with $p > q \geqslant 2$. It is trivial that $\|(p/q)^k\| \geqslant q^{-k}$ for all $k \geqslant 1$ in $\mathbf{Z}$. Mahler showed that, for any $\varepsilon > 0$, there is $c = c(p, q, \varepsilon) > 0$ such that $\|(p/q)^k\| \geqslant cq^{-\varepsilon k}$. For example, there are at most finitely many k with $\|(3/2)^k\| < (3^k + 2^k)/(4^k - 2^k)$; this inequality is significant as (for $k \geqslant 5$) it is implied by the number $g(k)$ in Waring's Problem *not* being given by

$$g(k) = 2^k + \left[\left(\frac{3}{2}\right)^k\right] - 2. \tag{9.1}$$

However, the c above is 'Thue-ineffective', so these exceptional exponents k cannot be determined (it is presumed that they do not exist).

Baker and Coates show that there are effective $\eta = \eta(p, q) < 1$ and $c = c(p, q) > 0$ such that

$$\left\|\left(\frac{p}{q}\right)^k\right\| \geqslant \frac{c}{q^{\eta k}}$$

for all k. They admit, however, that obtaining

$$\eta(3, 2) < \frac{\log(4/3)}{\log 2} = .415037\ldots,$$

as would be required to find all exceptions to (9.1), may need fundamentally new ideas.

The proofs use a refinement of the methods of Linear Forms I, Sharpenings I, II and Coates's earlier work on p-adic linear forms in logarithms (but not with p as above).

10. *Books*

Before Baker, there were three famous books on transcendental numbers written by leaders in the field: Siegel's *Transcendental numbers* (1949), Schneider's *Einführung in die transzendenten Zahlen* (1957), and Gelfond's *Transcendental and algebraic numbers* (1960). The dates reflect a certain irregular growth in the subject corresponding to various breakthroughs. Also Lang's *Introduction to transcendental numbers* (1966) was influential.

Baker's own *Transcendental number theory* [**40**] from 1975 is a worthy successor to these (although the writer well remembers the author's wail of agony over a particularly visible printing error on page 85). It formed the substance of an essay for the Adams Prize (1972) of Cambridge University.

After an introductory chapter containing probably some of the world's shortest proofs (one page for the transcendence of e, two pages for Lindemann–Weierstrass), he launches straight into linear forms in logarithms, and the subsequent four chapters give an excellent account of this topic and its various applications. They are followed by a chapter on elliptic functions including his own work. The next chapter describes Schmidt's subspace extension of Roth's Theorem. There follows a chapter on Mahler's Classification including Schmidt's beautiful proof that 'T-numbers do exist'. The next two chapters contain some of his own work in [**10**] and [**7**], and the book finishes with two chapters about the topic of algebraic independence. One is an account of Shidlovsky's fundamental work on E-functions and the other contains for example the proof, due independently to Brownawell and Waldschmidt, that at least one of e^e and e^{e^2} is transcendental. Nearly all the chapters contain the first account in book form of major results.

The book was eagerly devoured as soon as it appeared. The concision of the first chapter sets the tone for the rest, which makes it not so easy for beginners. It ran into a second edition in 1990, and at the time of writing a reprint is planned.

Some of the material, now updated, found its way into the first three chapters of his book *Logarithmic forms and diophantine geometry* [**49**] with Wüstholz in 2007. The other chapters are concerned more with developments for general commutative algebraic groups G. The simplest examples of these are the additive group $\mathbf{G}_{\mathrm{a}}$ and the multiplicative group $\mathbf{G}_{\mathrm{m}}$. The Quartet results can be naturally formulated in terms of $\mathbf{G}_{\mathrm{m}}^n$ and $\mathbf{G}_{\mathrm{a}} \times \mathbf{G}_{\mathrm{m}}^n$. The next simplest example is an elliptic curve E, and Baker's result in [**27**] involves in a similar way $\mathbf{G}_{\mathrm{a}} \times E^2$ or even $\mathbf{G}_{\mathrm{a}} \times E_1 \times E_2$. And given E, there is a unique $G = G(E)$ sitting inside an exact sequence

$$0 \longrightarrow \mathbf{G}_{\mathrm{a}} \longrightarrow G \longrightarrow E \longrightarrow 0$$

which is, however, not isomorphic to $\mathbf{G}_{\mathrm{a}} \times E$; similarly Baker's paper [**24**] involves (implicitly) $G(E_1) \times G(E_2)$ and even a quotient by a line inside.

