

## TRANSCENDENCE THEORY AND ITS APPLICATIONS

A. BAKER

Dedicated to Professor K. Mahler on his 75th birthday

(Received 19 December 1977)

Communicated by J. H. Coates

### Abstract

This is the text of an invited address given at the Summer Research Institute of the Australian National University in January 1978. It provides a survey of the subject matter of the title.

*Subject classification (Amer. Math. Soc. (MOS) 1970):* primary 10–02, 10 F 35; secondary 10 B 15.

### 1. Introduction

I should like to speak today about some of the latest developments that have taken place in connexion with the theory of linear forms in the logarithms of algebraic numbers. This is, of course, just one aspect of transcendence theory, but it has been an especially active area of development throughout the past decade, and it has found many applications relating to a wide variety of Diophantine problems. I shall discuss some of these applications later in my talk, but let me begin with some basic definitions.

### 2. Definitions

An algebraic number is a zero of a polynomial with integer coefficients; thus if  $\alpha$  is algebraic we have

$$a_0 \alpha^n + a_1 \alpha^{n-1} + \dots + a_n = 0$$

for some integers  $a_0, \dots, a_n$ . We can suppose that the polynomial on the left is irreducible, and that  $a_0, \dots, a_n$  are relatively prime; then  $n$  is called the degree of

$\alpha$  and  $\max(|a_0|, \dots, |a_n|)$  is called the height of  $\alpha$ . In particular, the height of a positive rational  $p/q$ , in lowest terms, is simply the maximum of  $p$  and  $q$ . A real or complex number that is not algebraic is said to be transcendental.†

We shall be concerned with linear forms

$$\Lambda = \beta_0 + \beta_1 \log \alpha_1 + \dots + \beta_n \log \alpha_n,$$

where the  $\alpha$ 's and  $\beta$ 's denote algebraic numbers. We shall assume that the  $\alpha$ 's are not 0 or 1, that the  $\beta$ 's are not all 0, and that the logarithms have their principal values; the latter assumption involves no real loss of generality.

### 3. Linear forms in logarithms

The first result on the non-vanishing of  $\Lambda$  goes back to the famous work of Hermite‡ of 1873 in which he proved that  $e$ , the natural base for logarithms, is transcendental; this implies that  $\Lambda \neq 0$  when  $n = 1$ ,  $\beta_0 = 1$ ,  $\beta_1 = -1$ . Hermite's work rested on the construction of simultaneous approximations to the exponential function, and it can be regarded as the main source of modern transcendence theory. The work was generalized by Lindemann in 1882; in his classic memoir, he proved that  $\Lambda \neq 0$  when  $n = 1$  for all  $\beta_0, \beta_1$ , not both 0, and this furnished, in particular, the transcendence of  $\pi = -i \log(-1)$ . The next major step was taken by Gelfond in 1929; he showed that  $\Lambda \neq 0$  when  $n = 2$ ,  $\beta_0 = 0$  and  $\beta_1/\beta_2$  is an imaginary quadratic irrational, whence, in particular,  $e^\pi = (-1)^{-i}$  is transcendental. Gelfond's argument has its origins in earlier studies on integral integer-valued functions, and it depends on an analysis of an extrapolation formula for the exponential function similar to that occurring in some well-known papers of Pólya and Hardy. In fact Hardy sharpened an earlier result of Pólya, or, as Hardy phrased it, put a result of Pólya in more pregnant form, by showing that among all transcendental integral functions which assume integer values for all non-negative integer values of the variable, that of least increase is the function  $2^z$ . Analogous results relating to integers in the Gaussian field were proved by Fukasawa in 1926, and it was Gelfond's subsequent studies in this connexion that gave rise to his striking discovery. Gelfond's result was extended to real quadratic irrationals  $\beta_1/\beta_2$  by Kuzmin in 1930, and this implied, in particular, that  $2^{\sqrt{2}}$  is transcendental. The result was further extended by Gelfond and Schneider independently in 1934;

† For an introduction to the properties of transcendental numbers, see, for instance, the author's monograph *Transcendental Number Theory* (Cambridge University Press, 1975).

