1

THE ORIGINS

1. Liouville's theorem

The theory of transcendental numbers was originated by Liouville in his famous memoir[†] of 1844 in which he obtained, for the first time, a class, *très-étendue*, as it was described in the title of the paper, of numbers that satisfy no algebraic equation with integer coefficients. Some isolated problems pertaining to the subject, however, had been formulated long before this date, and the closely related study of irrational numbers had constituted a major focus of attention for at least a century preceding. Indeed, by 1744, Euler had already established the irrationality of e, and, by 1761, Lambert had confirmed the irrationality of π. Moreover, the early studies of continued fractions had revealed several basic features concerning the approximation of irrational numbers by rationals. It was known, for instance, that for any irrational α there exists an infinite sequence of rationals p/q $(q > 0)$ such that[‡] $|\alpha - p/q| < 1/q^2$, and it was known also that the continued fraction of a quadratic irrational is ultimately periodic, whence there exists $c = c(\alpha) > 0$ such that $|\alpha - p/q| > c/q^2$ for all rationals p/q $(q > 0)$. Liouville observed that a result of the latter kind holds more generally, and that there exists in fact a limit to the accuracy with which any algebraic number, not itself rational, can be approximated by rationals. It was this observation that provided the first practical criterion whereby transcendental numbers could be constructed.

Theorem 1.1. *For any algebraic number α with degree $n > 1$, there exists $c = c(\alpha) > 0$ such that $|\alpha - p/q| > c/q^n$ for all rationals p/q $(q > 0)$.*

The theorem follows almost at once from the definition of an algebraic number. A real or complex number is said to be algebraic if it is a zero of a polynomial with integer coefficients; every algebraic

[†] *C.R.* **18** (1844), 883–5, 910–11; *J. Math. pures appl.* **16** (1851), 133–42. For abbreviations see page 130.

[‡] This is in fact easily verified; for any integer $Q > 1$, two of the $Q+1$ numbers 1, $\{q\alpha\}$ $(0 \leqslant q < Q)$, where $\{q\alpha\}$ denotes the fractional part of $q\alpha$, lie in one of the Q subintervals of length $1/Q$ into which $[0, 1]$ can be divided, and their difference has the form $q\alpha - p$.

number α is the zero of some such irreducible[†] polynomial, say P, unique up to a constant multiple, and the degree of α is defined as the degree of P. It suffices to prove the theorem when α is real; in this case, for any rational p/q $(q > 0)$, we have by the mean value theorem:

$$-P(p/q) = P(\alpha) - P(p/q) = (\alpha - p/q)\,P'(\xi)$$

for some ξ between p/q and α. Clearly one can assume that $|\alpha - p/q| < 1$, for the result would otherwise be valid trivially; then $|\xi| < 1 + |\alpha|$ and thus $|P'(\xi)| < 1/c$ for some $c = c(\alpha) > 0$; hence

$$|\alpha - p/q| > c\,|P(p/q)|.$$

But, since P is irreducible, we have $P(p/q) \neq 0$, and the integer $|q^n P(p/q)|$ is therefore at least 1; the theorem follows. Note that one can easily give an explicit value for c; in fact one can take

$$c^{-1} = n^2(1 + |\alpha|)^{n-1} H,$$

where H denotes the height of α, that is, the maximum of the absolute values of the coefficients of P.

A real or complex number that is not algebraic is said to be transcendental. In view of Theorem 1.1, an obvious instance of such a number is given by $\xi = \sum_{n=1}^{\infty} 10^{-n!}$. For if we write

$$p_j = 10^{j!} \sum_{n=1}^{j} 10^{-n!}, \qquad q_j = 10^{j!} \quad (j = 1, 2, \ldots),$$

then p_j, q_j are relatively prime rational integers and we have

$$|\xi - p_j/q_j| = \sum_{n=j+1}^{\infty} 10^{-n!}$$
$$< 10^{-(j+1)!}(1 + 10^{-1} + 10^{-2} + \ldots) = \tfrac{10}{9}\, q_j^{-j-1} < q_j^{-j}.$$

