BULLETIN OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 81, Number 6, November 1975

# THE ANALYTIC THEORY OF ALGEBRAIC NUMBERS

BY H. M. STARK

1. **The basics of algebraic number theory.** An algebraic number field is a field $K=\boldsymbol{Q}(\alpha)$ where $\alpha$ is a zero of an irreducible (over $\boldsymbol{Q}$) polynomial $f(x)$ with integral coefficients. The degree of $K$, which we denote by $n=n(K)=[K:\boldsymbol{Q}]$, is the degree of $f(x)$. We write the roots of $f(x)=0$ as $\alpha^{(1)}$, $\alpha^{(2)}, \cdots, \alpha^{(n)}$ in such a way that for $1 \leqq j \leqq r_1=r_1(K)$, $\alpha^{(j)}$ is real, while for $j>r_1$, $\alpha^{(j)}$ is complex. If we let $n=r_1+2r_2$, then it is customary to order the $r_2=r_2(K)$ complex conjugate pairs of roots so that for $r_1+1 \leqq j \leqq r_1+r_2$, $\overline{\alpha^{(j)}}=\alpha^{(j+r_2)}$. The $\alpha^{(j)}$'s are called the conjugates of $\alpha$ and the fields $K^{(j)}=\boldsymbol{Q}(\alpha^{(j)})$ are called the conjugate fields of $K$. If $r_2=0$, we say $K$ is totally real and if $r_1=0$, we say $K$ is totally complex.

The integers of $K$ are those elements of $K$ which are zeros of a polynomial with integer coefficients and leading coefficient 1. The integers of $K$ form a ring which we denote by $\mathfrak{o}$. As is well known, factorization of the integers of $K$ into prime integers is not necessarily unique. Various equivalent ways of remedying this have been used; we follow Dedekind's method. If $\alpha_1, \cdots, \alpha_k$ are elements of $K$, the set

$$\mathfrak{a}=[\alpha_1, \cdots, \alpha_k]=\left\{\sum_{i=1}^{k} a_i\alpha_i \mid a_i \in \boldsymbol{Z}\right\}$$

is called the module generated by $\alpha_1, \cdots, \alpha_k$ (today it would be called a finitely generated $\boldsymbol{Z}$ module). The ring $\mathfrak{o}$ is an example of such a module; on the other hand, $K$ is not an example since it is not finitely generated over $\boldsymbol{Z}$. If $\mathfrak{b}=[\beta_1, \cdots, \beta_m]$ is another module, we define the product $\mathfrak{ab}$ to be the module generated by the $km$ numbers $\alpha_i\beta_j$.

Since 1 is in $\mathfrak{o}$, we always have $\mathfrak{oa} \supset \mathfrak{a}$ for any module $\mathfrak{a}$. If $\mathfrak{oa}=\mathfrak{a}$ then we say $\mathfrak{a}$ is a fractional ideal of $K$. The nonzero fractional ideals of $K$ form an abelian group under multiplication with identity element $\mathfrak{o}$. An integral ideal, or just ideal for short, is a fractional ideal of $K$ which is contained in $\mathfrak{o}$. The integral ideals of $K$ are precisely the ideals of $\mathfrak{o}$ in the sense of ring theory today. Every fractional ideal is a quotient of two integral ideals and factorization of ideals into prime ideals is unique.

Among the fractional ideals of $K$ are the principal fractional ideals. If $\alpha$ is in $K$ then the principal fractional ideal generated by $\alpha$ is

$$(\alpha)=\alpha\mathfrak{o}=[\alpha]\mathfrak{o}.$$

This is a slightly expanded version of an invited address delivered at the Summer 1973 Meeting of the Society in Missoula, Montana.

*AMS (MOS) subject classifications* (1970). Primary 12-02, 12A50, 12A70.

*Key words and phrases.* Algebraic number theory, class-number, discriminant, Dirichlet unit theorem, Dedekind zeta function.

The nonzero principal fractional ideals form a subgroup of the group of nonzero fractional ideals; the quotient group is called the ideal class group of $K$, and the cosets are called the ideal classes of $K$. The order of the ideal class group of $K$ is denoted by $h=h(K)$. For algebraic number fields, $h(K)=1$ if and only if $\mathfrak{o}$ is a unique factorization domain.

