A generalisation of Artin's conjecture for primitive roots

by

K. R. Matthews (Brisbane, Australia)

I. Introduction. In 1927 Artin (see Artin [1], p. vii–ix) conjectured that if \( a \) is an integer other than \(-1\) or a perfect square, then the number \( N_a(x) \) of primes \( p \leq x \) such that \( a \) is a primitive root mod \( p \), satisfies an asymptotic formula of the form

\[
N_a(x) \sim A(a) \frac{x}{\log x} \quad (x \to \infty)
\]

for some positive constant \( A(a) \).

In 1967 Hooley (see [6]) showed that the truth of the Riemann hypothesis for the fields \( K \subset K \subset K \), \( K \) square-free, would imply the truth of Artin's conjecture.

Let \( a_1, \ldots, a_n \) be non-zero integers not \( \pm 1 \). In this paper the method of Hooley is used to obtain an asymptotic formula for the number \( N_{a_1,\ldots,a_n}(x) \) of primes \( p \leq x \) such that each of \( a_1, \ldots, a_n \) is a primitive root mod \( p \).

An equation

\[
N_{a_1,\ldots,a_n}(x) = \frac{x}{\log x} A(a_1, \ldots, a_n) + O \left( \frac{x}{\log^2 x (\log \log x)^{n-1}} \right)
\]

is obtained subject to the truth of the Riemann hypothesis for each of the fields \( K \subset K \subset K \) where \( K = \langle l_1, \ldots, l_n \rangle \) (the i.c.m. of \( l_1, \ldots, l_n \)) is square-free. Here

\[
A(a_1, \ldots, a_n) = \sum_{\mu(k)} \mu(k) \phi(k)
\]

where \( \phi(k) \) is the natural density of the primes \( q \equiv 1 \pmod{k}, q \nmid a_1 \ldots a_n \), such that for each prime \( p \mid k \), at least one of \( a_1, \ldots, a_n \) is a \( p \)-th power residue mod \( q \).
The constant $A(a_1, \ldots, a_n)$ is then converted to an infinite product, namely

\begin{equation}
A(a_1, \ldots, a_n) = \prod_{p>2} \left(1 - \frac{1 - \sigma(p)}{p}ight) \sum_{a_1 \equiv a_1 \equiv a_2 \equiv a_3 \equiv 0 \pmod{4}}^1 \sum_{a_4 \equiv a_5 \equiv \cdots \equiv a_n \equiv 0 \pmod{4}}^1 (-1)^{\sum_a} f(|a|)
\end{equation}

where

\begin{equation}
f(|a|) = \mu(|a|) \prod_{p|a} \frac{\sigma(p)}{1 - \sigma(p)}.
\end{equation}

(Here $a(b)$ is the square-free kernel of $b$.)

Denoting the finite sum by $S$, the positivity of $A(a_1, \ldots, a_n)$ is equivalent to that of $S$. A necessary and sufficient condition for $S$ to be positive is then obtained, namely the conjunction of the following two conditions:

\textbf{C}_1: \quad \text{if } a_1^6 \cdots a_n^6 = b^6, \ b \in \mathbb{Z}, \ \varepsilon_i = 0 \text{ or } 1, \text{ then } 2 \sum \varepsilon_i.

\textbf{C}_2: \quad \text{if } a_1^6 \cdots a_n^6 = -3b^6, \ b \in \mathbb{Z}, \ \varepsilon_i = 0 \text{ or } 1 \text{ and if } 2 \sum \varepsilon_i, \text{ then } \delta'(3), \ \text{the natural density of the primes } q \equiv 1 \pmod{3}, \ g \nmid a_1, \ldots, a_n, \text{ such that each of } a_1, \ldots, a_n, \text{ is a cubic non-residue mod } q, \text{ must be positive.}

An explicit formula is available for $\delta'(p) (\equiv \frac{1}{p-1} - \sigma(p))$ where $p$ is an odd prime. By applying the exclusion principle to Lemma 1 of Schinzel [10], p. 162, we have

\begin{equation}
\delta'(p) = \frac{1}{p^a(p-1)} \sum_{j=0}^{n-1} (-1)^j p^{n-1} \sigma_j
\end{equation}

where

\begin{equation}
\sigma_j = \sum_{1 \leq \delta_1 \leq \cdots \leq \delta_j \leq n} \delta_1 \delta_2 \cdots \delta_j
\end{equation}

and $\sigma_0 = 1$.

If $n = 1, 2$ or 3, or if $a_1$, $a_2$, $a_3$ are relatively prime in pairs, an examination of formula (1.6) allows us to replace the condition “$\delta'(3) > 0$” in $C_1$ by the statement “none of $a_1$, $a_2$, $a_3$ is a perfect cube.” However if $n > 3$ the situation is more complicated.

If $\mathcal{P}$ is the set of primes $p$ such that each of $a_1, \ldots, a_n$ is a primitive root mod $p$, we shall show in the next section that conditions $C_1$ and $C_2$ are each necessary for $\mathcal{P}$ to be infinite.

Finally I would like to acknowledge my indebtedness to Professors H. Halberstam and C. Hooley for suggesting the problem to me. I am also extremely grateful to Professor D. A. Burgess for much help while the main part of this work was carried out during a recent sabbatical year spent at the University of Nottingham where I was on leave from the University of Queensland. This paper forms part of a Ph. D. thesis to be submitted to the latter university. Finally, I wish to thank Professor C. S. Davis for improving the presentation of the manuscript.

2. Necessary conditions for $\mathcal{P}$ to be infinite. We prove that the falsity of $C_1$ or $C_2$ implies that $\mathcal{P}$ contains at most the element 2. For suppose that $C_1$ is false. Then we have

\begin{equation}
a_{i_1} \cdots a_{i_j} = b_{i_j}, \ b \in \mathbb{Z}, \ 1 \leq i_1 < \cdots < i_j \leq n
\end{equation}

and $j$ odd. Now if $p \in \mathcal{P}$ is an odd prime, then each of $a_1, \ldots, a_n$ is a quadratic non-residue mod $p$, and the Legendre symbol gives the following contradiction:

\begin{equation}
1 = \left(\frac{b_{i_j}^2}{p}\right) = \left(\frac{a_{i_1} \cdots a_{i_j}}{p}\right) = \left(\frac{a_{i_1}}{p}\right) \cdots \left(\frac{a_{i_j}}{p}\right) = (-1)^j = -1.
\end{equation}

Hence the only possible element of $\mathcal{P}$ is 2.

Now suppose that $C_2$ is false. Then we have

\begin{equation}
a_{i_1} \cdots a_{i_j} = -3b_{i_j}, \ b \in \mathbb{Z}, \ 1 \leq i_1 < \cdots < i_j \leq n
\end{equation}

with $j$ even and $\delta'(3) = 0$. If $p \in \mathcal{P}$ is an odd prime, an argument similar to the above gives $\left(\frac{-3}{p}\right) = 1$. Hence

\begin{equation}
\left(\frac{3}{p}\right) = \left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}}.
\end{equation}

But quadratic reciprocity gives

\begin{equation}
\left(\frac{3}{p}\right) = \left(\frac{p^{3}}{3}\right) = (-1)^{\frac{p^2-1}{2}}
\end{equation}

and consequently $\left(\frac{p^3}{3}\right) = 1$. Hence $p = 1 \pmod{3}$. Now as each of $a_1, \ldots, a_n$ is a primitive root mod $p$, in view of the congruence $p = 1 \pmod{3}$ just derived, each of $a_1, \ldots, a_n$ is a cubic non-residue mod $p$. Hence the condition $\delta'(3) = 0$ implies that $\mathcal{P}$ has density zero. However by Lemma 2.4 below, much more is true. For by that lemma we know that at least one of $a_1, \ldots, a_n$ is a cubic residue mod $p$, and hence there are no odd primes $p \in \mathcal{P}$.

It remains to prove Lemma 2.4. Our proof depends on a theorem of Elliott [4] (Theorem 1, p. 143) on the number of prime ideals $p$ which
have prescribed \( p \)-th power residue symbol values at each of \( n \) given non-zero rational integers.

**Definition 2.1.** Let \( p \) be a rational prime, \( a \) an algebraic integer of \( \mathbb{Q}(\zeta_1) \) and \( p \) a prime ideal of \( \mathbb{Q}(\zeta_1) \), \( p \not| [p,a] \). Then the \( p \)-th power residue symbol is defined by

\[
\left( \frac{a}{p} \right)_p = \alpha^p \left( \frac{a}{p} \right)_p, \quad \left( \frac{a}{p} \right)_p = 1.
\]

(See Landau [8], Definition 135, p. 295.)

We note that

\[
\left( \frac{a}{p} \right)_p = 1 \iff a^p \equiv \beta^p \left( \frac{a}{p} \right)_p, \quad \beta \in \mathbb{Q}(\zeta_1).
\]

We also have the equation

\[
\left( \frac{a_1 a_2}{p} \right)_p = \left( \frac{a_1}{p} \right)_p \left( \frac{a_2}{p} \right)_p
\]

if \( p \not| [p,a_1, a_2] \).

Elliott's result can now be stated as

**Lemma 2.1.** Let \( p \) be a rational prime and let \( a_1, \ldots, a_n \) be non-zero rational integers. Also let \( \varepsilon_1, \ldots, \varepsilon_n \) be \( p \)-th roots of unity. We define \( N(p, n) = N(p, n; \varepsilon_1, \ldots, \varepsilon_n) \) by

\[
N(p, n) = \sum_{i_1=1}^{p-1} \cdots \sum_{i_n=1}^{p-1} \left( \varepsilon_1^{i_1} \cdots \varepsilon_n^{i_n} \right)^{-1}.
\]

(2.4)

Let \( S(a, p, n) = S(a, p, n; \varepsilon_1, \ldots, \varepsilon_n) \) be the number of prime ideals \( p \) of the first degree which satisfy \( N(p) \leq a \), and for which the relations

\[
\left( \frac{a_j}{p} \right)_p = \varepsilon_j \quad (j = 1, \ldots, n)
\]

are satisfied. Then as \( a \to \infty \)

\[
S(a, p, n) = p^{-n} N(p, n) \pi(a) + O \left( \exp( -A \sqrt{\log a} ) \right)
\]

where \( A \) is a positive constant.