Many other transcendental numbers can be specified on the basis of Liouville's theorem; indeed any non-terminating decimal in which there occur sufficiently long blocks of zeros, or any continued fraction in which the partial quotients increase sufficiently rapidly, provides an example. Numbers of this kind, that is real ξ which possess a sequence of distinct rational approximations p_n/q_n $(n = 1, 2, \ldots)$ such that $|\xi - p_n/q_n| < 1/q_n^{\omega_n}$, where $\limsup \omega_n = \infty$, have been termed Liouville numbers; and, of course, these are transcendental. But other,

[†] That is, does not factorize over the integers or, equivalently, by Gauss' lemma, over the rationals.

less obvious, applications of Liouville's idea to the construction of transcendental numbers have been described; in particular, Maillet[†] used an extension of Theorem 1.1 concerning approximations by quadratic irrationals to establish the transcendence of a remarkable class of quasi-periodic continued fractions.[‡]

In 1874, Cantor introduced the concept of countability and this leads at once to the observation that 'almost all' numbers are transcendental. Cantor's work may be regarded as the forerunner of some important metrical theory about which we shall speak in Chapter 9.

2. Transcendence of *e*

In 1873, there appeared Hermite's epoch-making memoir entitled *Sur la fonction exponentielle*[§] in which he established the transcendence of e, the natural base for logarithms. The irrationality of e had been demonstrated, as remarked earlier, by Euler in 1744, and Liouville had shown in 1840, directly from the defining series, that in fact neither e nor e^2 could be rational or a quadratic irrational; but Hermite's work began a new era. In particular, within a decade, Lindemann succeeded in generalizing Hermite's methods and, in a classic paper,[||] he proved that π is transcendental and solved thereby the ancient Greek problem concerning the quadrature of the circle. The Greeks had sought to construct, with ruler and compasses only, a square with area equal to that of a given circle. This plainly amounts to constructing two points in the plane at a distance $\sqrt{\pi}$ apart, assuming that a unit length is prescribed. But, since all points capable of construction are defined by the intersection of lines and circles, it follows easily that their co-ordinates in a suitable frame of reference are given by algebraic numbers. Thus the transcendence of π implies that the quadrature of the circle is impossible.

The work of Hermite and Lindemann was simplified by Weierstrass[¶] in 1885, and further simplified by Hilbert,[††] Hurwitz[‡‡] and Gordan[§§] in 1893. We proceed now to demonstrate the transcendence of e and π in a style suggested by these later writers.

[†] See Bibliography.

[‡] Cf. *Mathematika*, **9** (1962), 1–8.

[§] *C.R.* **77**; = *Oeuvres* III, 150–81.

[||] *M.A.* **20** (1882), 213–25.

[¶] *Werke* II, 341–62.

[††] *Ges. Abh.* I, 1–4.

[‡‡] *Göttingen Nachrichten* (1893), 153–5.

[§§] *M.A.* **43** (1893), 222–5.

Theorem 1.2. *e is transcendental.*

Preliminary to the proof, we observe that if $f(x)$ is any real polynomial with degree m, say, and if

$$I(t) = \int_0^t e^{t-u} f(u)\,du,$$

where t is an arbitrary complex number and the integral is taken over the line joining 0 and t, then, by repeated integration by parts, we have[†]

$$I(t) = e^t \sum_{j=0}^{m} f^{(j)}(0) - \sum_{j=0}^{m} f^{(j)}(t). \tag{1}$$

Further, if $\bar{f}(x)$ denotes the polynomial obtained from f by replacing each coefficient with its absolute value, then

$$|I(t)| \leqslant \int_0^t |e^{t-u} f(u)|\,du \leqslant |t|\,e^{|t|}\bar{f}(|t|). \tag{2}$$