A nonzero fractional ideal of $K$ always has $n$ generators and $n$ is minimal. A set of $n$ generators of a fractional ideal $\mathfrak{a}$ is called a basis of $\mathfrak{a}$. If $\mathfrak{a}=[\alpha_1, \cdots, \alpha_n]$ is a fractional ideal, set

$$D(\mathfrak{a})=\det\begin{pmatrix}\alpha_1^{(1)} & \cdots & \alpha_n^{(1)}\\ \cdot & & \cdot\\ \cdot & & \cdot\\ \cdot & & \cdot\\ \alpha_1^{(n)} & \cdots & \alpha_n^{(n)}\end{pmatrix}^2.$$

The number $D(\mathfrak{a})$ is rational and is independent of the choice of basis for $\mathfrak{a}$. When $\mathfrak{a}=\mathfrak{o}$, we write just $D=D_K=D(\mathfrak{o})$; $D$ is called the discriminant of $K$. We define the norm $N(\mathfrak{a})$ of a fractional ideal $\mathfrak{a}$ by

$$D(\mathfrak{a})=D\cdot N(\mathfrak{a})^2,$$

with $N(\mathfrak{a})>0$. The norm of a fractional ideal is rational and indeed, the norm of an integral ideal $\mathfrak{a}$ is the order of the finite ring $\mathfrak{o}/\mathfrak{a}$ and is thus an integer. Norms are multiplicative so that $N(\mathfrak{a}\mathfrak{b})=N(\mathfrak{a})N(\mathfrak{b})$. Elements of $K$ also have norms. If $\alpha$ is in $K$ we define

$$N(\alpha)=\prod_{j=1}^{n}\alpha^{(j)}.$$

The norm of $\alpha$ is connected to the norm of the principal ideal generated by $\alpha$ by $N((\alpha))=|N(\alpha)|$. Every prime ideal $\mathfrak{p}$ divides a unique principal ideal of the form $(p)$ where $p$ is a rational prime. If we write $(p)$ as

$$(p)=\prod\mathfrak{p}_i^{e_i}, \quad \text{then} \quad p^n=N((p))=\prod N(\mathfrak{p}_i)^{e_i}$$

and so there is a positive integer $f_i$ such that $N(\mathfrak{p}_i)=p^{f_i}$. It also follows that $(p)$ is divisible by at most $n$ prime ideals. If any $e_i>1$, we say that $\mathfrak{p}_i$ is a ramified prime in $K$ and $p$ ramifies in $K$. It turns out that a necessary and sufficient condition that a prime $p$ ramify in $K$ is that $p \mid D$. A refinement of this involves an ideal $\mathfrak{d}$ of $K$ called the different of $K$. It has the two properties that $N(\mathfrak{d})=|D|$ and that a necessary and sufficient condition that a prime ideal $\mathfrak{p}$ is ramified in $K$ is that $\mathfrak{p}\mid\mathfrak{d}$.

We are now in position to define the zeta function of $K$. For complex $s$, we define

$$(1) \qquad \zeta_K(s)=\sum_{\mathfrak{a}} N(\mathfrak{a})^{-s}$$

where the summation is over all nonzero integral ideals of $K$. The Riemann zeta function is given by the special case, $\zeta(s)=\zeta_Q(s)$. We have, at least formally, an analytic expression of the unique factorization of ideals,

$$(2) \qquad \zeta_K(s)=\prod_{\mathfrak{p}}(1+N(\mathfrak{p})^{-s}+N(\mathfrak{p}^2)^{-s}+\cdots)=\prod_{\mathfrak{p}}(1-N(\mathfrak{p})^{-s})^{-1},$$

where the product is over all prime ideals of $K$. But now, for $\operatorname{Re} s>1$, we have

$$\sum_{\mathfrak{p}} |N(\mathfrak{p})^{-s}| \leqq \sum_{p} n\, |p^{-s}|$$

which is convergent. Thus (2) converges absolutely for $\operatorname{Re} s>1$ and, as a result, (1) converges absolutely for $\operatorname{Re} s>1$. This defines $\zeta_K(s)$ as an analytic function of $s$ for $\operatorname{Re} s>1$ which has no zeros in this region since none of the factors of the product in (2) have any zeros in this region. As we will soon see, the simple fact that the series for $\zeta_K(s)$ converges absolutely for $\operatorname{Re} s>1$ has some surprising consequences.

2. **Some initial applications of the zeta function.** In an algebraic number theory course at this point, one would usually prove Dirichlet's unit theorem, the finiteness of the class-number and Minkowski's discriminant theorem. We sketch here analytic proofs of these results. None of these proofs have appeared in print before although it seems likely that Hecke knew of the first two. It was Hecke who first extended $\zeta_K(s)$ to the entire $s$-plane. We begin by sketching his method.

If $\mathfrak{C}$ is an ideal class of $K$, we let

$$\zeta(s, \mathfrak{C}) = \sum_{\mathfrak{a} \in \mathfrak{C}} N(\mathfrak{a})^{-s}$$

where the summation is over all integral ideals of $\mathfrak{C}$ and the series converges absolutely for $\operatorname{Re} s>1$ by comparison with (1). Let $\mathfrak{b}$ be a fixed fractional ideal in $\mathfrak{C}^{-1}$. If $\mathfrak{a}$ is in $\mathfrak{C}$, $\mathfrak{a}\mathfrak{b}=(\alpha)$ is principal, and a necessary and sufficient condition that $\mathfrak{a}$ is integral is that $\alpha$ be an element of $\mathfrak{b}$. Thus

$$\zeta(s, \mathfrak{C}) = (N\mathfrak{b})^s \sum_{\alpha\in\mathfrak{b}}{}' |N(\alpha)|^{-s} \tag{3}$$

where $\sum'$ denotes the fact that only one generator of each nonzero ideal $(\alpha)$ is used in the summation. Since there are only finitely many roots of unity which satisfy irreducible equations over $Q$ of degree less than or equal to $n$, the number, $w$, of roots of unity in $K$ is finite. We now have

$$w\zeta(s, \mathfrak{C}) = (N\mathfrak{b})^s \sum_{\alpha\in\mathfrak{b}}{}'' |N(\alpha)|^{-s}$$

where $\sum''$ denotes the fact that whenever $\alpha \neq 0$ is used in the sum, so are all $w$ roots of unity in $K$ times $\alpha$ used and only these $w$ generators of $(\alpha)$ are used.