(The extra condition, that the prime ideals be of the first degree, is not present in Elliott's theorem, but the contribution of the prime ideals of degree greater than one is \( O(\sqrt{a}) \).)

The next result is not mentioned in Elliott's paper.

**Lemma 2.2.** With \( N(p, n) \) defined as in (2.4) we have

\[
N(p, n) = \sum_{r_1=1}^{p-1} \cdots \sum_{r_n=1}^{p-1} 1
\]

if a prime ideal \( p \) exists satisfying conditions (2.5); otherwise \( N(p, n) = 0 \).

**Proof.** (i) Let \( p \) be a prime ideal satisfying conditions (2.5). Then

\[
\varepsilon_1^{i_1} \cdots \varepsilon_n^{i_n} = \left( \frac{a_1^{i_1} \cdots a_n^{i_n}}{p} \right)_p.
\]

by (2.3). Hence

\[
N(p, n) = \sum_{r_1=1}^{p-1} \cdots \sum_{r_n=1}^{p-1} \left( \frac{a_1^{i_1} \cdots a_n^{i_n}}{p} \right)_p = \sum_{r_1=1}^{p-1} \cdots \sum_{r_n=1}^{p-1} 1,
\]

by (2.3).

(ii) If no prime ideal \( p \) exists satisfying conditions (2.5) then \( S(a, p, n) = 0 \) for each \( a \), and by virtue of (2.6) we must have \( N(p, n) = 0 \).

Before Lemma 2.4 can be proved, we need the following

**Lemma 2.3.** Let \( a \) be a rational integer, \( p \) and \( q \) be rational primes, \( q \equiv 1 (\mod p) \), \( q \not| a \). Also let \( p \) be a prime ideal in \( \mathbb{Q}(\zeta_1) \), \( p \not| [q] \). Then the congruence

\[
a = \beta^p (\mod p), \quad \beta \in \mathbb{Q}(\zeta_1)
\]

is soluble if and only if \( a \) is a \( p \)-th power residue \( \mod q \).

**Proof.** The assumption \( p \not| [q] \) and \( q \equiv 1 (\mod p) \) implies that \( p \) is of the first degree. Hence the integers of \( \mathbb{Q}(\zeta_1) \mod p \) form a field of \( q (= N(p)) \) elements. Hence if \( a = \beta^p (\mod p) \) we have

\[
a^{-1} a^p = \beta^{p-1} \equiv 1 (\mod p),
\]

and as \( q \not| p \), it follows that

\[
a^{-1} a^p = 1 (\mod q).
\]

Consequently \( a \) is a \( p \)-th power residue \( \mod q \).

**Lemma 2.4.** Let \( a_1, \ldots, a_n \) be non-zero rational integers, \( p \) a rational prime and suppose that a prime \( q \) exists with \( q \equiv 1 (\mod p) \), \( q \not| a_1 \cdots a_n \)
and such that each of $a_1, \ldots, a_n$ is a $p$-th power non-residue mod $q$. Then the natural density $d'(p)$ of the set $\mathcal{P}_i$ of such primes $q$, is positive.

Proof. The prime ideals in $\mathcal{O}(\sqrt{1})$ of the first degree and not dividing $[a_1 a_2 \ldots a_n]$ are grouped naturally together in sets of $p-1$ by the equation

$$[a] = p_1 \ldots p_{p-1}$$

where $q = N(p) = 1 mod p$.

By Lemma 2.3 and (2.2) the conditions

$$(2.8) \quad \left( \frac{a_j}{p_i} \right) \neq 1 \quad (j = 1, \ldots, n)$$

are either satisfied for all $p = p_i (i = 1, \ldots, p-1)$ or for no $p_i$. Hence

$$\sum_{q \in q^p} = \frac{1}{p-1} \sum_{q \in q^p} S(x, p, n; \epsilon_1, \ldots, \epsilon_n)$$

where $\sum'$ denotes a summation over all prime ideals $p$ of $\mathcal{O}(\sqrt{1})$ of the first degree, with $N(p) \leq x$ and satisfying the conditions (2.8). Hence

$$(2.9) \quad d'(p) = \lim_{x \to 0} \frac{1}{x} \sum_{q \in q^p} S(x, p, n; \epsilon_1, \ldots, \epsilon_n).$$

If $q_0$ is a prime satisfying the hypothesis of Lemma 2.4 and if $p_0$ is a prime ideal, then by Lemma 2.3 and (2.2) the conditions

$$\left( \frac{a_j}{p_0} \right) \neq 1 \quad (j = 1, \ldots, n)$$

are satisfied and if $\epsilon' = \left( \frac{a_j}{p_0} \right)$, (2.9) gives

$$d'(p) \geq \frac{1}{p^{n(p-1)}} N_0(p, n; \epsilon'_1, \ldots, \epsilon'_n) > 0,$$

by Lemma 2.2.

3. The fundamental equation of Hooley. With Hooley [6], Section 3, p. 210, we start from the observation that $a$ is a primitive root mod $p$ if and only if $p^a$ and for each prime $q \mid p-1$, $a$ is not a $q$th power residue mod $p$. Consequently if $R(q, p)$ denotes the statement

$$(3.1) \quad q \mid p-1 \text{ and at least one of } a_1, \ldots, a_n \text{ is a } q\text{th power residue mod } p,$$

it follows that $N_0(x) = N_0(a_1 a_2 \ldots a_n(x))$ is the number of primes $p \leq x$, $p^a \mid a_1 a_2 \ldots a_n$, such that $R(q, p)$ is false for all primes $q$. Let $N(x, \eta; \epsilon)$ be the number of primes $p \leq x$, $p^a \mid a_1 \ldots a_n$, such that $R(q, p)$ is false for all primes $q \leq \eta$. Then

$$(3.2) \quad N(x, \eta) = \sum_{k} \mu(k)P(x, k)$$

where $k$ runs through 1 and the square-free numbers composed entirely of primes $q \leq \eta$.

Let $\xi_1 = \frac{1}{\log x}$, $\xi_2 = \frac{1}{\log x}$, $\xi_3 = \frac{1}{\log x}$.

If $\eta_1 < \eta_2$ we let $M(x, \eta_1, \eta_2)$ be the number of primes $p \leq x$, $p^a \mid a_1 \ldots a_n$, such that $R(q, p)$ is true for at least one prime $q$, $\eta_1 < q \leq \eta_2$. Then with only slight changes to Hooley’s argument, the fundamental equation

$$(3.3) \quad N(x) = N(x, \xi_1) + O[M(x, \xi_1, \xi_2) + O \left( \frac{x \log \log x}{\log^2 x} \right)]$$

of Hooley follows.

We recall from (3.2) that

$$(3.4) \quad N(x, \xi_1) = \sum_{k} \mu(k)P(x, k)$$

and note that

$$(3.5) \quad k \leq \prod_{q \leq \xi_1} q = e^{\xi_1 \log x} \leq e^{\xi_1} = e^{1/2}.$$

We also observe that

$$(3.6) \quad M(x, \xi_1, \xi_2) \leq \sum_{\xi_1 \leq q} P(x, q).$$

4. A formula for $P(x, k)$. We recall from Section 2 that $P(x, k)$ counts the primes $p \leq x$, $p^a \mid a_1 \ldots a_n$, such that for all primes $q \leq p$ we have $p \equiv 1 \mod q$ and at least one of $a_1, \ldots, a_n$ is a $q$th power residue mod $p$. Thus $P(x, k)$ counts the primes $p \leq x$, $p^a \mid a_1 \ldots a_n$, $p \equiv 1 \mod k$, such that for all primes $q \mid k$ at least one of $a_1, \ldots, a_n$ is a $q$th power residue mod $p$. 

If \( l_1, \ldots, l_n \) are divisors of \( k \) (square-free) we let \( P(x, l_1, \ldots, l_n; k) \) denote the number of primes \( p \leq x \), \( p \nmid a_1 \ldots a_n \), \( p \equiv 1 \mod{k} \), such that each of the congruences
\[
a_i = a_i^1 \mod{p}, \ldots, a_n = a_n^0 \mod{p}
\]
is soluble. Then we have the following formula for \( P(x, k) \).

**Lemma 4.1.**

\[
P(x, k) = \mu(k) \sum_{l_1, \ldots, l_n} \mu(l_1) \ldots \mu(l_n) P(x, l_1, \ldots, l_n; k).
\]

**Proof.** (This is due to Prof. Burgess.) For each \( i, 1 \leq i \leq n \), a multiplicative function \( f_i(l) \) is defined by
\[
f_i(l) = \begin{cases} 1 & \text{if } a_i = a_i^1 \mod{p} \text{ is soluble}, \\ 0 & \text{otherwise}. \end{cases}
\]

Hence if \( f(q) \) is defined by
\[
f(q) = 1 - \prod_{i=1}^{n} (1 - f_i(q)),
\]
we have
\[
f(q) = \begin{cases} 1 & \text{if at least one of } a_i^q = a_i^1 \mod{p}, \\ 0 & \text{otherwise}. \end{cases}
\]

Hence
\[
P(x, k) = \sum_{d \mid k} \mu(d) \prod_{d \mid l} f(q).
\]

But
\[
\prod_{d \mid k} f(q) = \prod_{d \mid k} (1 - \prod_{i=1}^{n} (1 - f_i(q))) = \sum_{d \mid k} \mu(d) \prod_{d \mid q} \prod_{i=1}^{n} (1 - f_i(q))
\]
\[
= \sum_{d \mid k} \mu(d) \prod_{i=1}^{n} \sum_{d \mid l_i} \mu(l_i) f_i(l_i)
\]
\[
= \sum_{d \mid k} \mu(d) \sum_{d \mid l_1} \ldots \sum_{d \mid l_n} \mu(l_1) \ldots \mu(l_n) P(x, l_1, \ldots, l_n; k).
\]

Hence from (4.2)
\[
P(x, k) = \sum_{d \mid k} \mu(d) \sum_{d \mid l_1} \ldots \sum_{d \mid l_n} \mu(l_1) \ldots \mu(l_n) \prod_{l \mid l_i} f_i(l_i)
\]
\[
= \sum_{d \mid k} \mu(d) \sum_{d \mid l_1} \ldots \sum_{d \mid l_n} \mu(l_1) \ldots \mu(l_n) P(x, l_1, \ldots, l_n; k)
\]
\[
= \sum_{d \mid k} \mu(d) \sum_{d \mid l_1} \ldots \sum_{d \mid l_n} \mu(l_1) \ldots \mu(l_n) P(x, l_1, \ldots, l_n; k)
\]
\[
= \sum_{d \mid k} \mu(d) \sum_{d \mid l_1} \ldots \sum_{d \mid l_n} \mu(l_1) \ldots \mu(l_n) P(x, l_1, \ldots, l_n; k)
\]
\[
= \sum_{d \mid k} \mu(d) \prod_{d \mid l} f(q).
\]

The inner sum of (4.3) simplifies on making the substitution \( d = i(l_1, \ldots, l_n) \).