‡ Unless otherwise stated, references throughout this article can be found in the lists at the ends of the chapters of the book *Transcendence Theory: Advances and Applications* (Academic Press, London and New York, 1977). Proceedings of a conference held in Cambridge in 1976, edited by A. Baker and D. W. Masser.

they succeeded in covering all  $\beta_1, \beta_2$  with  $\beta_1/\beta_2$  irrational and thereby solved the famous seventh problem of Hilbert. One of the main features of the work is the construction, by means of Dirichlet's box principle, of an auxiliary function that vanishes to a high order at certain extrapolation points; some antecedents of the method can be found in the writings of Siegel and Mahler, and it has proved to be remarkably powerful.

In 1966, I generalized the Gelfond–Schneider Theorem to arbitrarily many logarithms; I showed in fact that  $\Lambda \neq 0$  when  $\beta_0 = 0$  and either  $\beta_1, \dots, \beta_n$  or  $\log \alpha_1, \dots, \log \alpha_n$  are linearly independent over the rationals, and, in 1967, I showed further that  $\Lambda \neq 0$  when  $\beta_0 \neq 0$ . The work depended on the construction of an auxiliary function of several complex variables in place of a function of a single variable as employed by Gelfond, and it also involved a new extrapolation technique. Here the range of extrapolation was extended and the order of the derivatives reduced, whereas, in previous work, the range was essentially fixed while the differential order increased.

Now the results described so far may be regarded as essentially qualitative in character, but the main applications of the theory rest on quantitative extensions of the work giving positive lower bounds for  $|\Lambda|$  in terms of the degrees and heights of the  $\alpha$ 's and  $\beta$ 's. The first results in this context were given by Morduchai-Boltovskoj in 1923 in the case  $n = 1$ , and by Gelfond in 1935 in the case  $n = 2$ ,  $\beta_0 = 0$ . In 1966, I obtained a positive lower bound for  $|\Lambda|$  of the desired kind in the general case, and many refinements have followed subsequently. It will suffice here if I simply record the latest results in this field; they include most of the earlier results as special cases, and their proofs can be found in the proceedings of the conference cited above. We shall suppose that  $\alpha_j$  and  $\beta_j$  have heights at most  $A_j$  ( $\geq 4$ ) and  $B$  ( $\geq 4$ ) respectively and that the field generated by the  $\alpha$ 's and  $\beta$ 's over the rationals has degree at most  $d$ . We put  $\Omega = \log A_1 \dots \log A_n$ , and  $\Omega' = \Omega/\log A_n$ . We have then

**THEOREM.** *If  $\Lambda \neq 0$  then  $|\Lambda| > (B\Omega)^{-C\Omega \log \Omega'}$ , where  $C = (16nd)^{200n}$ . Further, if  $\beta_0 = 0$  and  $\beta_1, \dots, \beta_n$  are rational integers, then in fact  $|\Lambda| > B^{-C\Omega \log \Omega'}$ .*

When the latter hypotheses concerning the  $\beta$ 's are satisfied, we say that we are dealing with the rational case; and this commonly arises in applications. As we shall see later, a crucial feature of the result is that it is best possible with respect to  $A_n$ ; indeed it is also best possible with respect to  $B$ , and in fact likewise with respect to the parameters  $A_1, \dots, A_{n-1}$ , except for the second order term  $\log \Omega'$ . Moreover the dependence of  $C$  on  $n$  is the most precise established to date; here, in particular, an idea of Shorey is utilized. On the other hand, no special significance attaches to the constant 200 occurring in the expression for  $C$ , and it could certainly

be substantially reduced if, for instance, one imposed minor restrictions on  $n$  or  $d$ ; this is especially relevant in connexion with computational work, and some useful improvements have already been obtained by van der Poorten, Loxton and others.†

There is also an extensive  $p$ -adic theory of linear forms in logarithms. The subject was initiated by Mahler in the 1930's when he obtained natural  $p$ -adic analogues of both the Hermite–Lindemann and the Gelfond–Schneider Theorems; indeed, in the course of these studies, Mahler laid the foundations of the  $p$ -adic theory of analytic functions which has been fundamental to all later works in the field. In fact most of the results referred to above have now been generalized  $p$ -adically; in particular, van der Poorten has recently established an estimate for the  $p$ -adic valuation of  $\Lambda$  of the same degree of precision as that quoted earlier in the Archimedean case, though subject to a mild restriction on one of the  $\beta$ 's. See his paper in the proceedings cited above.