A two or three page calculation then shows that with

$$\xi(s, \mathfrak{C}) = (|D|/(2^{2r_2}\pi^n))^{s/2}\Gamma(s/2)^{r_1}\Gamma(s)^{r_2}\zeta(s, \mathfrak{C}),$$

we have

$$w\xi(s, \mathfrak{C}) = \frac{n}{2}\int_0^\infty x^{(sn/2)-1}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \sum_{\alpha\in\mathfrak{b}}{}'' \exp\{-c(\mathfrak{b})xm(\alpha, v)\}\, dv\, dx \tag{4}$$

where $v=(v_1, \cdots, v_r)$ is an $r$ dimensional vector,

$$r = r(K) = r_1(K) + r_2(K) - 1,$$
$$c(\mathfrak{b}) = \pi/[|D|^{1/n}N(\mathfrak{b})^{2/n}],$$

and $m(\cdots)$ denotes the Stark mess function. This notation is used to denote something the reader would usually not care to see. In this case

$$m(\alpha, v)=\sum_{j=1}^{r_1}(\alpha^{(j)})^2e^{v_j}+2\sum_{j=r_1+1}^{r_1+r_2}|\alpha^{(j)}|^2\,e^{v_j},$$

where $v_{r+1}$ is defined by

$$\sum_{j=1}^{r_1}v_j+2\sum_{j=r_1+1}^{r_1+r_2}v_j=0.$$

We have written (4) under the assumption that $r_2>0$. If $r_2=0$, the right side of (4) is off by a factor of 2. In any event, all the integrations and interchanges that are necessary to reach (4) are valid for $\operatorname{Re} s>1$.

If $K$ has no units of infinite order, then the sum $\sum''_{\alpha\in\mathfrak{b}}$ is over all nonzero elements of $\mathfrak{b}$. But in this case, it is easy to show that for $r>0$, the inner $r$-fold integral in (4) would be divergent. Thus when $r>0$, $K$ has units of infinite order. A refinement of this argument produces $r$ multiplicatively independent units $\varepsilon_1,\cdots,\varepsilon_r$ in $K$. At this point, we may alter the right side of (4) by changing the $r$-fold inner integral to an integral over a certain $r$ dimensional parallelepiped $V$ determined by $\varepsilon_1,\cdots,\varepsilon_r$ and extending the range of summation on $\alpha$,

$$w\xi(s,\mathfrak{C})=\frac{n}{2}\int_0^\infty x^{(sn/2)-1}\int\cdots\int_V\sum_{\alpha\in\mathfrak{b}}{}'''\exp\{-c(\mathfrak{b})xm(\alpha,v)\}\,dv\,dx. \tag{5}$$

Here $\sum'''$ denotes the fact that when $\alpha\neq 0$ is used in the sum, so is $\varepsilon\alpha$ whenever $\varepsilon$ is a unit in the group $E$ generated by $\varepsilon_1,\cdots,\varepsilon_r$ and the roots of unity of $K$.

On the other hand, since $V$ is compact, the integral on the right of (5) may be shown to be convergent for $\operatorname{Re} s>1$ when $\sum'''_{\alpha\in\mathfrak{b}}$ is replaced by $\sum_{\alpha\in\mathfrak{b},\alpha\neq 0}$. But this means that $E$ is of finite index in the group of units of $K$ since otherwise this integral could be evaluated so as to involve a sum of the form on the right side of (3) except with each $(\alpha)$ occurring infinitely often and this would contradict the convergence of the integral. With a slight amount of work, it follows that the units $\varepsilon_1,\cdots,\varepsilon_r$ can be chosen so that $E$ is the whole group of units. This is Dirichlet's unit theorem. We assume in the sequel that $\varepsilon_1,\cdots,\varepsilon_r$ are so chosen.