For then
\[
\sum_{d \mid k} \mu(d) = \sum_{d \mid (l_1, \ldots, l_n)} \mu(d) = \mu(l_1, \ldots, l_n) \sum_{d \mid k} \mu(d)
\]
\[
= \begin{cases} \mu(k) & \text{if } \langle l_1, \ldots, l_n \rangle = k, \\ 0 & \text{otherwise}. \end{cases}
\]

Consequently (4.3) reduces to (4.1).

5. **An asymptotic formula for** \( P(x, l_1, \ldots, l_n; k) \). We recall that \( P(x, l_1, \ldots, l_n; k) \) counts those primes \( p \leq x \) which satisfy each of the conditions
\[
(5.1) \quad p = 1 \mod{k}, \quad p \nmid a_1, \quad a_i = a_i^1 \mod{p} \text{ soluble},
\]
where \( 1 \leq i \leq n \).

The argument of Hooley [6], Section 4, pp. 212–213, shows that (5.1) is equivalent to the statement
\[
(5.2) \quad p \nmid k a_1, \quad p \text{ factorises as a product of first degree prime ideals in } Q(V_1, V_a),
\]
where \( P(x, l_1, \ldots, l_n; k) \) counts the primes \( p \leq x \) which satisfy the condition
\[
(5.3) \quad p \nmid k a_1, \ldots, a_n, \quad p \text{ factorises as a product of first degree prime ideals in } K = Q(V_1, V_{a_1}, \ldots, V_{a_n}).
\]

We remark that such prime ideals are distinct, for \( p \nmid k a_1, \ldots, a_n \) implies that \( p \nmid \Delta(K) \), the discriminant of \( K \). (The primes dividing \( \Delta(K) \) are those which divide either \( \Delta(K) \) or \( \Delta(\mathcal{K}) \); also \( \Delta(Q(V_1, V_{a_1})) \) is formed from primes which divide \( k a_1 \).—see Hasse [5], p. 59, Satz 42, and Hooley [6], p. 213.)

Following Hooley, we write \( \pi(x, K) \) for the number of prime ideals \( p \) of \( K \) with \( N(p) < x \). Then
\[
\pi(x, K) = \pi_1(x, K) + \pi_2(x, K)
\]
where \( \pi_1(x, K) \) is the contribution to \( \pi(x, K) \) from the first degree prime ideals not dividing \( k a_1, \ldots, a_n \), and \( \pi_2(x, K) \) is the remaining contribution.

As \( K \) is a Galois extension of \( Q \), we know that each prime \( p \nmid k a_1, \ldots, a_n \) has either \( N(K)(1 = [K : Q]) \) distinct first degree prime ideal factors, or else factorises into distinct prime ideals of a higher degree, the number of factors being less than \( N(K) \). Consequently by (5.3)
\[
(5.5) \quad \pi_1(x, K) = N(K) P(x, l_1, \ldots, l_n; k)
\]
and

$$\pi_\rho(\sigma, K) \leq N(K) \omega(ka_1, \ldots, a_n) + N(K) \sum_{\nu \neq \rho, \rho_1} 1.$$  

Combining (5.4), (5.5) and (5.6) gives

$$N(K)P(x, l_1, \ldots, l_n; k) = \pi(x, K) + O(N(K) \omega(k)) + O(N(K)\sigma^2).$$  

6. A recursive formula for \([F(V_{a_1}, \ldots, V_{a_n}) : F]\). We shall eventually need a lower estimate for \(N(K) = [K : Q]\) where \(K = Q(V_1, V_{a_2}, \ldots, V_{a_n}, a_1, \ldots, a_n)\) are non-zero rational integers and \(k = \langle l_1, \ldots, l_n \rangle\).

Now \(N(K) = [K : Q(V_1)]_p(l)\), and hence it suffices to investigate \([F(V_{a_1}, \ldots, V_{a_n}) : F]\) where \(F\) is a number field containing all \(k\)-th roots of unity.

**Lemma 6.1.** Let \(F\) be a number field containing all \(k\)-th roots of unity, \(a_1, \ldots, a_n\) are non-zero elements of \(F\) and \(l_1, \ldots, l_n\) are divisible by \(k\), a square-free integer. Positive integers \(\lambda_1, \ldots, \lambda_n, \lambda'_1, \ldots, \lambda'_n\) are defined as follows: \(\lambda'_i\) is the product of the those primes \(p \mid l_i\) such that \(a_i = p^{\lambda_i} \beta_1 \cdots \beta_{\lambda'_i} \epsilon F\); \(l_i = \lambda_i / \lambda'_i\).

If \(1 \leq r \leq n\), \(\lambda'_r\) is the product of those primes \(p \mid l_r\) for which integers \(m_1, \ldots, m_r\) exist satisfying

$$a_r = a_1^{m_1} \cdots a_r^{m_r} \cdots a_n^{m_r} \beta^{m_r}, \quad \epsilon F,$$

where in addition, \(p \mid m_i \Rightarrow p \mid l_i\), for \(1 \leq i \leq r - 1\); also \(l_r = l_r / \lambda'_r\). Then

$$[F(V_{a_1}, \ldots, V_{a_n}) : F] = \lambda_1 / \lambda'_1.$$

**Proof.** Let \(J_0 = F, J_n = F(V_{a_1}, \ldots, V_{a_n})\), and for \(1 \leq r \leq n\) let

$$J_r = F(V_{a_1}, \ldots, V_{a_{n-r}}), \quad K_r = F(V_{a_1}, \ldots, V_{a_r}).$$

Then the following statements can be proved by induction on \(r\):

(i) \(J_r = K_r\) for \(0 \leq r \leq n\).

(ii) \(x^{\lambda_r} - a_r\) is irreducible over \(K_{r-1}\), for \(1 \leq r \leq n\).

Equation (6.1) now follows immediately, for from (ii) we have

$$[K_r : K_{r-1}] = \lambda_r, \quad 1 \leq r \leq n.$$  

and hence

$$[J_r : F] = [K_r : F] = \prod_{r=1}^n [K_r : K_{r-1}] = \prod_{r=1}^n \lambda_r.$$  

**Proof of (i).** The case \(r = 0\) needs no proof. Let \(1 \leq r \leq n\) and assume that \(J_{r-1} = K_{r-1}\). Then

$$J_r = J_{r-1}(V_{a_r}) = K_{r-1}(V_{a_r}).$$

We now write \(l_r = p_1 \cdots p_t, \) where \(p_1, \ldots, p_t\) are the prime factors of \(\lambda_r\). Then

$$K_{r-1}(V_{a_r}) = K_{r-1}(V_{a_1}, \ldots, V_{a_r}).$$  

But from construction of \(\lambda_r\), we have for \(p = p_1, \ldots, p_t\)

$$p^{\lambda_r} = p^{\lambda_r} \cdots p_t^{\lambda_r} \beta, \quad \epsilon F,$$

where \(\beta \in F\) and \(p \mid m_i \Rightarrow p \mid l_i\). Hence for such \(p\) we have \(V_{a_r} \epsilon K_{r-1}\) and hence \(K_{r-1}(V_{a_r}) = K_{r-1}\). From (6.3) it follows that

$$K_{r-1}(V_{a_r}) = K_{r-1}(V_{a_1}) = K_r,$$

and hence from (6.2) we have \(J_r = K_r\), completing the induction.

Before we can prove (ii) we need the following result which is basically Sata 150 of Hasse [5], pp. 218–221.

**Lemma 6.2.** Let \(F\) be a number field containing all \(k\)-th roots of unity and let \(a\) be a non-zero element of \(F\). Also assume that \(a^r - a\) is irreducible over \(F\), where \(l \mid k\). Then if \(\beta\) is a non-zero element of \(F\) with \(\beta = \gamma^p, \gamma \epsilon F(V_1)\) and \(p \mid k\), we have \(\beta = a^p - a\), where \(\epsilon F\) and \(p \mid m_i \Rightarrow p \mid l_i\), for \(1 \leq i \leq r\).

**Proof of (ii).** For \(1 \leq r \leq n\) let \(F_r\) denote the statement

(a) \(x^r - a_r\) is irreducible over \(K_{r-1}\) and

(b) if \(\beta\) is a non-zero element of \(F\) such that \(\beta = \gamma^p, \gamma \epsilon K_{r-1}\), then \(\beta = a^p - a\), where \(\epsilon F\) and \(p \mid m_i \Rightarrow p \mid l_i\), for \(1 \leq i \leq r\).

We use a well-known criterion for the irreducibility of \(a^r - a\) over a field \(H\), stated for example in Lang [9], Theorem 16, p. 221. For square-free \(l\) this states that \(a^r - a\) is irreducible over \(H\) if \(p \mid r \Rightarrow p \not\mid \beta^r, \beta \epsilon H\). (\(H\) is assumed to be of characteristic zero or prime to \(l\)).

We prove by induction on \(r\) that \(P_r\) holds for \(1 \leq r \leq n\). Our proof is based on that of Elliott [3], Lemma 3, pp. 134–135. When \(r = 1\), (a) follows from the construction of \(\lambda_1\) and the above mentioned criterion for irreducibility, while (b) reduces to the statement of Lemma 6.2.