#### 4. Diophantine equations

The theory of linear forms in logarithms has now had a considerable influence on the subject of Diophantine equations. First, some ten years ago, it was applied to give bounds for all the solutions of the Thue equation  $f(x, y) = m$ , where  $f$  denotes a binary form, the Mordell equation  $y^2 = x^3 + k$ , and more generally the hyperelliptic equation  $y^m = f(x)$ , where  $f$  denotes a polynomial with at least three simple zeros and  $m$  is an integer  $\geq 2$ . A little later it was further applied to determine effectively all the integer points on an arbitrary curve of genus 1, and, moreover, all elliptic curves with a given conductor. The latter theorem is due to Coates; it involved  $p$ -adic analysis and it can be viewed as an effective version of a well-known result of Mahler in the context of the Thue–Siegel theory.

Now all the work just described relates to equations in only two variables  $x$  and  $y$ , which signify either integers or rationals with denominators composed solely of powers of a fixed set of primes; and, for several years, it appeared that this binary character would be an essential feature of all the results obtained in this field. Recently, however, a remarkable new series of developments was begun; it arose from the sharpened form of the estimates for  $\Lambda$  mentioned earlier, involving, in particular, the best possible dependence on  $A_n$ . First it was noted that the estimates yielded at once an explicit upper bound for all solutions in integers  $x > 1$ ,  $y > 1$ ,  $n > 2$  of the equation  $ax^n - by^n = c$ , where  $a, b, c$  ( $\neq 0$ ) are any given integers. Next Tijdeman succeeded in effectively resolving the famous Catalan equation  $x^p - y^q = 1$  in integers  $x, y, p, q$  (all  $> 1$ ). Shortly afterwards, Schinzel and Tijdeman showed that one could likewise resolve the equation  $y^m = f(x)$ , of hyperelliptic

---

† Compare the papers by Waldschmidt in this volume.

type, in integers  $x, y > 1, m > 2$ ; and many novel results have followed. For instance, Shorey and Tijdeman have effectively dealt with the equation

$$y^m = x^n + x^{n-1} + \dots + 1$$

in integers  $x, y, m, n$  (all  $> 1$ ), when either  $x$  is fixed, or  $n+1$  or  $y$  has a fixed prime factor; and Györy, Tijdeman and Voorhoeve† have shown that, for any integer  $k \geq 6$ , the equation

$$y^m = 1^k + 2^k + \dots + x^k$$

has only finitely many solutions in integers  $x \geq 1, y \geq 1, m > 1$ , all of which can, in principle, be effectively determined. In another direction, van der Poorten‡ has used  $p$ -adic analysis to resolve effectively the equation  $x^p - y^q = z^r$  in integers  $x, y, z, p, q$  (all  $> 1$ ), where  $z$  is composed solely of powers of a fixed set of primes and  $r$  is the least common multiple of  $p$  and  $q$ . And recently Stewart§ has shown that the Fermat equation  $x^n + y^n = z^n$  has only finitely many solutions in positive integers  $x, y, z$  and  $n (> 2)$ , provided only that  $|x - y|$  is bounded.

To illustrate the latest advances, let us outline a proof of the last result. We can plainly assume that  $x, y, z$  are relatively prime and that  $x < y < z$ ; then  $z - x = kX^n, z - y = lY^n$  for some integers  $X, Y$  and some rationals  $k, l$  of the form  $1/d$  or  $2/d$  where  $d$  is a divisor of  $n$ . It follows that  $|kX^n - lY^n|$  is bounded, and so  $|\Lambda| \ll n/X^n$ , where

$$\Lambda = \log(k/l) + n \log(X/Y),$$

and the implied constant is absolute. This gives  $\log|\Lambda| \ll -n \log X + \log n$ . But we can assume that  $X > 1$ , and then, from the theorem quoted in Section 3, we obtain  $\log|\Lambda| \gg -(\log n)^3 \log X$ . A comparison of estimates shows at once that  $n$  is bounded; hence, in view of the work mentioned earlier on the hyperelliptic equation, also  $x, y, z$  are bounded, and this establishes the result. Clearly, in principle, all solutions can be determined by a finite amount of computation.

