ON A THEOREM OF KRONECKER AND A RELATED QUESTION OF LEHMER

by C.L. Stewart

§1. Kronecker's theorem.

Let $f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$ be a polynomial with integer coefficients of degree $n \geq 1$ and assume that $a_n \neq 0$. We define $M(f)$, the measure of the polynomial $f(x)$, to be

$$M(f) = |a_n| \prod_{i=1}^{n} \max\{1, |\alpha_i|\},$$

where $\alpha_1, \ldots, \alpha_n$ are the roots of $f(x)$. If α is an algebraic number then we define $M(\alpha)$ to be equal to $M(f)$ where f is the minimal polynomial of α. Equivalently, we have

$$M(\alpha) = \prod_{v} \max\{1, |\alpha|_v\}$$

where the product is taken over all the valuations $v^\dagger$ of the field $Q(\alpha)$. We remark that $M(f_1 f_2) = M(f_1)M(f_2)$ for any two polynomials f_1 and f_2 and that $M(\alpha\beta) \leq M(\alpha)M(\beta)$ for any two algebraic numbers α and β. Further we note that if α is a non-zero algebraic number which is not a unit then $M(\alpha) \geq 2$.

In 1857 Kronecker [8] proved that if α is a non-zero algebraic integer with $M(\alpha) = 1$ then α is a root of unity. The converse plainly holds. Kronecker observed that for any positive integer k, α^k is an algebraic integer of degree at most n all of whose conjugates are bounded by 1 in absolute value. Therefore α^k is the root of a polynomial $f_k(x) = x^\ell + a_1 x^{\ell-1} + \ldots + a_\ell$ with $1 \leq \ell \leq n$ and $|a_i| \leq \binom{\ell}{i}$ since the a_i are elementary symmetric functions of numbers having absolute value at most 1. Clearly there are only finitely many such polynomials and therefore for two distinct integers r and s we have $\alpha^r = \alpha^s$ whence

---

† The valuations are assumed to be normalized in the standard way.

$\alpha^{r-s} = 1$ and α is a root of unity as required.

Several different proofs of Kronecker's theorem have been given. For example, on noting that the powers of α all lie in the unit disc we can find, by means of the pigeon-hole principle, two distinct integers r and s for which $|\alpha^r-\alpha^s| < 2^{-n}$. The conjugates of $\alpha^r-\alpha^s$ satisfy $|\alpha_l{}^r-\alpha_l{}^s| \leq 2$ for $l = 1,\ldots,n$ and thus the norm of $\alpha^r-\alpha^s$ is less than 1 in absolute value, whence it is zero. Once again we conclude that $\alpha^r = \alpha^s$ and therefore that α is a root of unity.

Interestingly enough the theorem of Kronecker is implicit in Dirichlet's proof, published in 1846 [5], of the Dirichlet unit theorem. Dirichlet does not make this point explicit however.

§2. Lehmer's question.

In 1933 D.H. Lehmer (see p. 476 of [9]) asked the following question. Is it true that for every positive number ε there exists a polynomial $f(x)$, with integer coefficients, satisfying $1 < M(f) < 1+\varepsilon$? If the answer to Lehmer's question is no then Kronecker's theorem can be considerably strengthened, for then there exists a positive number ε_0 such that if α is any non-zero algebraic number with $M(\alpha) < 1+\varepsilon_0$ then α is a root of unity. Lehmer asked the above question in connexion with a method for finding large prime numbers. He considered the integers

$$\Delta_k = \prod_{l=1}^{n} (\alpha_l{}^k-1) \qquad \text{for } k = 1,2,\ldots,$$