Suppose now that e is algebraic, so that

$$q_0 + q_1 e + \ldots + q_n e^n = 0 \tag{3}$$

for some integers $n > 0$, $q_0 \neq 0$, $q_1, \ldots, q_n$. We shall compare estimates for

$$J = q_0 I(0) + q_1 I(1) + \ldots + q_n I(n),$$

where $I(t)$ is defined as above with

$$f(x) = x^{p-1}(x-1)^p \ldots (x-n)^p,$$

p denoting a large prime. From (1) and (3) we have

$$J = -\sum_{j=0}^{m} \sum_{k=0}^{n} q_k f^{(j)}(k),$$

where $m = (n+1)p-1$. Now clearly $f^{(j)}(k) = 0$ if $j < p$, $k > 0$ and if $j < p-1$, $k = 0$, and thus for all j, k other than $j = p-1$, $k = 0$, $f^{(j)}(k)$ is an integer divisible by $p!$; further we have

$$f^{(p-1)}(0) = (p-1)!\,(-1)^{np}\,(n!)^p,$$

whence, if $p > n$, $f^{(p-1)}(0)$ is an integer divisible by $(p-1)!$ but not by $p!$. It follows that, if also $p > |q_0|$, then J is a non-zero integer divisible by $(p-1)!$ and thus $|J| \geqslant (p-1)!$. But the trivial estimate $\bar{f}(k) \leqslant (2n)^m$ together with (2) gives

$$|J| \leqslant |q_1|\,e\bar{f}(1) + \ldots + |q_n|\,n e^n \bar{f}(n) \leqslant c^p$$

for some c independent of p. The estimates are inconsistent if p is sufficiently large and the contradiction proves the theorem.

[†] $f^{(j)}(x)$ denotes the jth derivative of f.

Theorem 1.3. *π is transcendental.*

Suppose the contrary, that π is algebraic; then also $\theta = i\pi$ is algebraic. Let θ have degree d, let $\theta_1 (=\theta), \theta_2, \ldots, \theta_d$ denote the conjugates of θ and let l signify the leading coefficient in the minimal polynomial[†] defining θ. From Euler's equation $e^{i\pi} = -1$, we obtain

$$(1+e^{\theta_1})(1+e^{\theta_2})\ldots(1+e^{\theta_d}) = 0.$$

The product on the left can be written as a sum of 2^d terms e^Θ, where

$$\Theta = \epsilon_1\theta_1 + \ldots + \epsilon_d\theta_d,$$

and $\epsilon_j = 0$ or 1; we suppose that precisely n of the numbers

$$\epsilon_1\theta_1 + \ldots + \epsilon_d\theta_d$$

are non-zero, and we denote these by $\alpha_1, \ldots, \alpha_n$. We have then

$$q + e^{\alpha_1} + \ldots + e^{\alpha_n} = 0, \tag{4}$$

where q is the positive integer $2^d - n$.

We shall compare estimates for

$$J = I(\alpha_1) + \ldots + I(\alpha_n),$$

where $I(t)$ is defined as in the proof of Theorem 1.2 with

$$f(x) = l^{np}x^{p-1}(x-\alpha_1)^p \ldots (x-\alpha_n)^p,$$

p again denoting a large prime. From (1) and (4) we have

$$J = -q\sum_{j=0}^{m} f^{(j)}(0) - \sum_{j=0}^{m}\sum_{k=1}^{n} f^{(j)}(\alpha_k),$$

where $m = (n+1)p-1$. Now the sum over k is a symmetric polynomial in $l\alpha_1, \ldots, l\alpha_n$ with integer coefficients, and it follows from two applications of the fundamental theorem on symmetric functions together with the observation that each elementary symmetric function in $l\alpha_1, \ldots, l\alpha_n$ is also an elementary symmetric function in the 2^d numbers $l\Theta$, that it represents a rational integer. Further, since $f^{(j)}(\alpha_k) = 0$ when $j < p$, the latter is plainly divisible by $p!$. Clearly also $f^{(j)}(0)$ is a rational integer divisible by $p!$ when $j \neq p-1$, and

$$f^{(p-1)}(0) = (p-1)!(-l)^{np}(\alpha_1\ldots\alpha_n)^p$$

[†] That is, the irreducible polynomial indicated earlier with relatively prime integer coefficients; the coefficient of x^d is called the leading coefficient, and it is assumed positive. The conjugates are the zeros of the polynomial.