Now in (5) the summation over $\alpha$ is over all $\alpha\neq 0$ in $\mathfrak{b}$. The summation over all $\alpha$ in $\mathfrak{b}$ gives rise to a multidimensional $\theta$-function. By utilizing the transformation formulae of the $\theta$-function, Hecke showed that

$$\begin{aligned}w\xi(s,\mathfrak{C})&=\frac{2^{r_1}R}{s(s-1)}\\&+\frac{n}{2}\int_1^\infty x^{(ns/2)-1}\int\cdots\int_V\sum_{\alpha\in\mathfrak{b},\alpha\neq 0}\exp\{-c(\mathfrak{b})xm(\alpha,v)\}\,dv\,dx\\&+\frac{n}{2}\int_1^\infty x^{[n(1-s)/2]-1}\int\cdots\int_V\sum_{\alpha\in\mathfrak{d}^{-1}\mathfrak{b}^{-1};\alpha\neq 0}\exp\{-c(\mathfrak{d}^{-1}\mathfrak{b}^{-1})xm(\alpha,v)\}\,dv\,dx,\end{aligned}\tag{6}$$

where the volume of $V$ is $2^{r_1}R$ and $R=R(K)$ is the regulator of $K$. One may express $R$ as an $r$ by $r$ determinant involving the units $\varepsilon_1, \cdots, \varepsilon_r$ and their conjugates. (6) is valid initially only for $\operatorname{Re} s>1$, but the integrals on the right converge for all $s$ and thus give an analytic continuation of $\xi(s, \mathfrak{C})$ to the entire $s$ plane. Again if $r_2=0$, the integrals on the right of (6) are off by a factor of 2. If $s$ is real, then everything in the integrals is positive, and so for $s$ real, $s>1$, we have

$$w\xi(s, \mathfrak{C}) > 2^{r_1}R/[s(s-1)], \tag{7}$$

which is a positive lower bound independent of $\mathfrak{C}$. But the sum over $\mathfrak{C}$ of the left side of (7) exists and so there are only finitely many $\mathfrak{C}$, i.e., the class-number of $K$ is finite.

The right-hand side of (6) is invariant when the pair $\mathfrak{b}$, $s$ is replaced by the pair $\mathfrak{d}^{-1}\mathfrak{b}^{-1}$, $1-s$. With

$$\xi(s)=(|D|/(2^{2r_2}\pi^n))^{s/2}\Gamma(s/2)^{r_1}\Gamma(s)^{r_2}\zeta_K(s),$$

this gives us the functional equation for $\zeta_K(s)$,

$$\xi(s)=\xi(1-s).$$

If we analyze the growth of $\xi(s)$ from (6), we find that the function $f(s)=s(s-1)\xi(s)$ is an entire function of order 1. The Hadamard product theorem then says that

$$f(s)=e^{A+Bs}\prod_{\rho}\{1-(s/\rho)\}e^{s/\rho} \tag{8}$$

where $\rho=\beta+i\gamma$ runs through the zeros of $f(s)$. The numbers $\rho$ are called the nontrivial zeros of $\zeta_K(s)$; since $\zeta_K(s)$ has no zeros with $\operatorname{Re} s>1$, it follows that $\beta$ is between 0 and 1.

If we logarithmically differentiate (8) and use the equation $f(s)=f(1-s)$, we can get rid of both $A$ and $B$,

$$\frac{f'(s)}{f(s)}=\sum_{\rho}\frac{1}{s-\rho}.$$

To ensure convergence, the $\rho$ and $\bar{\rho}$ terms must be grouped together; with this grouping we have

$$\frac{f'(s)}{f(s)}=\sum_{\rho}\frac{s-\beta}{(s-\beta)^2+\gamma^2}, \tag{9}$$

and this is valid for all $s$. On the other hand, the definition of $f(s)$ leads to

$$\frac{f'(s)}{f(s)}=\frac{1}{s}+\frac{1}{s-1}+\frac{1}{2}\log\left(\frac{|D|}{2^{2r_2}\pi^n}\right)+\frac{r_1}{2}\frac{\Gamma'(s/2)}{\Gamma(s/2)}+r_2\frac{\Gamma'(s)}{\Gamma(s)}+\frac{\zeta_K'(s)}{\zeta_K(s)}. \tag{10}$$

If $s$ is real, $s>1$, then from (9), $f'(s)/f(s)>0$ and from (2),

$$\frac{\zeta_K'(s)}{\zeta_K(s)}=-\sum_{\mathfrak{p}}\frac{\log(N\mathfrak{p})}{(N\mathfrak{p})^s-1}<0.$$

Therefore (10) implies that

$$(11)\quad \frac{1}{n}\log D > \frac{r_1}{n}\left[\log\pi - \frac{\Gamma'(s/2)}{\Gamma(s/2)}\right] + \frac{2r_2}{n}\left[\log(2\pi) - \frac{\Gamma'(s)}{\Gamma(s)}\right] - \frac{2}{n}\left(\frac{1}{s} + \frac{1}{s-1}\right).$$

With $s=3$, we get

$$\frac{1}{n}\log D > \frac{r_1}{n}(1.1) + \frac{2r_2}{n}(0.91) - \frac{2}{n}\cdot\frac{5}{6}$$
$$\geqq 0.91 - 5/(3n).$$