where $\alpha_1,\alpha_2,\ldots,\alpha_n$ are the roots of an irreducible polynomial. The prime factors of Δ_k satisfy certain congruence conditions. For example, if k is a prime then all prime divisors p of Δ_k which do not divide Δ_1 or the discriminant of $Q(\alpha)$ have the property that one of $\{p,p^2,p^3,\ldots,p^n\}$ is congruent to $1 \pmod k$. If $|\Delta_k|$ is not too large with respect to k then the above congruence condition considerably restricts the possible prime factors of Δ_k and because of this allows one to factor Δ_k. The rate of growth of $|\Delta_k|$ is approximately $(M(\alpha))^k$ whence Lehmer's interest in small values of $M(\alpha)$. As an example he takes $n = 3$, α_1, α_2 and α_3 to be the roots of x^3-x-1, so that $M(\alpha) = 1.32471795\ldots$, and he finds that $\Delta_{127} = 3\ 233\ 514\ 251\ 032\ 733$ is a prime number.

The smallest value of $M(f)$ larger than 1 which Lehmer found was associated with the polynomial

$$f_0(x) = x^{10}+x^9-x^7-x^6-x^5-x^4-x^3+x+1 .$$

In this case $M(f_0) = \alpha_0 = 1.17628081\ldots$ where α_0 is the largest real root of $f_0(x)$. α_0 is a Salem number, a real algebraic integer larger than 1 having one conjugate on the unit circle and all other conjugates, apart from itself, on or inside the unit circle. It is an open question whether the set of Salem numbers is dense in $[1,\infty)$ although it is suspected that they are not dense. This would follow from a negative answer to the question of Lehmer. On the basis of a computer search Boyd [4] has even conjectured that α_0 is the smallest Salem number. In the general case no polynomial $f(x)$ with $1 < M(f) < \alpha_0$ has been found to date. We remark, however, that no extensive search

has been made. In fact to make an exhaustive computer search for the smallest values assumed by $M(f)$ when f runs through the polynomials with relatively low degree, less than 25 say, appears to be a quite difficult task. This is because, at least with the present arguments, the number of polynomials to be tested grows exponentially with the degree.

The problem of determining the set of values taken by $M(f)$ when f runs through the set of all polynomials with integer coefficients has arisen in ergodic theory. D.A. Lind [10] has recently proved that the set of possible values for the entropy of a continuous algebraic automorphism of a separable compact group is a countable subset of $[0,\infty]$ if the answer to Lehmer's question is no, while if the answer is yes it is all of $[0,\infty]$. Lind shows that the only type of group automorphism having a small positive entropy is an ergodic group automorphism, S say, of a torus. If the torus has finite dimension d then S induces a rational vector space isomorphism T of Q^d, here Q is the set of rational numbers (see pp. 213-215 of [10]). The characteristic polynomial $f(x)$ associated with T has integer coefficients and the entropy of S is given by the logarithm of $M(f)$, whence the connexion with Lehmer's question.

§3 Smyth's theorem.

Lehmer's question has been answered in the negative when the polynomials f are assumed to be non-reciprocal. A reciprocal polynomial $f(x)$ of degree n is a polynomial satisfying $f(x) \equiv x^n f(x^{-1})$. The

cyclotomic polynomials, with the exception of $x-1$, are reciprocal. In 1971 C.J. Smyth [15] proved that if $f(x)$ is a polynomial with integer coefficients which does not have 0 or 1 as a root and which is not reciprocal, then $M(f) \geq \beta_0$ where β_0 is the real root of x^3-x-1 and hence is 1.32471795... .

One immediate consequence of Smyth's theorem is that β_0 is the smallest Pisot-Vijayaraghavan number. A P.V. number is a real algebraic integer larger than 1 all of whose conjugates, apart from itself, lie strictly within the unit circle. It was shown by Salem [11] in 1944 that there exists a smallest P.V. number and in the same year Siegel [14] proved that β_0 is the smallest. The proofs of Salem, Siegel and Smyth follow the same general pattern. To illustrate this pattern we shall prove a weaker version of Smyth's result to the effect that if $f(z)$ is a non-reciprocal polynomial with integer coefficients which does not have 0 or 1 as a root then $M(f) \geq \sqrt{5}/2$.