is a rational integer divisible by $(p-1)!$ but not by $p!$ if p is sufficiently large. Hence, if $p > q$, we have $|J| \geqslant (p-1)!$. But from (2) we obtain

$$|J| \leqslant |\alpha_1|\, e^{|\alpha_1|}\bar{f}(|\alpha_1|)+\dots+|\alpha_n|\, e^{|\alpha_n|}\bar{f}(|\alpha_n|) \leqslant c^p$$

for some c independent of p. The estimates are inconsistent for p sufficiently large, and the contradiction proves the theorem.

3. Lindemann's theorem

The two preceding theorems, that is the transcendence of e and π, are special cases of a much more general result which Lindemann sketched in his original memoir of 1882, and which was later rigorously demonstrated by Weierstrass.

Theorem 1.4. *For any distinct algebraic numbers $\alpha_1, \dots, \alpha_n$ and any non-zero algebraic numbers $\beta_1, \dots, \beta_n$ we have*

$$\beta_1 e^{\alpha_1}+\dots+\beta_n e^{\alpha_n} \neq 0.$$

It follows at once from Theorem 1.4 that $e^{\alpha_1}, \dots, e^{\alpha_n}$ are algebraically independent for all algebraic $\alpha_1, \dots, \alpha_n$ linearly independent over the rationals; this form of the result is generally known as Lindemann's theorem. As further immediate corollaries of Theorem 1.4, one sees that $\cos\alpha$, $\sin\alpha$ and $\tan\alpha$ are transcendental for all algebraic $\alpha \neq 0$, and moreover $\log\alpha$ is transcendental for algebraic α not 0 or 1.

Suppose now that the theorem is false, so that

$$\beta_1 e^{\alpha_1}+\dots+\beta_n e^{\alpha_n} = 0. \tag{5}$$

One can clearly assume, without loss of generality, that the β's are rational integers, for this can be ensured by multiplying (5) by all the expressions obtained on allowing $\beta_1, \dots, \beta_n$ on the left to run independently through their respective conjugates and then further multiplying by a common denominator. Furthermore, one can assume that there exist integers $0 = n_0 < n_1 < \dots < n_r = n$, such that $\alpha_{n_t+1}, \dots, \alpha_{n_{t+1}}$ is a complete set of conjugates for each t, and

$$\beta_{n_t+1} = \dots = \beta_{n_{t+1}}.$$

For certainly $\alpha_1, \dots, \alpha_n$ are zeros of some polynomial with integer coefficients and degree N, say, and if $\alpha_{n+1}, \dots, \alpha_N$ denote the remaining zeros, we have

$$\Pi(\beta_1 e^{\alpha_{k_1}}+\dots+\beta_N e^{\alpha_{k_N}}) = 0,$$

where the product is over all permutations $k_1, \dots, k_N$ of $1, \dots, N$ and

$\beta_{n+1} = \ldots = \beta_N = 0$. The left-hand side can be expressed as an aggregate of terms $\exp(h_1\alpha_1 + \ldots + h_N\alpha_N)$ with integer coefficients, where $h_1, \ldots, h_N$ are integers with sum $N!$, and clearly $h_1\alpha_{k_1} + \ldots + h_N\alpha_{k_N}$ taken over all permutations $k_1, \ldots, k_N$ of $1, \ldots, N$ is a complete set of conjugates; the condition concerning the equality of the β's follows by symmetry. Note also that, after collecting terms with the same exponents, one at least of the new coefficients β will be non-zero; this is readily confirmed by considering the coefficient of the term with exponent that is highest according to the ordering of the complex numbers $z = x+iy$ given by $z_1 < z_2$ if $x_1 < x_2$ or $x_1 = x_2$ and $y_1 < y_2$.