Hence we assume \(1 \leq r \leq n\) and that \(P_r\) is valid for \(s < r\). We argue indirectly and assume that \(a^r - a_r\) is reducible over \(K_{r-1}\). Then for some \(p \mid l_r\) we have

$$a_r = \beta^p, \quad \beta \epsilon K_{r-1},$$  

and

$$a_r = \beta^p, \quad \beta \epsilon K_{r-1}.$$
A generalisation of Artin’s conjecture

Proof. Let \( F \) denote the ring of integers of \( F \). Then the tower formula for discriminant ideals (see Cassels and Fröhlich [2], Proposition 7(ii), p. 17) gives

\[
\begin{align*}
\mathfrak{b}(F(a)) &= \mathfrak{b}(F)^g N_{F/Q}(\mathfrak{b}(F(a)/F)),
\end{align*}
\]

where \( \mathfrak{b}(F(a)) \) is an “absolute” discriminant and \( \mathfrak{b}(F(a)/F) \) is a “relative” discriminant. Now

\[
\begin{align*}
\mathfrak{b}(F(a)/F) &= N_{F/Q}(g'(a)) F,
\end{align*}
\]

where \( g'(a) = a^x - a \). (See Cassels and Fröhlich [2], Proposition 6(ii), p. 17.) Hence

\[
\mathfrak{b}(F(a)/F) = N_{F/Q}(g'(a)) F = N_{F/Q}(k^a - 1) F = k^a a^x - 1 F.
\]

Also \( F[a] \) is a finitely generated \( \overline{F} \) sub-module of \( F(a) \) and hence \( \mathfrak{b}(F[a]/F) \) divides \( \mathfrak{b}(F(a)/F) \), and hence divides \( k^a a^x - 1 F \). (See Cassels and Fröhlich [2], Corollary 1, p. 12.) Hence \( N_{F/Q}(\mathfrak{b}(F(a)/F)) \) divides \( N_{F/Q}(k^a a^x - 1 F) \), and consequently divides \( (k^a a^x - 1)^{\mathfrak{b}(F[a]/F)} Z \), as \( a \neq Z \).

Equation (7.3) now shows that \( \mathfrak{b}(F(a)) \) divides \( \mathfrak{b}(F)^g(k^a a^x - 1)^{\mathfrak{b}(F[a]/F)} Z \), and hence (7.2) holds.

**Lemma 7.2.** Let \( K = F(V_{a_1}, \ldots, V_{a_n}) \) where \( F \) is a number field containing all \( h \)-th roots of unity, \( a_1, \ldots, a_n \) are non-zero rational integers and \( i_1, \ldots, i_n \) are divisors of \( k \), a square-free integer. Then

\[
\begin{align*}
D(K) &= D(K)[a_1^{i_1} \ldots a_n^{i_n}] K^{[K:Q]}.
\end{align*}
\]

Proof. From Lemma 6.1 we recall that certain divisors \( \lambda_1, \ldots, \lambda_n \) of \( i_1, \ldots, i_n \), respectively were constructed with the property that

\[
\begin{align*}
K = F(V_{a_1}, \ldots, V_{a_n}) \quad \text{and} \quad a^h - a \text{ is irreducible over} \quad K_{i-1}.
\end{align*}
\]

Hence Lemma 7.1 may be applied with \( F, k \) and \( a \) replaced by \( K_{i-1}, \lambda_i \) and \( a_i \), respectively. Writing \( \Delta_i = D(K_i) \) we have

\[
\begin{align*}
\Delta_i &= \Delta_i(k_i^{[K_i:Q]} a_i^{i_k}) \quad \text{for} \quad 1 \leq i \leq n.
\end{align*}
\]

It follows by induction on \( i \) that

\[
\begin{align*}
\Delta_i &= \Delta_i(k_i^{[K_i:Q]} a_i^{i_k}) \quad \text{for} \quad 1 \leq i \leq n.
\end{align*}
\]

To prove Lemma 7.2 we consider (7.6) with \( i = n \). We have

\[
\begin{align*}
\Delta(K) &= \Delta_i(k_i^{[K_i:Q]} a_i^{i_k}) \quad \text{for} \quad 1 \leq i \leq n.
\end{align*}
\]

and (7.4) follows since \( \lambda_i | i_i \) for \( 1 \leq i \leq n \) and \( \lambda_1 \ldots \lambda_n = [K:F] \) by Lemma 6.1.

7. An upper estimate for \( D(K) \). The argument of Hooley [6] (Section 5) dealt with the Dedekind zeta function of the field \( Q(V_1, V_2) \). To enable the argument of that section to carry over to \( K = Q(V_1, V_{a_1}, \ldots, V_{a_n}) \), it suffices to verify that

\[
\begin{align*}
|D(K)| \leq L^{A(N_K)}
\end{align*}
\]

for some positive constant \( A \). This is a consequence of Lemma 7.3 below.

**Lemma 7.1.** Let \( F \) be a number field and let \( a^2 = a \), a non-zero rational integer. Then if \( a^h - a \) is irreducible over \( F \), we have

\[
\begin{align*}
\Delta(F(a)) &= \Delta(F)^g(k^a)^{[F(a):Q]}.
\end{align*}
\]
We remark that a formula, similar to (8.5) but with a weaker error term, may be proved without any Riemann hypothesis, using the prime ideal theorem (see Landau [7], Satz 191, p. 110) instead of (8.1). Hence we have

**Lemma 8.1.** Let $k$ be square-free. Then the primes $p \equiv 1 \pmod{k}$ such that for each prime $q|k$ at least one of $a_1, \ldots, a_n$ is a $q$-th power residue mod $p$, have a natural density $c(k)$ given by (8.4).

The expression for $P(x, k)$ given by (8.5) is now substituted in (3.4) to give

\[
N(x, \xi^k) = \frac{1}{N(K)} \frac{1}{k^x} \sum_{\substack{p \leq x \atop p \equiv 1 \pmod{k}}} \mu(p) e(k/p) + O \left( x^{\delta + \epsilon} \sum_{k \leq x^{1/2+\epsilon}} |\mu(k)| \right)
\]

\[
= \frac{1}{N(K)} \frac{1}{k^x} \sum_{\substack{p \leq x \atop p \equiv 1 \pmod{k}}} \mu(p) e(k/p) + O \left( x^{\delta + \epsilon} \right)
\]

To show that the series $\sum_{k=1}^{\infty} \mu(k) e(k)$ converges absolutely we need the following lower estimate for $N(K)$.

**Lemma 8.2.** Let $a_1, \ldots, a_n$ be non-zero rational integers not $\pm 1$ and let $h_i$ be the largest positive integer such that $a_i$ is a perfect $h_i$-th power. Also let $k = \langle h_1, \ldots, h_n \rangle$ be square-free.

\[
K = \mathcal{O}(V_{h_1}, V_{h_2}, \ldots, V_{h_n}), \quad N(K) = [K:Q].
\]

Then

\[
\frac{1}{N(K)} \leq \frac{\prod_{i=1}^{n} (h_i, k)}{k^x}.
\]

**Proof.** From Lemma 6.1 we have

\[
\frac{1}{N(K)} = \frac{1}{\varphi(k)} \sum_{l_1 \ldots l_n} \prod_{i=1}^{n} (h_i, l_i).
\]

The construction of $\lambda^*_1, \ldots, \lambda^*_n$ reveals the following chain of inequalities:

\[
\lambda^*_1 \leq (h_1, l_1), \quad \lambda^*_2 \leq (h_2, l_2)(l_1, l_2), \quad \lambda^*_3 \leq (h_3, l_3)(l_2, l_3)(l_1, l_3), \quad \text{and so on.}
\]

We prove by induction that for $1 \leq j \leq n$

\[
\frac{1}{l_1 \cdots l_n} \leq \prod_{i=1}^{n} \frac{(h_i, l_i)}{(l_i, l_j)}.
\]
Inequality (8.8) is clearly true when \( r = 1 \). Consequently let \( 1 \leq r < n \) and assume that (8.8) holds. Then
\[
\frac{\lambda_1^r \cdots \lambda_{r+1}^r}{\lambda_1 \cdots \lambda_{r+1}} \leq \left( \prod_{t=1}^{r} (\lambda_t, \lambda_t) \right) \left( \prod_{t=1}^{r} (\lambda_t, \lambda_t) \right) \left( \prod_{t=1}^{r} (\lambda_t, \lambda_t) \right) = \prod_{t=1}^{r+1} (\lambda_t, \lambda_t) = \frac{\lambda_1^r \cdots \lambda_{r+1}^r}{\lambda_1 \cdots \lambda_{r+1}}.
\]
This completes the induction.

**Lemma 8.3.** Let \( a_1, \ldots, a_n \) be non-zero rational integers not \( \pm 1 \) and let
\[ K = Q(V_1, V_{a_1}, \ldots, V_{a_n}), \quad N(K) = [K : Q]. \]
Then if \( \phi(k) \) is defined (as in (8.4)) by
\[ \phi(k) = \mu(k) \sum_{l_1 \mid k_1} \sum_{l_2 \mid k_2} \frac{\mu(l_1) \cdots \mu(l_n)}{N(K)}, \]
the series \( \sum_{k=1}^{\infty} \phi(k) \) is absolutely convergent.

**Proof.** From (8.7) we have
\[
|\phi(k)| \leq \sum_{l_1 \mid k_1} \sum_{l_2 \mid k_2} \frac{1}{\varphi(l)} = \left( \frac{2^n - 1}{n} \right) \frac{\varphi(k)}{k \varphi(k)} \leq \frac{\varphi(k)}{k \varphi(k)}.
\]
Now
\[
\frac{\varphi(k)}{k \varphi(k)} = \frac{k^{\log \log k}}{k^2} \quad \text{for each } k > 0,
\]
and hence
\[
\sum_{k=1}^{\infty} \frac{\varphi(k)}{k \varphi(k)} = \sum_{k=1}^{\infty} \frac{\varphi(k)}{k^2 \varphi(k)}
\]
converges by comparison with
\[
\sum_{k=1}^{\infty} \frac{\log \log k}{k^{2+\varepsilon}}, \quad 0 < \varepsilon < 1.
\]
Consequently \( \sum_{k=1}^{\infty} \phi(k) \) converges absolutely by (8.9).