Accordingly, we assume that $M(f) < 2$ and, since $M(f_1f_2) = M(f_1)M(f_2)$ and the measure of a polynomial is always at least 1, that $f(z)$ is irreducible. Let $\alpha_1,\ldots,\alpha_n$ be the roots of $f(z)$ and put $r(z) = z^n f(z^{-1})$. Since $f(z)$ is not reciprocal, the quotient $f(z)/r(z)$ is not constant and we may expand it in a power series as

$$\frac{f(z)}{r(z)} = a_0+a_k z^k+a_\ell z^\ell+\ldots \quad , \quad (0 < k < \ell<\ldots), \tag{1}$$

where the a_1's are non-zero integers. Note that both the leading coefficient and the constant coefficient of $f(z)$ have absolute value 1 since $M(f) < 2$; thus $|a_0| = 1$. Furthermore, $f(z)$ has no roots α on

the unit circle for if $\alpha\bar{\alpha} = 1$, where $\bar{\alpha}$ denotes the complex conjugate of α, then $\sigma(\alpha)\sigma(\bar{\alpha}) = 1$ for all embeddings σ of $Q(\alpha)$ into the complex numbers; hence if α is any root of $f(z)$ then α^{-1} is also a root of $f(z)$, and therefore either $f(z)$ is reciprocal or $\alpha = 1$ contrary to our assumptions. Thus we have $f(z)/(r(z)) = g(z)/(h(z))$ where

$$(2) \qquad g(z) = \prod_{|\alpha_j|<1} \left(\frac{z-\alpha_j}{1-\bar{\alpha}_j z}\right) = c+c_1 z+c_2 z^2+\dots ,$$

and

$$(3) \qquad h(z) = \prod_{|\alpha_j|>1} \left(\frac{1-\bar{\alpha}_j z}{z-\alpha_j}\right) = d+d_1 z+d_2 z^2+\dots .$$

On comparing the series (1), (2) and (3) we find that $a_k d+a_0 d_k = c_k$. We have $|a_0| = 1$ and, since a_k is a non-zero integer, $|a_k| \geq 1$ whence

$$(4) \qquad \max\{|d_k|,|c_k|\} \geq |d|/2.$$

Both $g(z)$ and $h(z)$ are holomorphic in a neighbourhood containing the unit disc. Thus, by Parseval's identity,

$$\frac{1}{2\pi} \int_0^{2\pi} |g(e^{i\theta})|^2 d\theta = |c|^2+|c_1|^2+|c_2|^2+\dots$$

and therefore, since $g(z)$ has absolute value 1 on the unit circle, $|c_k|^2 \leq 1-|c|^2$. Similarly, we find that $|d_k|^2 \leq 1-|d|^2$. From (4) and the observation that $|c| = |d| = M(f)^{-1}$ we conclude that $M(f) \geq \sqrt{5}/2$ as required. Smyth obtained his best possible result by a more refined use of Parseval's identity than that given above.

Schinzel [12] has used Smyth's theorem for his work concerning the number of irreducible factors of a polynomial. In particular, he has proved that any trinomial $f(x) = a_n x^n + a_m x^m + a_0$ has at most $(\log(a_n^2 + a_m^2 + a_0^2))/(\log \beta_0 + o(1))$ irreducible non-cyclotomic factors.

§4 Recent advances concerning Lehmer's question.

Lehmer's question in the general case remains open despite some progress in the last few years. In 1971 Blanksby and Montgomery [3] improved considerably upon previous estimates [2], [13] by showing that if $f(x)$ is a polynomial of degree n with

$$M(f) < 1 + (52n \log 6n)^{-1}, \tag{5}$$

then $M(f) = 1$. Let $\alpha_1, \ldots, \alpha_n$ be the roots of $f(x)$ and assume that $f(x)$ is irreducible. They show that if (5) holds then, for some positive integer k,