Hence for $n \geqq 2$, we have

$$n^{-1}\log D > 0.07,$$

and so $D > 1$ and, indeed,

$$(12)\qquad D > (1.07)^n.$$

If $n$ is large, we may take $s$ closer to 1 in (11) and improve our estimate in (12). The ultimate result may be expressed as

$$\frac{1}{n}\log D \geqq \frac{r_1}{n}\left[\log\pi - \frac{\Gamma'(\frac{1}{2})}{\Gamma(\frac{1}{2})}\right] + \frac{2r_2}{n}\left[\log(2\pi) - \frac{\Gamma'(1)}{\Gamma(1)}\right] + o(1)$$
$$\geqq r_1 n^{-1}(3.108) + 2r_2 n^{-1}(2.415) + o(1)$$
$$\geqq 2.415 + o(1)$$

as $n \to \infty$. Thus for large degrees,

$$(13)\qquad D^{1/n} > (22)^{r_1/n}(11)^{2r_2/n}.$$

This is in fact better than Minkowski's original result, which numerically was

$$D^{1/n} > (7.389)^{r_1/n}(5.803)^{2r_2/n}$$

when $n$ is large.

3. **The Brauer-Siegel theorem.** There are many standard applications of zeta functions to algebraic number theory that must be skipped here for lack of time. The most obvious of these is the prime ideal theorem which says that the number of prime ideals of a field $K$ with norm $\leqq x$ is asymptotic to $x/\log x$ as $x \to \infty$. This theorem was covered in my Las Vegas address [**4**]. I would also like to mention the topic of values of $L$-series at $s=1$; the reader is referred to [**6**] for further details on this topic.

There is a Riemann hypothesis for $\zeta_K(s)$. It says that the only zeros of $\zeta_K(s)$ in the region $\mathrm{Re}\, s > 0$ lie on the line $\mathrm{Re}\, s = \frac{1}{2}$. As with the ordinary Riemann hypothesis (the case $K = \mathbf{Q}$), no progress has been made. A very important special case is the following

CONJECTURE. *If $s$ is real, $s > 1 - (4\log|D|)^{-1}$, then $\zeta_K(s) \neq 0$.*

Even this is not known to be true for all $K$ but it is known that the conjecture can only be wrong once for each $K$.

LEMMA. *There is at most one real zero of* $\zeta_K(s)$ *between* $1-(4\log|D|)^{-1}$ *and* 1; *if it exists, it is a simple zero of* $\zeta_K(s)$.

The proof of the Lemma may be based on (9) and (10). If we take $s$ real and slightly greater than 1, then two zeros of $\zeta_K(s)$ very near 1 make $f'(s)/f(s)$ in (9) larger than the $1/(s-1)$ term on the right side of (10) allows. (The $\Gamma$ and $\zeta$ terms are negative.) If $\zeta_K(s_0)=0$ for some real $s_0$ between $1-(4\log|D|)^{-1}$ and 1, we will call $s_0$ *the exceptional zero* of $\zeta_K(s)$.

Suppose for the moment that the conjecture is true for a particular $K$; i.e., $\zeta_K(s)$ does not have an exceptional zero. In this case, thanks to the first order pole of $\zeta_K(s)$ at $s=1$, $\zeta_K(s)$ is less than or equal to zero at $s_0=1-(4\log|D|)^{-1}$. We now sum (6) over the $h$ different classes $\mathfrak{C}$ and set $s=s_0$. On the left side we find $w\xi(s_0)$ which is less than or equal to zero; on the right we find $2^r hR/s_0(s_0-1)$ which is negative while the integral terms are clearly positive. Indeed the integral terms may be shown to be moderately large. This means that $hR$ cannot be too small and in fact one obtains an estimate of the following form,

$$hR > c_1^{-n}s_0(1-s_0)\,|D|^{s_0/2} > c_2^{-n}(1-s_0)\,|D|^{1/2} \tag{14}$$

where $c_1$ and $c_2$ are large constants (effectively computable). This is an excellent lower estimate of $hR$ but it depends on the conjecture. If the conjecture is false, the best we can do is let $s_0$ be the exceptional zero of $\zeta_K(s)$. We then get (14) again but the result is much poorer if $s_0$ is very close to 1; indeed the result is useless if $s_0$ is within $|D|^{-1/2}$ of 1.

The Brauer-Siegel theorem gets around this, but in an ineffective way. The main idea of the proof is to make the exceptional zeros of two different zeta functions contradict each other by showing that they are both zeros of the zeta function of the composite field. When successful, this argument yields a result of the form

$$hR > c_3(\varepsilon)^{-1}\,|D|^{(1/2)-\varepsilon} \tag{15}$$

where $c_3(\varepsilon)$ is a large number depending upon $\varepsilon>0$ and is ineffective since it depends upon a hypothetical counterexample to the conjecture. Before now, no proof of (15) was known for which $c_3(\varepsilon)$ could be explicitly given for any $\varepsilon<\frac{1}{2}$.