The fourth chapter supplies the basic theory of such general G.

As already mentioned, Baker had to work a bit in [**27**] to obtain enough zeros for a final contradiction. This indeed was the main obstacle to generalizing even to three elliptic curves. It was overcome by Wüstholz, and the fifth chapter is devoted to the zero estimate or 'multiplicity estimate' that did the trick. The resulting extension of Baker's works on linear independence to general G was Wüstholz's 'Analytic Subgroup Theorem', and this is the topic of the sixth chapter.

The seventh chapter is concerned with linear independence measures, both improved versions for linear forms in logarithms (as in [**45**] above) and versions for general G. Some consequences of the latter more in the realm of diophantine geometry (due to Wüstholz and the writer) are also discussed, such as the existence of 'small' isogenies between abelian varieties and polarizations of abelian varieties, as well as effective versions of Serre's Open Image Theorem for elliptic curves, the Tate Conjecture for abelian varieties and some of Faltings's Finiteness Theorems.

We pause to make clearer the connection between isogenies and (4.20), (4.21). Suppose, for example, that E and E' are elliptic curves with periods ω_1, ω_2 and ω_1', ω_2', respectively, as in (6.4). An isogeny between E, E' leads to relations

$$\beta\omega_1' = b_{11}\omega_1 + b_{12}\omega_2, \quad \beta\omega_2' = b_{21}\omega_1 + b_{22}\omega_2, \tag{10.1}$$

as in (4.20). The correct analogue of (4.21) then leads to an upper bound on the degree of a connecting isogeny.

The book closes with concise accounts of other diophantine topics such as Schmidt's Subspace Theorem and the André–Oort Conjecture; the latter takes the reader very near to a great deal of current work.

Baker had already published his *A concise introduction to the theory of numbers* [**43**] in 1984. This attractive volume is much more elementary and suitable for a first course in Number Theory in general. He then developed [**43**] into *A comprehensive course in number theory* [**50**] in 2012. This is over twice as long and goes up to the graduate level.

Two conferences that Baker organized led to published proceedings: the first (Cambridge 1976) edited by Baker and the writer (and between the two of us we refereed practically all the 16 papers there), mentioned in [**42**]; and the second (Durham 1986) edited by Baker alone (where a less intensive programme than usual produced more mathematical advances than usual), mentioned in [**44**].

11. *Fields medal*

This medal was for many years the supreme prize in mathematics, and still manages very well to hold its own among various lifetime awards. It is given only to people aged at most 40, two or three or four of them every four years at the International Congress. At the time of writing, there have been 60 winners since it began in 1936.

Baker won the prize at Nice in 1970, just 13 days after his 31st birthday, on the basis of his outstanding work on linear forms in logarithms and its consequences. See [**30**] for the Proceedings Volume containing his address. The citation was:

> *'Generalized the Gelfond–Schneider theorem* (*the solution to Hilbert's seventh problem*). *From this work he generated transcendental numbers not previously identified.'*

(Indeed at the time of writing that is the only accomplishment listed in his somewhat minimal Wikipedia entry.) Turán in his talk at Nice spoke of this:

> *'The analytic prowess displayed by Baker could hardly receive a higher testimonial.'*

However, any reader of this obituary will see that Baker went far beyond that. And Turán himself recognized the fact in his closing words:

> *'To conclude, I remark that his work exemplifies two things very convincingly. Firstly, that beside the worthy tendency to start a theory in order to solve a problem it pays also to attack specific difficult problems directly. Particularly is this the case with such problems where rather singular circumstances do not make it probable that a solution would fall out as an easy consequence of a general theory. Secondly, it shows that a direct solution of a deep problem develops itself quite naturally into a healthy theory and gets into early and fruitful contact with other significant problems of mathematics. So, let the two different ways of doing mathematics live in peaceful coexistence for the benefit of our science.'*

Baker was presented to Président Pompidou at the Elysée Palace and later wrote: '...*and was quite impressed by his feat of memory when making a welcoming scientific address.*'

He was the third British winner of the Fields Medal (after Roth in 1958 and Atiyah in 1966).

He received many other honours, including the prestigious Adams Prize of Cambridge University, the election to the Royal Society and the Academia Europaea; and he was made an honorary fellow of University College, London, a foreign fellow of the Indian Academy of

Science, a foreign fellow of the National Academy of Sciences, India, an honorary member of the Hungarian Academy of Sciences, and a fellow of the American Mathematical Society.

Alan Baker single-handedly transformed several areas of number theory. He achieved a major breakthrough in transcendence, and applied it to obtain a new and important large class of transcendental numbers (opening the way to the subsequent discovery of several other such classes); developed quantitative versions and applied them to the effective solutions of many classical diophantine equations as well as the first effective improvement on Liouville's 1844 result on diophantine approximation and the resolution of the celebrated Gauss Conjectures of 1801 on class numbers, not only $h = 1$ but also $h = 2$, of imaginary quadratic fields; and started the study of extensions to elliptic curves (opening the way to later generalizations to abelian varieties and commutative group varieties and in turn their applications to old and new problems in diophantine geometry).

Despite this, his cousin describes him as 'extremely modest'. This is confirmed by another extract from the funeral tribute of Eva McLean:

> 'His most striking characteristic, however, was a genuine modesty. He really believed that his brilliant achievement was merely down to hard work and determination. At the same time he spoke in great awe of those mathematicians whom he regarded as the truly greats.'

References

⟨**1**⟩ R. Balasubramanian (ed.), Special commemorative volume in honour of Alan Baker, *Hardy–Ramanujan J.* 42 (2019).

⟨**2**⟩ Yu. Bilu, Y. Bugeaud and M. Mignotte, *The problem of Catalan* (Springer, Berlin, 2014).

⟨**3**⟩ P. Colmez and J-P. Serre (eds), *Correspondance Serre–Tate I*, Documents Mathématiques 13 (Société Mathématique de France, Paris, 2015).

⟨**4**⟩ V. Dimitrov and P. Habegger, 'Galois orbits of torsion points near atoral sets', Preprint, 2019, arxiv.org/abs/1909.06051.

⟨**5**⟩ T. Estermann, *Introduction to modern prime number theory*, Cambridge Tracts in Mathematics 41 (Cambridge University Press, Cambridge, 1961).

⟨**6**⟩ N. I. Feldman, 'Improved estimate for a linear form of the logarithms of algebraic numbers', *Math. USSR–Sb.* 6 (1968) 393–406.

⟨**7**⟩ D. M. Goldfeld, 'Gauss' class number problem for imaginary quadratic fields', *Bull. Amer. Math. Soc.* 13 (1985) 23–37.

⟨**8**⟩ D. Masser, 'Alan Baker', Preprint, 2020, arxiv.org/abs/2010.10256.

⟨**9**⟩ S. Simonson, 'The mathematics of Levi ben Gershon, the Ralbag', 2000, https://u.cs.biu.ac.il/~tsaban/Pdf/MathofLevi.pdf.

⟨**10**⟩ H. M. Stark, 'A complete determination of the complex quadratic fields of class-number one', *Michigan Math. J.* 14 (1967) 1–27.

⟨**11**⟩ H. M. Stark, 'A transcendence theorem for class-number problems', *Ann. of Math.* (2) 94 (1971) 153–173.