The finite sum \( \sum_{k=1}^{\infty} \mu(k) \phi(k) \) occurring in (8.6) may now be replaced by \( \sum_{k=1}^{\infty} \mu(k) \phi(k) \) with an error that is estimated by

**Lemma 8.4.** Let
\[ S(z) = \sum_{k=1}^{\infty} \frac{d^z(k)}{k \varphi(k)}. \]
Then
\[
S(z) \leq a^{-1}(\log a)^{2n-1}.
\]

**Proof.** (Kindly supplied by Mr. M. Croft.) By a Stieltjes integration
\[
S(z) = \int a \frac{1}{t} \left( \sum_{d < t \varphi(d)} \frac{k}{\varphi(d)} \right) \frac{dt}{t} - \frac{1}{z} \left( \sum_{d < t \varphi(d)} \frac{k}{\varphi(d)} \right) \frac{dt}{t}.
\]
Write
\[
\Sigma_1 = \sum_{d < t \varphi(d)} \frac{k}{\varphi(d)} \left( \frac{a}{t} \right) \frac{dt}{t}.
\]
Then
\[
\Sigma_1 \leq \sum_{d < t \varphi(d)} \sum_{k \mid d} \frac{d^z(k)}{k \varphi(k)} \left( \frac{1}{t} \frac{1}{t} \right) = \sum_{d < t \varphi(d)} \frac{d^z(k)}{k \varphi(k)} \left( \frac{1}{t} \frac{1}{t} \right) = \sum_{d < t \varphi(d)} \frac{d^z(k)}{k \varphi(k)} \left( \frac{1}{t} \frac{1}{t} \right).
\]
by Wilson [12].

Hence
\[
\Sigma_1 \leq t(\log t)^{2n-1}.
\]
Hence from (8.11), (8.12) and (8.13) we have
\[ S(z) \leq \int a \frac{1}{t} (\log t)^{2n-1} dt \leq a^{-1}(\log a)^{2n-1}. \]
From (8.6) and (8.10) we deduce that

\[
N(x, \xi_1) = \lim_{x \to \infty} \sum_{k=1}^{\infty} \mu(k) e(k) \log \xi_1 + O\left(\frac{x}{\log x} \sum_{k=1}^{\infty} \frac{d^2(k)}{k \log(k)}\right) + O(x^{\theta + \epsilon})
\]

\[
= \lim_{x \to \infty} \sum_{k=1}^{\infty} \mu(k) e(k) \log \xi_1 + O\left(\frac{x}{\log x} \frac{\log \xi_1^{\theta + \epsilon}}{\xi_1}\right) + O(x^{\theta + \epsilon})
\]

\[
= \lim_{x \to \infty} \sum_{k=1}^{\infty} \mu(k) e(k) \log \xi_1 + O\left(\frac{x}{\log x} \left(\log \log x\right)^{\theta + \epsilon} \xi_1\right)
\]

From (3.3) it remains to estimate \(M(x, \xi_1, \xi_2)\) from above, using (3.6). From (3.3) we have

\[
P(x, q) = e(q) \log x + O(x^{\theta + \epsilon}) \leq \frac{1}{(q-1)} \log x + O(x^{\theta + \epsilon})
\]

by (8.9).

Hence from (3.6), following Hooley [6], end of Section 6, we have (recalling that \(\xi_1 = \frac{x}{\log x}\) and \(\xi_2 = x^{\theta + \epsilon} \log^{-1} x\))

\[
M(x, \xi_1, \xi_2) \leq \sum_{1 \leq q \leq x} \frac{\log x}{q} + O(x^{\theta + \epsilon} \log x)
\]

\[
= O\left(\frac{x}{\log x} \sum_{1 \leq q \leq x} \frac{1}{q}\right) + O(x^{\theta + \epsilon} \sum_{1 \leq q \leq x} \frac{1}{q})
\]

\[
= O\left(\frac{x}{\xi_1} \log \xi_1 + O\left(\frac{x^{\theta + \epsilon} \xi_2 \log \xi_2}{\log \xi_1}\right) = O\left(\frac{x}{\log x}\right),
\]

which is the estimate required.

Finally from (3.3), (8.14) and (8.15) we have

\[
N(x) = \sum_{k=1}^{\infty} \mu(k) e(k) + O\left(\frac{x}{\log x} \log \xi_1^{\theta + \epsilon} \xi_2\right),
\]

where \(e(k)\) is defined by (8.4) and has an interpretation given by Lemma 8.1.

The reader is reminded that (8.16) has been derived on the assumption that none of \(a_1, \ldots, a_n\) is \(\pm 1\) and that the Riemann hypothesis holds for each of the fields \(\mathbb{Q}(V, V_{a_1}, \ldots, V_{a_n})\) where \(k = \langle t_1, \ldots, t_n\rangle\) is square-free.

9. Another formula for \([F(V_{a_1}, \ldots, V_{a_n}):F]\).

Lemma 9.1. Let \(F\) be a number field containing all \(k\)-th roots of unity, \(a_1, \ldots, a_n\) are non-zero elements of \(F\), and \(l_1, \ldots, l_n\) are divisors of \(k\). Then

\[
[F(V_{a_1}, \ldots, V_{a_n}):F] = l_1 \cdots l_n \prod_{\beta \in \mathbb{F}} \prod_{\alpha \in \mathbb{F}^*} (\alpha^l_{\beta})^{\dim \mathbb{F}/\mathbb{F}^*}
\]

Proof. (9.1) is a consequence of the formula

\[
[F(V_{a_1}, \ldots, V_{a_n}):F] = \prod_{\beta \in \mathbb{F}} \prod_{\alpha \in \mathbb{F}^*} (\alpha^l_{\beta})^{\dim \mathbb{F}/\mathbb{F}^*}
\]

For \(F(V_{a_1}, \ldots, V_{a_n}) = F(V_{a_1}^{l_1}, \ldots, V_{a_n}^{l_n})\) and by (9.2) with \(a_1, \ldots, a_n\) replaced by \(a_1^{l_1}, \ldots, a_n^{l_n}\), we have

\[
[F(V_{a_1}, \ldots, V_{a_n}):F] = \prod_{\beta \in \mathbb{F}} \prod_{\alpha \in \mathbb{F}^*} (\alpha^l_{\beta})^{\dim \mathbb{F}/\mathbb{F}^*}
\]

\[
= \prod_{\beta \in \mathbb{F}} \prod_{\alpha \in \mathbb{F}^*} (\alpha^{l_1 \cdots l_n})^{\dim \mathbb{F}/\mathbb{F}^*}
\]

which gives (9.1).

To prove (9.2) we argue as follows.

Let \(F^*\) and \(F^{k*}\) denote the multiplicative groups of non-zero elements of \(F\) and the \(k\)-th powers of the elements of \(F^*\) respectively. Following Hasse [5], pp. 223–233, we let \(\{a_1, \ldots, a_n, y^k\}\) denote the multiplicative group generated by \(a_1, \ldots, a_n\) and \(F^{k*}\). Then the reader is referred to Hasse for a proof of the fact that the Galois group of \(F(V_{a_1}, \ldots, V_{a_n})\) over \(F\) is isomorphic to \(\{a_1, \ldots, a_n, y^k\}/F^{k*}\). (See Hasse [5], Satz 152, p. 223.) In particular

\[
[F(V_{a_1}, \ldots, V_{a_n}):F] = \prod_{\beta \in \mathbb{F}} \prod_{\alpha \in \mathbb{F}^*} (\alpha^l_{\beta})^{\dim \mathbb{F}/\mathbb{F}^*}
\]

However, it is not difficult to prove that if \(S\) is the abelian group formed by all \(n\)-tuples of residue classes mod \(k\), while \(T\) is the subgroup of \(S\) formed by those \(n\)-tuples \((r_1, \ldots, r_n)\) of residues mod \(k\) which satisfy

\[
a_1^{r_1} \cdots a_n^{r_n} = \beta^k, \quad \beta \in \mathbb{F}
\]
then $S/T$ is isomorphic to $\{a_1, \ldots, a_n, y^{l/b}\}^{F^{ab}}$. Now

\[(9.4) \quad |S| = b^k \quad \text{and} \quad |T| = \sum_{\delta \in \mathbb{Q}^{h \delta}} \ldots \sum_{\delta \in \mathbb{Q}^{h \delta}} 1,
\]

and hence (9.2) follows from (9.3) and (9.4).

A "transcendental" proof in the case when $F = \mathbb{Q}(\sqrt{1})$ may be constructed from Lemma 1, p. 162, of Schinzel [11].

10. **Expressing $c(k)$ in terms of the multiplicative function $c'(k)$.** It is convenient to state some results on $\mathbb{Q}(\sqrt{1})$.

**Lemma 10.1.** Let $k$ be a square-free positive integer and let $a$ be a non-zero rational integer. Then

(i) $a = \beta^h, \beta \in \mathbb{Q}(\sqrt{1})$ implies $k$ is odd, $a = b^k, b \in \mathbb{Z}$, or $k$ is even, $a = b^{k/2}, b \in \mathbb{Z}, V \beta \in \mathbb{Q}(\sqrt{1})$,

(ii) if $k$ is even,

$$a = \beta^{k/2}, \beta \in \mathbb{Q}(\sqrt{1}) \Rightarrow a = b^{k/2}, b \in \mathbb{Z},$$

(iii) $V \in \mathbb{Q}(\sqrt{1}) \Rightarrow \alpha(a) = 1(\mod 4)$ and $\alpha(a) = 1(\mod 4)$.

Proof. For (i) and (iii) see Schinzel [11], Lemmas 3 and 4, p. 162.

To prove (ii) assume that $a = \beta^{k/2}, \beta \in \mathbb{Q}(\sqrt{1})$. Then $a^2 = \beta^{k/2}$ and by (i) $a^2 = b^{k/2}, b \in \mathbb{Z}$. Hence $b = \epsilon^{k/2}, \epsilon \in \mathbb{Z}$, as $k/2$ is odd. Then $a^2 = (\epsilon^{k/2})^2$ and $a = (\pm \epsilon^{k/2})$ as required.