With today's knowledge (about Artin $L$-series, in particular), the argument leading to (15) can be carried out for two types of fields. First, we may derive (15) for all fields $K$ of a fixed degree $n$, in which case $c_3(\varepsilon)$ depends upon $n$ also. Second (this is the more frequently quoted case), we may derive (15) for a sequence $\mathcal{K}$ of fields $K$ which are normal over $\mathbf{Q}$ and such that $n(K)^{-1}\log|D_K|\to\infty$ as $K$ runs through $\mathcal{K}$. Because of the first case, we may restrict ourselves in the discussion of the second case to sequences $\mathcal{K}$ such that $n(K)\to\infty$ as $K$ runs through $\mathcal{K}$. In this case, $c_3(\varepsilon)$ depends upon $\mathcal{K}$ also. In either case, (15) represents the interesting half of the Brauer-Siegel theorem.

The dull half of the Brauer-Siegel theorem says that

$$hR < c_4(\varepsilon)\,|D|^{(1/2)+\varepsilon} \tag{16}$$

where $c_4(\varepsilon)$ is effectively computable for each $\varepsilon > 0$. In the first case, $c_4(\varepsilon)$ depends upon $n$ also, and in the second case, $c_4(\varepsilon)$ depends upon $\mathcal{K}$ also. However, because of (16), we may recast (15) in a more picturesque form and this is the way the result is usually phrased.

BRAUER-SIEGEL THEOREM (FIRST VERSION). *If K runs through a sequence of fields of fixed degree, then*

$$\log[h(K)R(K)] \sim \tfrac{1}{2}\log|D_K|. \tag{17}$$

BRAUER-SIEGEL THEOREM (SECOND VERSION). *If K runs through a sequence of normal extensions of* $\boldsymbol{Q}$ *such that* $n(K) \to \infty$, $n(K)^{-1}\log|D_K| \to \infty$, *then* (17) *holds.*

It is to be emphasized that up to now, both versions of the Brauer-Siegel theorem have been completely ineffective. To make them effective, it is necessary to specify $c_3(\varepsilon)$ in (15) for each $\varepsilon$, $0 < \varepsilon < \frac{1}{2}$.

Perhaps the most important application of the Brauer-Siegel theorem is the case that $K$ is totally complex, $k$ is totally real, $[K:k]=2$. Then the regulators of $K$ and $k$ are essentially the same (the unit theorem says that the units of $k$ form a subgroup of finite index in $K$ and this index may be easily estimated). Thus

$$h(K)/h(k) > c_5(\varepsilon)^{-1}\,|D_K|^{(1/2)-\varepsilon}/|D_k|^{(1/2)+\varepsilon}.$$

Now $|D_K| = D_k^2 f$, where $f$ is an integer, and so

$$h(K) > c_5(\varepsilon)^{-1}\,|D_k|^{(1/2)-3\varepsilon} f^{(1/2)-\varepsilon}. \tag{18}$$

Here $c_5(\varepsilon)$ depends also upon the degree of $k$ and is ineffective. If $k$ is fixed and $K$ runs through all totally complex quadratic extensions of $k$, then $|D_K| \to \infty$ and so $f \to \infty$. Therefore *as K runs through all totally complex quadratic extensions of a fixed totally real field k,* $h(K) \to \infty$. This result depends upon the first version of the Brauer-Siegel theorem and is ineffective.

It was Heilbronn's proof of this last result for $k = \boldsymbol{Q}$ that motivated Siegel's part of the Brauer-Siegel theorem. Siegel's theorem is the first version of the Brauer-Siegel theorem for $n(K)=2$ and is also ineffective. However, it now turns out that the case of fields of degree 2 is where the entire difficulty lies in the Brauer-Siegel theorem. Every other case of the Brauer-Siegel theorem can be made somewhat effective, and in some instances the whole result can be made effective.

The key to this new development is most easily found if we assume an unproved conjecture of Artin. If $M$ is a normal extension of $\boldsymbol{Q}$ then

$$\zeta_M(s) = \zeta(s)\prod_{\chi} L(s,\chi)^{m(\chi)}$$

where $\chi$ runs through the nontrivial irreducible characters of the Galois group $G = G(M/\boldsymbol{Q})$ of $M$ and $m(\chi)$ is the degree of $\chi$ which is a positive integer. The functions $L(s,\chi)$ are Artin $L$-series and are conjectured to be entire functions of $s$. Assuming this to be true, if $s_0$ is a real simple zero of

$\zeta_M(s)$ between 0 and 1, then $\zeta(s_0)\neq 0$, and so $L(s_0, \chi)=0$ for exactly one $\chi$, say $\chi=\chi_2$. The character $\chi_2$ must be real ($L(s_0, \bar{\chi}_2)=0$ also) and the corresponding $m(\chi_2)=1$.