⟨**12**⟩ I. Vardi, 'Archimedes' cattle problem', *Amer. Math. Monthly* 105 (1998) 305–319.

⟨**13**⟩ G. WÜSTHOLZ (ed.), *A panorama of number theory or the view from Baker's garden* (Cambridge University Press, Cambridge, 2002).

⟨**14**⟩ G. WÜSTHOLZ, 'Obituary of Alan Baker FRS', *Acta Arithmetica* 189 (2019) 309–345.

⟨**15**⟩ 'Obituary: Alan Baker', *The Times*, 1 March 2018.

⟨**16**⟩ 'Sums of three cubes', https://en.wikipedia.org/wiki/Sums_of_three_cubes.

Selected publications of Alan Baker

1. 'Continued fractions of transcendental numbers', *Mathematika* 9 (1962) 1–8.

2. 'Approximations to the logarithms of certain rational numbers', *Acta Arith.* 10 (1964) 315–323.

3. 'On an analogue of Littlewood's Diophantine approximation problem', *Michigan Math. J.* 11 (1964) 247–250.

4. 'On Mahler's classification of transcendental numbers', *Acta Math.* 111 (1964) 97–120.

5. 'Rational approximations to certain algebraic numbers', *Proc. Lond. Math. Soc.* (3) 14 (1964) 385–398.

6. 'Rational approximations to $\sqrt[3]{2}$ and other algebraic numbers', *Quarterly. J. Math.* 15 (1964) 375–383.

7. 'On some Diophantine inequalities involving the exponential function', *Canadian J. Math.* 17 (1965) 616–626.

8. 'Power series representing algebraic functions', *J. Lond. Math. Soc.* 40 (1965) 108–110.

9. 'A note on the Padé table', *Indag. Math.* 28 (1966) 596–601.

10. 'On a theorem of Sprindžuk', *Proc. Roy. Soc. London* A 292 (1966) 92–104.

11. 'Linear forms in the logarithms of algebraic numbers', *Mathematika* 13 (1966) 204–216.

12. 'Linear forms in the logarithms of algebraic numbers II', *Mathematika* 14 (1967) 102–107.

13. 'Linear forms in the logarithms of algebraic numbers III', *Mathematika* 14 (1967) 220–228.

14. 'A note on integral integer-valued functions of several variables', *Proc. Cambridge Philos. Soc.* 63 (1967) 715–720.

15. 'On Mahler's classification of transcendental numbers II: Simultaneous Diophantine approximation', *Acta Arith.* 12 (1967) 281–288.

16. 'On some diophantine inequalities involving primes', *J. reine angew. Math.* 228 (1967) 166–181.

17. 'Simultaneous rational approximations to certain algebraic numbers', *Proc. Cambridge Phil. Soc.* 63 (1967) 693–702.

18. 'Contributions to the theory of Diophantine equations. I. On the representation of integers by binary forms', *Phil. Trans. Roy. Soc. London* A 263 (1968) 173–191.

19. 'Contributions to the theory of Diophantine equations. II. The Diophantine equation $y^2 = x^3 + k$', *Phil. Trans. Royal Soc. London* A 263 (1968) 193–208.

20. 'Linear forms in the logarithms of algebraic numbers IV', *Mathematika* 15 (1968) 204–216.

21. 'The diophantine equation $y^2 = ax^3 + bx^2 + cx + d$', *J. Lond. Math. Soc.* 43 (1968) 1–9.

22. 'A remark on the class number of quadratic fields', *Bull. Lond. Math. Soc.* 1 (1969) 98–102.

23. 'Bounds for the solutions of the hyperelliptic equation', *Proc. Cambridge Phil. Soc.* 65 (1969) 439–444.

24. 'On the quasi-periods of the Weierstrass ζ-function', *Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II* (1969) 145–157.

25. (with H. DAVENPORT) 'The equations $3x^2 - 2 = y^2$ and $8x^2 - 7 = z^2$ ', *Quarterly J. Math.* 20 (1969) 129–137.