Let now l be any positive integer such that $l\alpha_1, \ldots, l\alpha_n$ and $l\beta_1, \ldots, l\beta_n$ are algebraic integers,[†] and let

$$f_i(x) = l^{np}\{(x-\alpha_1)\ldots(x-\alpha_n)\}^p/(x-\alpha_i) \quad (1 \leqslant i \leqslant n),$$

where p denotes a large prime. We shall compare estimates for $|J_1 \ldots J_n|$, where

$$J_i = \beta_1 I_i(\alpha_1) + \ldots + \beta_n I_i(\alpha_n) \quad (1 \leqslant i \leqslant n),$$

and $I_i(t)$ is defined as in the proof of Theorem 1.2, with $f = f_i$. From (1) and (5) we have

$$J_i = -\sum_{j=0}^{m}\sum_{k=1}^{n}\beta_k f_i^{(j)}(\alpha_k),$$

where $m = np-1$. Further, $f_i^{(j)}(\alpha_k)$ is an algebraic integer divisible[‡] by $p!$ unless $j = p-1$, $k = i$; and in the latter case we have

$$f_i^{(p-1)}(\alpha_i) = l^{np}(p-1)!\prod_{\substack{k=1\\k\neq i}}^{n}(\alpha_i-\alpha_k)^p,$$

so that it is an algebraic integer divisible by $(p-1)!$ but not by $p!$ if p is sufficiently large. It follows that J_i is a non-zero algebraic integer divisible by $(p-1)!$. Further, by the initial assumptions, we have

$$J_i = -\sum_{j=0}^{m}\sum_{t=0}^{r-1}\beta_{n_{t+1}}\{f_i^{(j)}(\alpha_{n_t+1}) + \ldots + f_i^{(j)}(\alpha_{n_{t+1}})\},$$

and here each sum over t can be expressed as a polynomial in α_i with rational coefficients independent of i; for clearly, since $\alpha_1, \ldots, \alpha_n$ is a complete set of conjugates, the coefficients of $f_i^{(j)}(x)$ can be expressed in this form. Thus $J_1 \ldots J_n$ is rational, and so in fact a rational integer

[†] An algebraic number is said to be an algebraic integer if the leading coefficient in its minimal defining polynomial is 1; if α is an algebraic number and l is the leading coefficient in its minimal polynomial, then $l\alpha$ is an algebraic integer.

[‡] That is, the quotient is an algebraic integer.

divisible by $((p-1)!)^n$. Hence we have $|J_1 \ldots J_n| \geqslant (p-1)!$. But (2) gives

$$|J_i| \leqslant \sum_{k=1}^{n} |\alpha_k \beta_k|\, e^{|\alpha_k|} \bar{f}(|\alpha_k|) \leqslant c^p,$$

for some c independent of p, and the inequalities are inconsistent if p is sufficiently large. The contradiction proves the theorem.

The above proofs are simplified versions of the original arguments of Hermite and Lindemann and their motivation may seem obscure; indeed there is no explanation *a priori* for the introduction of the functions I and f. A deeper insight can best be obtained by studying the basic memoir of Hermite where, in modified form, the functions first occurred, but it may be said that they relate to generalizations, concerning simultaneous approximation, of the convergents in the continued fraction expansion of e^x. Further light on the topic will be shed by Chapters 10 and 11. Lindemann's theorem formed the summit of the accomplishments of the last century, and our survey of the period is herewith concluded.