$$\prod_{j=1}^{n} |\alpha_j^k - 1| < 1. \tag{6}$$

Thus the α_j's are roots of unity since the above product is an integer and hence 0. Put $i\theta_j = \log(\alpha_j/|\alpha_j|)$ for $j = 1, \ldots, n$. The product in (6) grows approximately as $M(f)^k$ and is small only when there are terms $k\theta_j$ which are close to 0 modulo 2π. One way of finding an integer k such that the $k\theta_j$'s are all close to 0 modulo 2π, and thus such that (6) holds for $M(f)$ sufficiently small, is to use the pigeon-hole principle. Ideally one would like to choose k to be small. However, the pigeon-hole

principle is too crude in this context and it only leads to a proof that there exists a positive number c such that if $M(f) < 1+c^{-n}$ then $M(f) = 1$. The more efficient approach of Blanksby and Montgomery depends upon an averaging argument. They consider the function

$$g(x_1,\ldots,x_n) = -\sum_{j=1}^{n} \log \left| \rho_j \exp(2\pi i x_j) - 1 \right|,$$

where $\rho_1,\ldots,\rho_n$ are appropriately chosen positive numbers less than 1. By using the non-negativity of the Fejér kernel applied to coefficients of terms in the Fourier expansions of functions of the form

$$\sum_{k=1}^{K} \left(1 - \frac{k}{K+1}\right) g(kx_1,\ldots,kx_n),$$

they deduce that for some small positive integer k, $g(k\theta_1,\ldots,k\theta_n)$ is so large that (6) holds. Their method of proof is related to work of Turán concerning estimates for power sums; see [16] for an illustration of this link.

In 1977 the author found a new method of attacking Lehmer's question based upon ideas from the theory of transcendental numbers. We proved, [17]: if α is a non-zero algebraic integer of degree n, at least 2, and $M(\alpha) < 1+(10^4 n \log n)^{-1}$ then α is a root of unity. While the constant is less precise, the dependence on n in the above inequality is the same as that given by Blanksby and Montgomery. We construct an exponential polynomial $g(z)$ of the form,

$$g(z) = e^{-10z} \sum_{k=1}^{K} \sum_{d=1}^{n} a_{k,d} \alpha^{d} e^{(\log \alpha^{r} k)z},$$

where the integers $r_1,\ldots,r_k$ and the real number θ are chosen so that $|\mathrm{Im}(\log \alpha^{r_k})-\theta|$ is small for $k = 1,\ldots,K$; here $\mathrm{Im}(x)$ denotes the imaginary part of x and $\log x$ denotes the principal value of the logarithm of x. Further, the $a_{k,d}$ are integers, not all of which are zero, with small absolute value. They are chosen, by means of a modified version of Siegel's lemma concerning integer solutions of systems of linear equations, so that $g(u) = 0$ for the first U positive integers. By construction $g(z)$ grows slowly as a function of z. This fact, combined with the zeros $g(z)$ has, allows one to show, using the maximum modulus principle, the norm inequality and the estimate for $M(\alpha)$, that $g(u) = 0$ for all positive integers u. It then follows directly that α is a root of unity as required.

This approach was taken up by Dobrowolski [7] recently (see also [6]). He proved there exists a positive number c such that if α is a non-zero algebraic integer of degree n, (> 2), and

$$M(\alpha) < 1 + c\left(\frac{\log\log n}{\log n}\right)^3 ,$$

then α is a root of unity. Dobrowolski's theorem is the most precise response given to date to Lehmer's question. Let $\alpha = \alpha_1,\ldots,\alpha_n$ be the conjugates of α and let $f(x)$ be the irreducible polynomial associated with α. Dobrowolski employs the following congruence relation: for any prime number p,

$$(7) \qquad \prod_{i=1}^{n} |f(\alpha_i^p)| \equiv 0 \pmod{p^n}.$$

He constructs a polynomial $g(x)$ which is divisible by a high power of $f(x)$ and yet which has coefficients which are small in absolute value. Thus

$$g(x) = f(x)^k h(x) = a_0 + a_1 x + \ldots + a_N x^N,$$

where the $|a_i|$'s are small and $h(x)$ is a polynomial with integer coefficients. To effect this construction he appeals to a version of Siegel's lemma. Dobrowolski uses the fact that the $|a_i|$'s are small to show that $g(\alpha_i^p)$ is also small for several prime numbers p. He then employs (7) and the norm inequality to deduce that $g(x)$ is zero at the points α_i^p for $i = 1,\ldots,n$ and for at least $\frac{N}{n} + 1$ prime numbers p. Since $g(x)$ has degree N, we have $\alpha_i^p = \alpha_j^q$ where p and q are prime numbers; if $p \neq q$ then α is a root of unity as required, otherwise the result follows by induction on the degree of α. In the next section we shall give a short proof which illustrates Dobrowolski's argument.