From (9.1) and the definition of $c(k)$ given in (8.4) we have

\[(10.1) \quad c(k) = \frac{\mu(k)}{\varphi(k)} \sum_{i \in \mathbb{Z}^{(k)}} \ldots \sum_{i \in \mathbb{Z}^{(k)}} \mu(l_i) \ldots \mu(l_n) d(l_1, \ldots, l_n; k),
\]

where

\[(10.2) \quad d(l_1, \ldots, l_n; k) = \prod_{i=1}^{l_1} \ldots \prod_{n}^{l_n} 1,
\]

\[(10.3) \quad d'(l_1, \ldots, l_n; k) = \sum_{i=1}^{l_1} \ldots \sum_{i=1}^{l_n} 1,
\]

and

\[(10.4) \quad c'(k) = \frac{\mu(k)}{\varphi(k)} \sum_{i=1}^{l_1} \ldots \sum_{i=1}^{l_n} \mu(l_i) \ldots \mu(l_n) d'(l_1, \ldots, l_n; k).
\]

Then from Lemma 10.1(i) we have immediately the following

**Lemma 10.2.** Let $k$ be an odd square-free integer. Then

$$e(k) = c'(k).$$

(For even $k$ the relation between $c(k)$ and $c'(k)$ is much more complicated.)

We will later need the fact that $c'(k)$ is multiplicative. This is a consequence of the following two lemmas:

**Lemma 10.3.** Let $l = \langle l_1, \ldots, l_n \rangle, m = \langle m_1, \ldots, m_n \rangle$ and $(l, m) = 1$.

Then

$$d(l_1 m_1, \ldots, l_n m_n; lm) = d(l_1, \ldots, l_n; l) d'(m_1, \ldots, m_n; m).$$

**Lemma 10.4.** Let $f$ satisfy

$$f(l_1, m_1, \ldots, l_n, m_n; lm) = f(l_1, \ldots, l_n; l) f(m_1, \ldots, m_n; m)$$

whenever $l = \langle l_1, \ldots, l_n \rangle, m = \langle m_1, \ldots, m_n \rangle$ and $(l, m) = 1$. Then the function $g$ defined by

$$g(k) = \sum_{i=1}^{l_1} \ldots \sum_{i=1}^{l_n} f(l_1, \ldots, l_n; k)$$

is multiplicative.

**Proof** of Lemma 10.3. Let $a_1^{(l_1 m_1)} \ldots a_n^{(l_n m_n)} = b^k, b \in \mathbb{Z}$ and $a_1^{(l_1 m_1)} \ldots a_n^{(l_n m_n)} = c^k, c \in \mathbb{Z}$. Then

$$a_1^{(l_1 m_1 + m_1 l_1)} \ldots a_n^{(l_n m_n + m_n l_n)} b^{(l_1 m_1 + m_1 l_1) \ldots b^{(l_n m_n + m_n l_n)}} = (bc)^m.$$

Now the residues $l_i m_i + \mu_i l_i \mod l_i$ are in 1-1 correspondence with the ordered pairs $(l_i, \mu_i)$ of residues mod $l_i$ and $m_i$, respectively. Consequently it remains to show that every $(r_1, \ldots, r_n)$ which contributes to $d'(l_1, \ldots, l_n; l m_n; l m)$ arises in the above way. Accordingly we assume that

\[(10.5) \quad a_1^{(l_1 m_1 + m_1 l_1)} \ldots a_n^{(l_n m_n + m_n l_n)} = d^m, \quad d \in \mathbb{Z},
\]

and define $n$-tuples $(l_1, \ldots, l_n)$ and $(\mu_1, \ldots, \mu_n)$ of residues by the congruences

\[l_i m_i + \mu_i l_i = r_i \mod l_i m_i, \quad 1 \leq i \leq n.
\]

From (10.5) we deduce that

$$a_1^{(l_1 m_1 + m_1 l_1) \ldots a_n^{(l_n m_n + m_n l_n)}} = (d')^m.$$
and hence
\[(a_i^{(l_i)} \ldots a_n^{(l_n)})^m = (d''^{k'})^{ \frac{1}{m}}, \quad d'' \in \mathbb{Z}.\]

It follows from (10.6) and \((l, m) = 1\) that \(d'' = f^m, f \in \mathbb{Z}\). Hence from (10.6)
\[a_i^{(l_i)} \ldots a_n^{(l_n)} = \xi f^m,
\]
where \(\xi\) is an \(m\)-th root of unity, necessarily \(\pm 1\). Hence
\[a_i^{(l_i)} \ldots a_n^{(l_n)} = (\pm f)^m\]
and \((l_1, \ldots, l_n)\) contributes to \(d'(l_1, \ldots, l_n; l)\); similarly for \((\mu_1, \ldots, \mu_n)\).
This completes the proof of Lemma 10.3.

Proof of Lemma 10.4. Let \((l, m) = 1\). Then
\[g(lm) = \sum_{k_1 \mid l} \sum_{k_2 \mid m} f(k_1, \ldots, k_n; lm).
\]

We now write \(k_i = l_i m_i\) where \(l_i | l\) and \(m_i | m\). Then
\[\langle l_1, \ldots, l_n \rangle = \langle m_1, \ldots, m_n \rangle = \langle m_1 \rangle \text{ and } \langle m_1, \ldots, m_n \rangle = m.
\]

Hence
\[g(lm) = \sum_{l_1 | l} \sum_{m_1 | m} \sum_{l_2 | l} \sum_{m_2 | m} f(l_1, m_1; \ldots, l_n, m_n; lm).
\]

Thus for \((l, m) = 1, m \geq 1\),
\[D(m_1, \ldots, m_n; k) = \sum_{m_1, \ldots, m_n}^{m_1, \ldots, m_n} \frac{1}{a_i^{(l_i)} \ldots a_n^{(l_n)} = k}.
\]

Then with \(a(k)\) defined as in (10.1), we have
\[a(k) = c'(k)[2] c''(k).
\]

The next result will be needed when \(c''(k)\) is expanded further. It will also prove useful when the vanishing of \(\chi(2) = 1 - c''(2)\) is considered later.

Lemma 10.7. Let \(\mathbb{F}\) be a field and let \(p\) be a prime. For \(1 \leq i_1 < \ldots < i_j < n\) let
\[\tau(i_1, \ldots, i_j) = \sum_{i_1 < \ldots < i_j < n} 1.
\]

Also let
\[\sigma_n = \sum_{1 \leq i_1 < \ldots < i_j < n} \tau(i_1, \ldots, i_j).
\]

Then by Lemma 10.1 (ii) we may assume that \(d'' \in \mathbb{Z}\), and the rest of the argument is the same as before, with \((\lambda_1, \ldots, \lambda_n)\) contributing to \(d'(l_1, \ldots, l_n; k)[2]\).

(ii) Assume that
\[a_i^{(l_i)} \ldots a_n^{(l_n)} = (d'')^l, \quad d'' \in \mathbb{Q}(\sqrt{1}).
\]

Then from Lemma 10.1 (iii) we have \(\chi(a^{[2]}) = 1 (\text{mod} \, 4)\) and \(\chi(a^{[2]}) = k\) where \(a = a_i^{(l_i)} \ldots a_n^{(l_n)}\). But \(\chi(a^{[2]}) = \chi(a)\) and hence by Lemma 10.1 (iii) again, we have \(a = f^m, f \in \mathbb{Q}(\sqrt{1})\). Hence \((\mu_1, \ldots, \mu_n)\) contributes to \(D(m_1, \ldots, m_n; k)\).

This completes the proof of Lemma 10.5.

The next result is proved in a fashion similar to the proof of Lemma 10.4.

Lemma 10.8. Let \(k\) be an even square-free integer and let
\[c''(k) = - \sum \sum \mu(m_1) \ldots \mu(m_n) D(m_1, \ldots, m_n; k)
\]

where \(D(m_1, \ldots, m_n; k)\) is defined (as in (10.7)) by
\[D(m_1, \ldots, m_n; k) = \sum m_1 \ldots m_n 1.
\]

Then with \(a(k)\) defined as in (10.1), we have
\[a(k) = c'(k)[2] c''(k).
\]
Then if $r^*(i_1, \ldots, i_j)$ and $a_i^*$ are defined similarly, but with none of $r_1, \ldots, r_i$ divisible by $p$, we have

$$\sum_{j=0}^n (-1)^j p^{n-j} a_j = \sum_{j=0}^n (-1)^j (p-1)^{n-j} a_j^*.$$ \hspace{1cm} (10.12)

**Proof.** The following identity holds:

$$\tau(i_1, \ldots, i_j) = 1 + \sum_{1 \leq i_1 < \cdots < i_j \leq n} \tau^*(i_1, \ldots, i_j) + \sum_{1 \leq i_1 < i_2 \leq n} \tau^*(i_1, i_2) + \cdots + \tau^*(i_1, \ldots, i_j).$$

Hence

$$\tau = \sum_{1 \leq i_1 < \cdots < i_j \leq n} \tau(i_1, \ldots, i_j) = \sum_{1 \leq i_1 < i_2 \leq n} \tau(i_1, i_2) + \sum_{1 \leq i_1 < i_2 < \cdots < i_j \leq n} \tau^*(i_1, \ldots, i_j) = \left( \begin{array}{c} n \end{array} \right) + \sum_{j=1}^n \left( \begin{array}{c} n-j \end{array} \right).$$

But the inner sum on the right of (10.13) is the number of subsets of a given set of $n$ elements, each subset containing $j$ elements and containing a given subset of $r$ elements. This number is $\binom{n-r}{j-r}$. Hence

$$\tau_j = \left( \begin{array}{c} n \end{array} \right) + \sum_{j=1}^n \left( \begin{array}{c} n-j \end{array} \right) \tau_j^* = \left( \begin{array}{c} n \end{array} \right) + \sum_{j=1}^n \left( \begin{array}{c} n-j \end{array} \right) \tau_j^*.$$ \hspace{1cm} (10.14)

Hence from (10.14) we have

$$\sum_{j=0}^n (-1)^j p^{n-j} a_j = \sum_{j=0}^n (-1)^j p^{n-j} \sum_{r=0}^n \left( \begin{array}{c} n-j \end{array} \right) \tau_j^* = \sum_{j=0}^n \tau_j^* \sum_{j=0}^n (-1)^j p^{n-j} \left( \begin{array}{c} n-j \end{array} \right) = \sum_{j=0}^n (-1)^j (p-1)^{n-j} \tau_j^*,$$

as asserted.