2

LINEAR FORMS IN LOGARITHMS

1. Introduction

In 1900, at the International Congress of Mathematicians held in Paris, Hilbert raised, as the seventh of his famous list of 23 problems, the question whether an irrational logarithm of an algebraic number to an algebraic base is transcendental. The question is capable of various alternative formulations; thus one can ask whether an irrational quotient of natural logarithms of algebraic numbers is transcendental, or whether α^β is transcendental for any algebraic number $\alpha \neq 0, 1$ and any algebraic irrational β. A special case relating to logarithms of rational numbers can be traced to the writings of Euler more than a century before, but no apparent progress had been made towards its solution. Indeed, Hilbert expressed the opinion that the resolution of the problem lay farther in the future than a proof of the Riemann hypothesis or Fermat's last theorem.

The first significant advance was made by Gelfond[†] in 1929. Employing interpolation techniques of the kind that he had utilized previously in researches on integral integer-valued functions,[‡] Gelfond showed that the logarithm of an algebraic number to an algebraic base cannot be an imaginary quadratic irrational, that is, α^β is transcendental for any algebraic number $\alpha \neq 0, 1$ and any imaginary quadratic irrational β; in particular, this implies that $e^\pi = (-1)^{-i}$ is transcendental. The result was extended to real quadratic irrationals β by Kuzmin[§] in 1930. But it was clear that direct appeal to an interpolation series for $e^{\beta z}$, on which the Gelfond–Kuzmin work was based, was not appropriate for more general β, and further progress awaited a new idea. The search for the latter was concluded successfully by Gelfond[||] and Schneider[¶] independently in 1934. The arguments they discovered were applicable for any irrational β and, though differing in detail, both depended on the construction of an auxiliary function that vanished at certain selected points. A similar technique had been used a few years earlier by Siegel in the course of investigations on the

[†] *C.R.* 189 (1929), 1224–8.

[‡] *Tôhoku Math. J.* 30 (1929), 280–5.

[§] *I.A.N.* 3 (1930), 583–97.

[||] *D.A.N.* 2 (1934), 1–6; *I.A.N.* 7 (1934), 623–4.

[¶] *J.M.* 172 (1934), 65–9.

Bessel functions.[†] Herewith Hilbert's seventh problem was finally solved.

The Gelfond–Schneider theorem shows that for any non-zero algebraic numbers α_1, α_2, β_1, β_2, with $\log \alpha_1$, $\log \alpha_2$ linearly independent over the rationals, we have

$$\beta_1 \log \alpha_1 + \beta_2 \log \alpha_2 \neq 0.$$

It was natural to conjecture that an analogous theorem would hold for arbitrarily many logarithms of algebraic numbers, and, moreover, it was soon realized that such a result would be capable of wide application. The conjecture was proved by the author[‡] in 1966, and the demonstration will be the subject of the present chapter.

Theorem 2.1. *If $\alpha_1, \ldots, \alpha_n$ are non-zero algebraic numbers such that*[§] *$\log \alpha_1, \ldots, \log \alpha_n$ are linearly independent over the rationals, then 1, $\log \alpha_1, \ldots, \log \alpha_n$ are linearly independent over the field of all algebraic numbers.*

The proof depends on the construction of an auxiliary function of several complex variables which generalizes the function of a single variable employed originally by Gelfond. Functions of several variables were utilized by Schneider[‖] in his studies concerning Abelian integrals but, for many years, there appeared to be severe limitations to their serviceability in wider settings. The main difficulty concerned the basic interpolation techniques. Work in this connexion had hitherto always involved an extension in the order of the derivatives while leaving the points of interpolation fixed; however, when dealing with functions of several variables, 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. The proof of Theorem 2.1 involves an extrapolation procedure, special to the present context, in which the range of interpolation is now extended while the order of the derivatives is reduced. Refinements and generalizations will be discussed in the next chapter and applications of the results to various branches of number theory will be the theme of Chapters 4 and 5.

[†] *Abh. Preuss Akad. Wiss.* (1929), No. 1; cf. ch. 11.
[‡] *Mathematika*, **13** (1966), 204–16; **14** (1967), 102–7, 220–8.
[§] Here the logarithms can take any fixed values.
[‖] *J.M.* **183** (1941), 110–28.