26. 'An estimate for the $\wp$-function at an algebraic point', *Amer. J. Math.* 92 (1970) 619–622.

27. 'On the periods of the Weierstrass $\wp$-function', *Symposia Mathematica*, INDAM, Rome 1968/69, vol. IV (Academic Press, London, 1970) 155–174.

28. (with J. COATES) 'Integer points on curves of genus 1', *Proc. Cambridge Phil. Soc.* 67 (1970) 595–602.

29. (with W. M. SCHMIDT) 'Diophantine approximation and Hausdorff dimension', *Proc. Lond. Math. Soc.* (3) 21 (1970) 1–11.

30. 'Effective methods in the theory of numbers', *Actes du Congrès International des Mathématiciens*, Nice, 1970, vol. I (Gauthier-Villars, Paris, 1971) 19–26.

31. 'Imaginary quadratic fields with class number 2', *Ann. of Math.* (2) 94 (1971) 139–152.

32. 'On the class number of imaginary quadratic fields', *Bull. Amer. Math. Soc.* 77 (1971) 678–684.

33. (with A. SCHINZEL) 'On the least integers represented by the genera of binary quadratic forms', *Acta Arith.* 18 (1971) 137–144.

34. (with H. M. STARK) 'On a fundamental inequality in number theory', *Ann. of Math.* (2) 94 (1971) 190–199.

35. 'A sharpening of the bounds for linear forms in logarithms', *Acta Arith.* 21 (1972) 117–129.

36. 'A sharpening of the bounds for linear forms in logarithms II', *Acta Arith.* 24 (1973) 33–36.

37. 'A central theorem in transcendence theory', *Diophantine approximation and its applications* (ed. C. F. Osgood; Academic Press, New York, 1973) 1–23.

38. (with B. J. BIRCH and E. A. WIRSING) 'On a problem of Chowla', *J. Number Theory* 5 (1973) 224–236.

39. 'A sharpening of the bounds for linear forms in logarithms III', *Acta Arith.* 27 (1975) 247–252.

40. *Transcendental number theory*, Cambridge Mathematical Library (Cambridge University Press, Cambridge 1975); 2nd edn (1990).

41. (with J. COATES) 'Fractional parts of powers of rationals', *Math. Proc. Cambridge Phil. Soc.* 77 (1975) 269–279.

42. 'The theory of linear forms in logarithms', *Transcendence theory: advances and applications*, Cambridge, 1976 (eds A. Baker and D. Masser; Academic Press, London, 1977) 1–27.

43. *A concise introduction to the theory of numbers* (Cambridge University Press, Cambridge, 1984).

44. (with C. L. STEWART) 'On effective approximations to cubic irrationals', *New advances in transcendence theory*, Durham, 1986 (ed. A. Baker; Cambridge University Press, Cambridge, 1988) 1–24.

45. (with G. WÜSTHOLZ) 'Logarithmic forms and group varieties', *J. reine angew. Math.* 442 (1993) 19–62.

46. 'Logarithmic forms and the *abc*-conjecture', *Number theory* (eds K. Györy, A. Pethö and V. T. Sos; de Gruyter, Berlin, 1998) 37–44.

47. 'Experiments on the *abc*-conjecture', *Publ. Math. Debrecen* 65 (2004) 253–260.

48. 'On an arithmetical function associated with the *abc*-conjecture', *Diophantine geometry*, CRM Series 4 (ed. U. Zannier; Ed. Norm., Pisa, 2007) 25–33.

49. (with G. WÜSTHOLZ) *Logarithmic forms and Diophantine geometry*, New Mathematical Monographs 9 (Cambridge University Press, Cambridge, 2007).

50. *A comprehensive course in number theory* (Cambridge University Press, Cambridge, 2012).

David Masser
Department of Mathematics and Computer Science
University of Basel
Spiegelgasse 1
Basel 4051
Switzerland

david.masser@unibas.ch