The theory of Artin $L$-series says that corresponding to such a character there is a quadratic field $F$ contained in $M$ such that $\zeta_F(s)=\zeta(s)L(s, \chi_2)$ and so $s_0$ is a zero of $\zeta_F(s)$ also. Furthermore, if $K$ is any subfield of $M$, $\zeta_K(s)$ has a product decomposition into $L$-series of the form

$$\zeta_K(s)=\zeta(s)\prod_{\chi} L(s, \chi)^{m(K,\chi)}$$

where $m(K, \chi)\geqq 0$. One finds that $m(K, \chi_2)>0$ if and only if $F\subset K$. Thus $\zeta_K(s_0)=0$ if and only if $F\subset K$. The upshot is that *if $s_0$ is a real simple zero of $\zeta_M(s)$ between 0 and 1 and if*

(i) *each $L(s, \chi)$ is analytic at $s_0$,*

*then*

(ii) *$M$ has a quadratic subfield $F$ such that $\zeta_F(s_0)=0$, and*

(iii) *if $K$ is a subfield of $M$ then $\zeta_K(s_0)=0$ if and only if $F\subset K$.*

Recently Heilbronn [**1**] proved (ii) without having to assume (i). His proof does not yield the important second part of the result above. However, I have found another version of his result that does enable one to prove (iii) also. The key step is, in fact, a proof [**5**] (via group representation theory) that each $L(s, \chi)$ is analytic at any simple zero of $\zeta_M(s)$. In particular, (i) is true and (ii) and (iii) follow.

This result has startling consequences for the Brauer-Siegel theorem. Let $K$ be a field and $M$ the splitting field of $K$ over $\boldsymbol{Q}$. We may estimate $D_M$ and find that for $n=n(K)$,

$$|D_M|<|D_K|^{n!}. \tag{19}$$

Let $g(K)=1$ if $K$ is normal over $\boldsymbol{Q}$ ($M=K$) and $g(K)=n!$ otherwise. Suppose $s_0$ is a real zero of $\zeta_K(s)$ between $1-(4g(K)\log|D_K|)^{-1}$ and 1. A result of Aramata (rediscovered by Brauer) states that $\zeta_M(s)/\zeta_K(s)$ is entire and so $s_0$ is a zero of $\zeta_M(s)$ also. In fact $s_0$ is in the range $1-(4\log|D_M|)^{-1}\leqq s_0<1$ and so $s_0$ is a simple zero of $\zeta_M(s)$. Therefore there is a quadratic subfield $F$ of $K$ such that $\zeta_F(s_0)=0$. As a result, $s_0$ does not exist if $K$ has no quadratic subfields. In particular, the Brauer-Siegel theorem is completely effective for fields of odd degree.

But even if $K$ does have a quadratic subfield, we may say something effective about the Brauer-Siegel theorem. If $F$ is a quadratic subfield of $K$ and $\zeta_K(s_0)=0$ where $s_0$ is real, then

$$s_0<1-(c_6^{-1}/|D_F|^{1/2})$$

where $c_6$ is a large constant that is effectively computable (actually $\pi/6$ will probably work). (This is the $\varepsilon=\frac{1}{2}$ case of Siegel's theorem which is effective; however this case gives no useful information about $F$.) But

$$|D_K|=|D_F|^{n/2}f_1$$

for some positive integer $f_1$, where we are still using $n=n(K)$. Therefore

$$|D_F|^{1/2} \leqq |D_K|^{1/n}$$

and hence

$$s_0 < 1 - (c_6^{-1}/|D_K|^{1/n}).$$

In particular,

$$\begin{aligned} 1 - s_0 &> \min(1/(4g(K)\log|D_K|),\ (c_6^{-1}/|D_K|^{1/n})) \\ &> (ng(K)\,|D_K|^{1/n})^{-1} \cdot \min(e/4,\ c_6^{-1}). \end{aligned}$$

Thus, from (14),

$$h(K)R(K) > (c_7^{-1}c_2^{-n}/ng(K))|D|^{(1/2)-(1/n)} \tag{20}$$

where $c_7$ is effectively computable. This shows that the second version of the Brauer-Siegel theorem is effective also. In fact our methods lead to the following result.

THEOREM 1. *Both versions of the Brauer-Siegel theorem are effective for fields of odd degree. The second version of the Brauer-Siegel theorem is effective. The first version of the Brauer-Siegel is partially effective for fields of degree $n(K)=n>2$ in that we may specify $c_3(\varepsilon)$ in (15) for $\varepsilon \geqq n^{-1}$. In any event, the first version of the Brauer-Siegel theorem is a corollary of Siegel's theorem.*

Because we have made the first version of the Brauer-Siegel theorem partially effective, we can now derive an effective class-number result. It has the general appearance of (18) except that the estimate is made somewhat better by the use of $L$-series [**5**].

THEOREM 2. *Let $K$ be a totally complex field of degree $2n$ containing a totally real field $k$ of degree $n$. We let $f$ denote the positive integer such that $|D_K|=|D_k|^2 f$ and we let $g(k)=1$ if $k/\mathbf{Q}$ is normal, $g(k)=n!$ otherwise. For $\varepsilon$ in the range $0<\varepsilon \leqq \frac{1}{2}$, and for sufficiently large $c_8(\varepsilon)$ (where $c_8(\varepsilon)$ is effectively computable and independent of $n$),*

$$h(K) > (ng(k)c_8(\varepsilon)^n)^{-1}\,|D_K|^{(1/2)-(1/n)-\varepsilon} f^{(1/2)-1/(2n)}. \tag{21}$$

As a corollary, we obtain the following result.