§5 The set of values of $M(f)$.

An interesting question which is more general than the question asked by Lehmer is the following. What is the structure of the set $R = \{M(f) \mid f(x)$ a polynomial with integer coefficients$\}$? For example, is there a smallest element of $R\backslash\{1\}$? What does the set of limit points of R look like? If the answer to Lehmer's question is yes then it is easy to see that R is dense in $[1,\infty)$. If the answer is no, as I suspect,

then the structure of R is likely to be more complicated.

In letters to the author both Boyd and Smyth have remarked that R contains some quite small limit points. Put $f_n(x) = x^{2n}-x^{2n-1}-x^n-x+1$ for $n = 1,2,\ldots$. They showed that $\lim_{n\to\infty} M(f_n) = 1.285\ldots$. Boyd has found an even smaller limit point of R. He has shown that if $f_n(x) = x^{2n}-x^{2n-1}+x^{n+1}-x^n+x^{n-1}-x+1$ then $\lim_{n\to\infty} M(f_n) = 1.255\ldots$. Let $L(f)$ denote the sum of the absolute values of the coefficients of $f(x)$. If $\{f_n(x)\}_{n=1}^{\infty}$ is a sequence of irreducible polynomials with measure larger than 1 satisfying $\lim_{n\to\infty} M(f_n) = 1$, then $\lim_{n\to\infty} L(f_n) = \infty$. For, as we remarked with M. Mignotte and M. Waldschmidt, if $f(x)$ is an irreducible polynomial then either $M(f) = 1$ or $M(f) > 1+(6L(f))^{-1}$. Let $\alpha_1,\ldots,\alpha_n$ be the roots of $f(x)$. We observe that if α_i^p is a root of $f(x)$ for some prime number p then either $\alpha_i = 0$ or α_i is a root of unity. In both cases $M(f) = 1$. Otherwise we have, from (7), that

$$p^n \le \prod_{i=1}^{n} |f(\alpha_i^p)|, \qquad \text{for any prime number } p.$$

It is easily seen that

$$\prod_{i=1}^{n} |f(\alpha_i^p)| \le L(f)^n M(f)^{np},$$

whence

$$p \le L(f)M(f)^p.$$

By means of Bertrand's postulate we choose p to be between $2L(f)$ and

$4L(f)$ so that

$$2 \le M(f)^{4L(f)}.$$

Putting $M(f) = 1+x$ and noting that $\log(1+x) < x$ for $x > 0$ we see that $x > (\log 2)/(4L(f))$ as required.

§6 An elliptic analogue of Lehmer's question.

Let E be an elliptic curve expressed in Weierstrass normal form,

$$y^2 = 4x^3 - g_2x - g_3,$$

where g_2 and g_3 are algebraic numbers with $g_2^3 \neq 27g_3^2$. The set of algebraic points of E together with the point at infinity on E form an additive group $E(A)$. An algebraic point $P = (\alpha,\beta)$ of E is a point whose coordinates α and β are algebraic numbers. We define a height function for the algebraic points of E which is analogous to $M(\alpha)$ for algebraic numbers α; we put

$$H(P) = \prod_v \max\{1, |\alpha|_v, |\beta|_v\},$$

where the product is taken over all normalized valuations v of the field $Q(\alpha,\beta)$. Also, we define the height of the point at infinity on E to be 1. We then define the more tractable Tate height, $\hat{H}(P)$, by

$$\hat{H}(P) = \lim_{n\to\infty} (H(2^nP))^{2^{-2n}}.$$

The Tate height has the property that $\hat{H}(mP) = (\hat{H}(P))^{m^2}$ for all positive

integers m. Further, the set of algebraic points of E with bounded Tate height whose coordinates have bounded degree is finite. Therefore, $\hat{H}(P) = 1$ if and only if P is a point of finite order in $E(A)$. This is the analogue of Kronecker's theorem for algebraic numbers. Points of finite order in $E(A)$ correspond to roots of unity in the algebraic numbers.