**Lemma 10.8.** Let $k$ be an even square-free integer and let $\tau(k)$ be defined as in (10.9). Then

$$\tau(k) = \frac{1}{2n} \sum_{j=1}^n (-1)^j \sum_{1 < i_1 < \cdots < i_j \leq n} \frac{1}{\tau(s_{i_1}, \ldots, s_{i_j})},$$

where $\tau(s_{i_1}, \ldots, s_{i_j})$ is the product of all $\tau_i$'s for which $i_1, \ldots, i_j$ are not divisible by $p$.

**Proof.**

$$\tau(k) = \frac{1}{2n} \sum_{j=1}^n (-1)^j \sum_{1 < i_1 < \cdots < i_j \leq n} \frac{1}{\tau(s_{i_1}, \ldots, s_{i_j})} = \frac{1}{2n} \sum_{j=1}^n (-1)^j \sum_{1 < i_1 < \cdots < i_j \leq n} \frac{1}{\tau^*(s_{i_1}, \ldots, s_{i_j})} + \cdots + \frac{1}{\tau^*(s_{i_1}, \ldots, s_{i_j})},$$

$$= \frac{1}{2n} \sum_{j=1}^n (-1)^j \sum_{1 < i_1 < \cdots < i_j \leq n} \frac{1}{\tau(s_{i_1}, \ldots, s_{i_j})},$$

by Lemma 10.7.

Similarly

$$\tau(2) = \frac{1}{2n} \sum_{j=1}^n (-1)^j \sum_{1 < i_1 < \cdots < i_j \leq n} \frac{1}{\tau(s_{i_1}, \ldots, s_{i_j})} = \frac{1}{2n} \sum_{j=1}^n (-1)^j \sum_{1 < i_1 < \cdots < i_j \leq n} \frac{1}{\tau^*(s_{i_1}, \ldots, s_{i_j})},$$

which reduces to (10.15) by Lemma 10.1 (iii).

From Lemma 10.6 and the multiplicativity of $\tau(k)$ we deduce immediately from Lemma 10.8 the formula for $\tau(k)$, $k$ even:

**Lemma 10.9.** Let $k$ be an even square-free integer. Then

$$\tau(k) = \frac{1}{2n} \sum_{j=1}^n (-1)^j \sum_{1 < i_1 < \cdots < i_j \leq n} \frac{1}{\tau(s_{i_1}, \ldots, s_{i_j})},$$

$$= \frac{1}{2n} \sum_{j=1}^n (-1)^j \sum_{1 < i_1 < \cdots < i_j \leq n} \frac{1}{\tau^*(s_{i_1}, \ldots, s_{i_j})},$$

which reduces to (10.15) by Lemma 10.1 (iii).

From Lemma 10.6 and the multiplicativity of $\tau(k)$ we deduce immediately from Lemma 10.8 the formula for $\tau(k)$, $k$ even:

**Lemma 10.9.** Let $k$ be an even square-free integer. Then

$$\tau(k) = \frac{1}{2n} \sum_{j=1}^n (-1)^j \sum_{1 < i_1 < \cdots < i_j \leq n} \frac{1}{\tau(s_{i_1}, \ldots, s_{i_j})},$$

$$= \frac{1}{2n} \sum_{j=1}^n (-1)^j \sum_{1 < i_1 < \cdots < i_j \leq n} \frac{1}{\tau^*(s_{i_1}, \ldots, s_{i_j})},$$

which reduces to (10.15) by Lemma 10.1 (iii).

From Lemma 10.6 and the multiplicativity of $\tau(k)$ we deduce immediately from Lemma 10.8 the formula for $\tau(k)$, $k$ even:

**Lemma 10.10.** The series $\sum_{k=1}^\infty \mu(k) \tau(k)$ is absolutely convergent.
Proof. First let \( k \) be an even square-free integer. Then by (10.17) we have

\[
\sigma'(k) = \sigma(k) + O(\sigma(k/2)) = O \left( \frac{d^n(k)}{k^p(k)} \right) + O \left( \frac{d^n(k/2)}{\frac{1}{2}k^p(k/2)} \right) \text{ by (8.9),}
\]

\[
= O \left( \frac{d^n(k)}{k^p(k)} \right).
\]

This relation holds also when \( k \) is odd as \( \sigma'(k) = \sigma(k) \) then holds. Hence \( \sum_{k=1}^{\infty} \mu(k)\sigma'(k) \) is absolutely convergent by comparison with \( \sum_{k=1}^{\infty} \frac{d^n(k)}{k^p(k)} \).

11. An infinite product for \( \sum_{k=1}^{\infty} \mu(k)\sigma'(k) \). The lemma below follows from the multiplicativity of \( \sigma'(k) \) and the absolute convergence of \( \sum_{k=1}^{\infty} \mu(k)\sigma'(k) \).

**Lemma 11.1.** For each prime \( p \) let

\[
\chi(p) = 1 - \sigma'(p) = 1 - \sigma(p).
\]

Then

\[
\sum_{k=1}^{\infty} \mu(k)\sigma'(k) = \prod_{p} \chi(p).
\]

**Lemma 11.2.**

\[
\chi(2) = \frac{1}{2^n} \sum_{a_1, \ldots, a_n = 0}^{1} (1 - \chi(2)) \sum_{a_1 \cdot \ldots \cdot a_n = \alpha} (-1)^{2^n}. \tag{11.3}
\]

Proof. We have \( \chi(2) = 1 - \sigma'(2) \). Hence from (10.16)

\[
\chi(2) = \frac{1}{2^n} \sum_{j=0}^{n} (-1)^j \sum_{1 \leq i_1 < \ldots < i_j \leq n \atop a_i \cdot b_i = a_j} 1
\]

and this is equivalent to (11.3).

**Lemma 11.3.** With \( \chi(p) \) defined by (11.1) we have

\[
\sum_{k=1}^{\infty} \mu(k)\sigma(k) = \frac{1}{2^n} \prod_{p > 1} \chi(p) \sum_{a_1 = 0}^{1} \ldots \sum_{a_n = 0}^{1} (1 - \chi(2)) \sum_{a = (a_1 \cdot \ldots \cdot a_n) = \alpha \text{ (mod } 4)} (-1)^{2^n} f(|a|) \tag{11.4}
\]

where

\[
f(|a|) = \mu(|a|) \prod_{p | a} \frac{\sigma'(p)}{1 - \sigma'(p)}.
\]

Proof. From Lemmas 10.2, 10.9 and 11.1 we have

\[
\sum_{k=1}^{\infty} \mu(k)\sigma(k) = \sum_{k=1}^{\infty} \mu(k)\sigma'(k) - \frac{1}{2^n} \sum_{j=1}^{n} (-1)^j \sum_{a \cdot b = \alpha \text{ (mod } 4) \atop a \neq 1} \sum_{\substack{1 \leq i_1 < \ldots < i_j \leq n \leq \alpha \neq 0}} \mu(\alpha)\sigma'(\alpha) \sum_{k=1}^{\infty} \mu(k)\sigma'(k)
\]

\[
= \prod_{p > 1} \chi(p) \frac{1}{2^n} \sum_{j=1}^{n} (-1)^j \sum_{\substack{1 \leq i_1 < \ldots < i_j \leq n \leq \alpha \neq 0 \atop a \cdot \alpha = \alpha \text{ (mod } 4) \atop a \neq 1}} \mu(\alpha)\sigma'(\alpha) \sum_{k=1}^{\infty} \mu(k)\sigma'(k)
\]

\[
= \frac{1}{2^n} \prod_{p > 1} \chi(p) \left( 2^n \chi(2) + \sum_{j=1}^{n} (-1)^j \sum_{\substack{1 \leq i_1 < \ldots < i_j \leq n \leq \alpha \neq 0 \atop a \cdot \alpha = \alpha \text{ (mod } 4) \atop a \neq 1}} f(|a|) \right)
\]

by Lemma 11.2.

The product \( \prod_{p > 1} \chi(p) \) is positive. For the absolute convergence of \( \sum \sigma(p) \) implies that \( \prod_{p > 1} \chi(p) \) vanishes if and only if a factor \( \chi(p) \) vanishes, \( p > 2 \). However as \( \sigma(p) \) is the natural density of the primes \( q = 1 \text{ (mod } p) \) such that at least one of \( a_1, \ldots, a_n \) is a \( p \text{th power residue mod } q \) (see Lemma 8.1) it follows that

\[
0 \leq \sigma(p) \leq \frac{1}{p-1}
\]

and hence

\[
\chi(p) \geq 1 - \frac{1}{p-1} > 0 \text{ if } p > 2.
\]

Consequently by (11.4) the vanishing of \( \sum_{k=1}^{\infty} \mu(k)\sigma(k) \) is equivalent to the vanishing of the finite sum occurring in (11.4). The reader may
be amused to verify that this sum vanishes (as it must by Section 2) if $C_i$ is false. If $C_i$ holds, some simplification of the finite sum is possible. It is convenient to state

**Lemma 11.4.** Let $S(a_1, \ldots, a_n)$ be the set of $n$-tuples $(e_1, \ldots, e_n)$, $e_i = 0$ or 1, satisfying $a_1^{e_1} \cdots a_n^{e_n} = b^k$, $b \in \mathbb{Z}$. Then

(i) $C_n$ implies that $\chi(2) = \frac{1}{2^n} |S(a_1, \ldots, a_n)|$,

(ii) $C_n$ implies that the number of solutions $(e_1, \ldots, e_n)$, $e_i = 0$ or 1 of $\chi(a_1^{e_1} \cdots a_n^{e_n}) = \chi(a_1^{e_1} \cdots a_n^{e_n})$ is $|S(a_1, \ldots, a_n)|$,

(iii) if $C_n$, then $\chi(a_1^{e_1} \cdots a_n^{e_n}) = \chi(a_1^{e_1} \cdots a_n^{e_n})$ implies that $\sum_{e_i} = \sum_{e_i \equiv (\text{mod} 2)}$.