THEOREM 3. *If $k$ is a fixed totally real field other than $\mathbf{Q}$, then when $K$ runs through all complex quadratic extensions of $k$, $h(K)\to\infty$ effectively. If $n$ is fixed, $n>2$, then given $h$, there are only finitely many totally real fields $k$ of degree $n$ which have any totally complex quadratic extensions of class number $h$ and these $k$ may be effectively determined.*

Indeed, since $c_8(\varepsilon)$ in Theorem 2 is independent of $n$, it is possible to replace the words "totally complex quadratic" in the last theorem by "totally complex abelian".

4. **Postscript.** It is now almost two years since the talk on which this paper is based was given. There have been some significant advances on two

of the topics covered in this paper. With reference to the estimate (13) of discriminants, the best that the geometry of numbers has produced is the estimate of Mulholland [**2**],

$$|D|^{1/n} \geqq (32.561)^{r_1/n}(15.775)^{2r_2/n} + o(1) \quad \text{as } n \to \infty, \tag{22}$$

which is better than (13). Odlyzko has recently refined (11) so as to produce estimates better than this. His best estimates are complicated but he has produced a very simple refinement [**3**] of (11),

$$|D|^{1/n} \geqq (50)^{r_1/n}(19)^{2r_2/n} + o(1) \quad \text{as } n \to \infty$$

which is already better than (22).

The analytic methods we used in [**5**] to get (21) already suffice to produce a reasonable lower estimate for the associated $L$-function which is independent of the degree of $k$. If $k$ is totally real of degree $n$, $K$ is a totally complex quadratic extension of $k$, $|D_K| = D_k^2 f$, then

$$\zeta_K(s) = \zeta_k(s)L(s, \chi)$$

where $\chi$ is a real character of $k$ whose conductor has norm $f$ and the associated $L$-series, $L(s, \chi)$, is entire. From Lemma 5 of [**5**] and p. 150, lines 4–7 of [**5**], we see that for $s$ real, $3/2 \leqq s \leqq 2$,

$$L(1, \chi) > [c_9 n g(k)]^{-1} m(s)^n \cdot [D_k^{(1/2)(s-1)+1/n} f^{1/(2n)}] \tag{23}$$

where $c_9$ is effectively computable and

$$m(s) = \pi^{s/2}/[\Gamma(s/2)\zeta(s)].$$

Since $\pi > \pi^2/6 = \zeta(2)$, $m(s) > 1$ for $s$ near 2. For example $m(1.66) > 1$. If we take $s = 1.66$ in (23) and assume that $k/\boldsymbol{Q}$ is normal so that $g(k) = 1$, then for $n > n_0$, where $n_0$ is effectively computable,

$$L(1, \chi) > 1/(D_k f)^{1/3}. \tag{24}$$

An estimate of this form has been previously derived only ineffectively, and even then some sort of extra hypothesis has been needed to eliminate from consideration those fields with small discriminants. Class-number results follow from (24) because $h(K)$ is essentially $(D_k f)^{1/2} L(1, \chi)$. However, the correct estimate is actually [**5**, equation (31)],

$$h(K) \geqq (h(k)/(2\pi)^n)(D_k f)^{1/2} L(1, \chi)$$

and for variable $n$, the factor $(2\pi)^{-n}$ is deadly. With better choices of $s$, a fair amount of the $(2\pi)^{-n}$ can be absorbed, but not all of it. Odlyzko's methods, however, give enough of an improvement on (23) so that the restriction in Theorem 3 that $n$ be fixed can be removed for $k$ normal over $\boldsymbol{Q}$. He proves [**3**] that given $h$, there is an effectively computable $n_0(h)$ such that if $k$ is a totally real field of degree $n$ and $K$ is a totally complex quadratic extension of $k$ (or even a totally complex abelian extension of $k$) with $h(K) = h$, then $n \leqq n_0(h)$.

## References

1. H. Heilbronn, *On real zeros of Dedekind $\zeta$-functions*, Canad. J. Math. **25** (1973), 870–873. MR **48** #6061.

2. H. P. Mulholland, *On the product of n complex homogeneous linear forms*, J. London Math. Soc. **35** (1960), 241–250. MR **22** #4703.

3. A. O. Odlyzko, *Some analytic estimates of class numbers and discriminants*, Invent. Math. (to appear).

4. H. M. Stark, *Some applications of analysis to number theory*, Address to the M.A.A. (Las Vegas, 1972), Amer. Math. Monthly (to be submitted).

5. ———, *Some effective cases of the Brauer-Siegel theorem*, Invent. Math. **23** (1974), 135–152.

6. ———, *Values of L-functions at $s=1$*. II, *Artin L-functions with rational characters*, Advances in Math. (to appear).