D. Masser and a student of his, M. Anderson, have investigated the elliptic analogue of Lehmer's question: Is there a positive number ϵ, depending only on g_2 and g_3, such that $\hat{H}(P) > 1+\epsilon$, for every point P of $E(A)$ which is not a point of finite order? A simple counting argument shows, see [1], that if P is a point of infinite order in $E(A)$ whose coordinates generate a field of degree n over the rationals then

$$\hat{H}(P) > 1+c^{-n^2},$$

where c is a positive number which depends only on g_2 and g_3. By means of a proof, apparently similar to [17], which uses techniques from transcendence theory, Anderson [1] has obtained a considerable improvement on the above estimate in the case that E has complex multiplication. He has proved that in this case if P is a point of infinite order in $E(A)$ and if the coordinates of P generate a field of degree n over the rationals then

$$\hat{H}(P) > 1+c(n \log 2n)^{-3},$$

where c is a positive number which depends only on E.

REFERENCES

[1] M. Anderson, Ph.D. Thesis, The University of Nottingham, 1978.

[2] P.E. Blanksby, A note on algebraic integers, J. Number Theory 1 (1969), pp. 155-160.

[3] P.E. Blanksby and H.L. Montgomery, Algebraic integers near the unit circle, Acta Arith. 28 (1971), pp. 355-369.

[4] D. W. Boyd, Small Salem numbers, Duke Math. J. 44 (1977), pp. 315-328.

[5] G. Lejeune Dirichlet, Werke, (Zur Theorie der complexen Einheiten, (1846)), Chelsea, New York 1969, pp. 639-644.

[6] E. Dobrowolski, On the maximal modulus of conjugates of an algebraic integer, Bull. Acad. Polon. Sci. (to appear).

[7] E. Dobrowolski, On a question of Lehmer, Acta Arith. (to appear).

[8] L. Kronecker, Zwei Sätze über Gleichungen mit ganzzahligen Coefficienten, J. Reine Angew, Math. 53 (1857), pp. 173-175.

[9] D.H. Lehmer, Factorization of certain cyclotomic functions, Ann. Math. (2) 34 (1933), pp. 461-479.

[10] D.A. Lind, Skew products with group automorphisms, Israel J. Math. 28 (1977), pp. 205-248.

[11] R. Salem, A remarkable class of algebraic integers. Proof of a conjecture of Vijayaraghavan, Duke Math. J. 11 (1944), pp. 103-108.

[12] A. Schinzel, Reducibility of lacunary polynomials III, Acta Arith. 34 (1978), pp. 227-266.

[13] A. Schinzel and H. Zassenhaus, A refinement of two theorems of Kronecker, Mich. Math. J. 12 (1965), pp. 81-85.

[14] C.L. Siegel, Algebraic integers whose conjugates lie in the unit circle, Duke Math. J. 11 (1944), pp. 597-602.

[15] C.J. Smyth, On the product of the conjugates outside the unit circle of an algebraic integer, Bull. London Math. Soc. 3 (1971), pp. 169-175.

[16] C.J. Smyth, Some inequalities for certain power sums, Acta Math. Acad. Sci. Hung., 28 (1976), pp. 271-273.

[17] C.L. Stewart, Algebraic integers whose conjugates lie near the unit circle, Bull. Soc. Math. France 106, 1978, (to appear).

Department of Pure Mathematics
University of Waterloo
Waterloo, Ontario
Canada
N2L 3G1