A similar argument applies to the Catalan equation  $x^p - y^q = 1$ . Here we can assume that  $p, q$  are odd primes, and it will suffice to treat the case  $p > q$ . Then  $x = kX^q + 1, y = lY^p - 1$  for some integers  $X, Y$ , where  $k$  is 1 or  $1/p$  and  $l$  is 1 or  $1/q$ . Plainly we have  $|p \log x - q \log y| \ll y^{-q}$ , and thus  $|\Lambda| \ll (kX^q)^{-1}$ , where

$$\Lambda = p \log k - q \log l + pq \log(X/Y).$$

But now an application of the theorem quoted in Section 3 shows that  $q \ll (\log p)^4$ . Further it is clear that the linear form

$$\Lambda' = p \log(x/Y^q) - q \log l$$

† To appear.

‡ *Acta Arithmetica* 33 (1977), 195–207.

§ *Mathematika* 24 (1977), 130–132. A similar result was obtained independently by Inkeri and van der Poorten.

satisfies  $|\Lambda'| \ll (LY^p)^{-1}$ , and hence, from a second application of the theorem, we obtain  $p \ll q(\log p)^3$ . Thus  $p$  is bounded, and, as in the previous proof, the required result now follows from the theorem referred to earlier on the hyperelliptic equation. It is easy to write down explicit bounds for  $p, q, x, y$ , but they are large, and there would seem to be no likelihood, at present, of dealing with the remaining possibilities even with a computing machine. I might remark, however, that examples in this field are known where bounds of the order of  $10^{10500}$  have been reduced to manageable figures by means of simple observations from the theory of Diophantine approximation.†

### 5. Elliptic and Abelian functions

Finally, I should like to mention briefly some recent generalizations of the theory of linear forms in logarithms in the context of elliptic curves.

We define now

$$\Lambda = \beta_0 + \beta_1 u_1 + \dots + \beta_n u_n,$$

where the  $\beta$ 's are again algebraic numbers, not all 0, and  $u_1, \dots, u_n$  are algebraic points of a Weierstrass elliptic function  $\wp(z)$  with algebraic invariants; by an algebraic point of  $\wp$  we mean a number  $u$  such that either  $\wp(u)$  is algebraic or  $u$  is a pole of  $\wp$ . Schneider proved in 1937 that  $\Lambda \neq 0$  when  $n = 2$ ,  $\beta_0 = 0$ , provided that  $\wp(u_1 z)$ ,  $\wp(u_2 z)$  are algebraically independent. In 1968, I proved that  $\Lambda \neq 0$  when  $n = 2$ ,  $\beta_0 \neq 0$  in the case when  $u_1$  and  $u_2$  are fundamental periods of  $\wp(z)$ ; in 1975, Masser showed that  $\Lambda \neq 0$  in the case when  $\wp(z)$  has complex multiplication and  $u_1, \dots, u_n$  are linearly independent over the corresponding complex quadratic field. Subsequently a number of authors have given quantitative refinements of this result; see the paper by Anderson in the proceedings cited earlier. The estimates at present, however, are somewhat weaker than those established in the case of linear forms in logarithms. The main reason for the extra difficulty encountered in the elliptic case is that the Weierstrass functions have order 2, and not 1 as for the exponential function. Thus it is necessary to use a doubly-dense sequence of extrapolation points, and not just the integers as suffices in the earlier context. This leads one to some rather delicate questions relating, for instance, to the division value properties of the Weierstrass functions; in particular, some recent work of Bashmakov and Ribet on the subject has proved useful. The methods can be generalized, as Masser, Lang and Coates have shown, to deal also with Abelian functions, provided that one makes suitable assumptions concerning complex multiplications; and, furthermore, the theory can be carried over to the  $p$ -adic domain for a wide class of primes  $p$ , as has been demonstrated recently by Bertrand and Flicker.

† See A. Baker and H. Davenport, *Quart. J. Math. Oxford, Ser (2)*, **20** (1969), 129–137.

Well, I should say before closing that I have touched upon only a few of the topics to which transcendence theory has now been successfully applied. More especially, I have made no mention of the penetrating role it has played in connexion with class number problems or questions concerning divisor properties of arithmetical sequences. Nevertheless, I hope that I have said enough to give you an adequate impression of this fast-growing subject, which has plainly moved well beyond the rudimentary researches of Liouville and Cantor with which it all began.

Trinity College  
Cambridge  
England