**Proof.** (i) follows immediately from (11.3); the proofs of (ii) and (iii) are straightforward and are left to the reader.

Let $G(a_1, \ldots, a_n)$ be the set of integers (square-free) of the form $a = \chi(a_1^{e_1} \cdots a_n^{e_n}) = 1 \pmod{4}$, $e_i = 0$ or 1. Then from Lemma 11.4 (iii), assuming $C_n$, the expression $(-1)^{\sum_{e_i}}$ depends on $a$ only; and we may define unambiguously a function $\omega(a)$ on $G(a_1, \ldots, a_n)$ by the formula

$$\omega(a) = (-1)^{\sum_{e_i}}.$$  

(11.7)

It follows from Lemma 11.4 (ii) that

$$\sum_{a \in G(a_1, \ldots, a_n)} (-1)^{\sum_{e_i}} f(|a|) = |S(a_1, \ldots, a_n)| \sum_{a \in G(a_1, \ldots, a_n)} \omega(a) f(|a|) = 2^n \chi(2) \sum_{a \in G(a_1, \ldots, a_n)} \omega(a) f(|a|),$$  

by Lemma 11.4 (i). Hence from (11.4) and (11.8) we have

**Lemma 11.5.** On the assumption that $C_n$ holds

$$\sum_{a \in G(a_1, \ldots, a_n)} \mu(a) c(k) = \prod_{p \nmid a} \sum_{a \in G(a_1, \ldots, a_n)} \omega(a) f(|a|),$$  

where $f(|a|)$ is defined by (11.5) and where $\prod_{p \nmid a} \chi(p) > 0$.

We remark that $G(a_1, \ldots, a_n)$ is closed under the associativity operation $a \otimes b = \chi(ab)$; also $a \otimes a = 1$. Hence (see Ledermann [10], Lemma 2, p. 47) $G(a_1, \ldots, a_n)$ is isomorphic to a group $G \times \cdots \times G$, the direct product of $t$ cyclic groups of order 2. We observe also that

$$\omega(a \otimes b) = \omega(a) \omega(b)$$  

for all $a$ and $b$ in $G(a_1, \ldots, a_n)$.

12. The non-vanishing of $\sum_{a \in G(a_1, \ldots, a_n)} \omega(a) f(|a|)$. For each $a \in G(a_1, \ldots, a_n)$ let

$$h(a) = \omega(a) f(|a|).$$

(12.1)

In this section we prove that $\sum_{a \in G(a_1, \ldots, a_n)} h(a)$ is positive if $C_n$ holds. But first we need some information about $h(a)$.

**Lemma 12.1.** For each $a \in G(a_1, \ldots, a_n)$ we have

(i) $h(a) \leq 1$,

(ii) $C_n$ implies that $h(a) \neq -1$.

**Proof.** (i) follows from the inequalities

$$|h(a)| = \left\{ \begin{array}{ll}
\frac{c(p)}{1 - c(p)} & p \pmod{a} \\
\frac{1}{p - 2} & p \pmod{a}.
\end{array} \right.$$

(12.2)

(See inequalities (11.6).)

To prove (ii) we assume that $h(a) = -1$ for some $a \in G(a_1, \ldots, a_n)$. Then inequalities (12.2) imply that $a = -3$ and $c(3) = \frac{1}{3}$. Also, if $3 f(3) = -1$ we have $\omega(-3) = 1$. Consequently (11.7), $-3 = \chi(a_1^{e_1} \cdots a_n^{e_n})$ where $2|\sum_{e_i}$. Hence $a_1^{e_1} \cdots a_n^{e_n} = -3b^k$, $b \in \mathbb{Z}$ where $2|\sum_{e_i}$; this together with $c(3) = \frac{1}{3}$, implies that $C_n$ is false.

The reader may be amused to prove that if $C_n$ fails to hold, then

$$\sum_{a \in G(a_1, \ldots, a_n)} h(a)$$

vanishes (as it must by Section 2).

Let $a_1, \ldots, a_t$ be a basis for the group $G(a_1, \ldots, a_n)$. Then the elements of $G(a_1, \ldots, a_n)$ are the numbers $a = a_1 \otimes \cdots \otimes a_t$, $e_i = 0$ or 1, and hence by (11.9) we have

$$\sum_{a \in G(a_1, \ldots, a_n)} h(a) = \sum_{a_1 = 0}^{1} \cdots \sum_{a_t = 0}^{1} \omega(a_1) \cdots \omega(a_t) f(|a| \otimes \cdots \otimes |a|).$$

It remains to express $|a_1| \otimes \cdots \otimes |a_t|$ as a square-free integer as follows. Write each $|a_i|$ as a product of square-free integers

$$|a_i| = \prod_{i} A_{i_1}$$

where $A_{i_1}$ runs through all $t$-tuples $(\eta_1, \ldots, \eta_t)$, where $\eta_i = 0$ or 1, with $\eta_i = 1$, and where $A_{[\eta_1, \ldots, \eta_t]}$ is the product (possibly empty) of those primes which divide each $|a_i|$ where $\eta_i = 1$. Then (if $e_i = 0$, $\eta_i = 1$, $\forall i$, $1 \leq i \leq t$, the canonical factorisation of $|a| = |a_1| \otimes \cdots \otimes |a_t|$ is found by replacing each $A_{i_1}$ by 1 in the product

$$\left(\prod_{i} A_{i_1}\right)^{\eta_1} \cdots \left(\prod_{i} A_{i_1}\right)^{\eta_t}.$$
We get a factorisation of the type $a = \prod A_\lambda$ where $\lambda$ runs over certain $2^{l-1}$ $t$-tuples. It is important to notice that the term $A_{\lambda_{i_1}, \ldots, \lambda_{i_t}}^{\lambda_{i_1} \ldots \lambda_{i_t}}$ is present in the above factorisation of $[s]$. For this factorisation we have

$$h(a) = (\omega(a_1))^{\lambda_1} \cdots (\omega(a_t))^{\lambda_t} \prod_i f(A_i) = \prod_i g(A_i)$$

where $g(A_i) = \omega(a_i)f(A_i)$ if $\lambda = \lambda_i = (0, \ldots, 1, \ldots, 0)$ for some $i$, and where $g(A_i) = f(A_i)$ otherwise. We observe that Lemma 12.1 implies that

$$|g(A_i)| \leq 1$$

and that on the assumption that $C_2$ holds

$$\prod g(A_i) \neq -1.$$  

We can now prove

**Lemma 12.2.** If $C_2$ holds then

$$\sum_{a \in \mathbb{Z}^{n_1} \ldots \mathbb{Z}^{n_t}} h(a) > 0.$$  

Our proof depends on the following result kindly supplied by John Campbell.

**Lemma 12.3.** Let $p_i$ be the polynomial

$$p_i = \prod_{k=1}^{t} \left( 1 + \prod_{k} x_k \right)$$

where $k$ runs through all $t$-tuples $(\eta_1, \ldots, \eta_t)$ with $\eta_i = 0$ or 1 and $\eta_i = 1$. A related polynomial $g_i$ is defined by replacing $x_k$ by 1 in all monomials in $p_i$ other than the monomial 1, which are formed by multiplying out the $t$ terms in $p_i$. We write

$$g_i = 1 + m_1 + \ldots + m_{n-1},$$

where $m_1, \ldots, m_{n-1}$ are monomials (each of degree $2^{l-1}$ in fact). Then if all variables $x_k$ take on real values satisfying $|x_k| < 1$, with the added restriction that $m_k \neq -1$ for $j = 1, \ldots, 2^{l-1}$, we have $g_i > 0$.

**Proof.** From the construction of $g_i$ we observe that $p_i = g_i$ if every $x_k = \pm 1$.

(i) Let $l$ be the minimum of $g_i$ when all variables satisfy $|x_k| < 1$. Then $l \geq 0$. For $g_i$ is a linear function of each $x_k$, and $g_i$ attains its minimum for at least one assignment of values $x_k = \pm 1$. But for such values of $x_k$ we have

$$l = g_i = p_i = \prod_{k=1}^{n} \left( 1 + \prod_{k} x_k \right) \geq 0.$$

(ii) We complete the proof by showing that $g_i = 0 \Rightarrow m_j = -1$ for some $j$.

We write for each $x_k$, $g_i = f + gx_k$ where $f$ and $g$ are independent of $x_k$. Then if $g_i = 0$ and $-1 < x_k < 1$, we must have $g = 0$, otherwise $g_i$ could be made to assume negative values, contrary to (i). We then replace each $x_k$ by zero. Hence if $g_i = 0$ we may assume that each $x_k$ is either zero or $\pm 1$. Consequently from (10.5) we have

$$0 = q_i = 1 + m_1 + \ldots + m_{n-1},$$

where each $m_k$ is either zero or $\pm 1$, and hence $m_j = -1$ for at least one $j$.

This completes the proof of Lemma 12.3.

To prove Lemma 12.2 we observe that the earlier remarks of this section show that $\sum_{a \in \mathbb{Z}^{n_1} \ldots \mathbb{Z}^{n_t}} h(a) = q(t)$, where the variables in $g_i$ satisfy $x_k = g(A_i)$. Conditions (12.3) and (12.4) are then precisely those of Lemma 12.3, which in turn gives $g_i > 0$.

**13. The theorems.** On combining (8.16) with Lemmas 11.5 and 12.2 we have the following theorem.

**Theorem 13.1.** Let $a_1, \ldots, a_n$ be non-zero rational integers and assume

(i) that if $a_1^2, \ldots, a_n^2 = b_1, b_2 \in \mathbb{Z}, c_i = 0 \text{ or } 1$, then $b_2 \Sigma c_i$,

(ii) that if $a_1^a, \ldots, a_n^a = -3b_1, b_2 \in \mathbb{Z}, c_i = 0 \text{ or } 1$, and if $b_2 \Sigma c_i$, then $\delta^3 [3]$, the natural density of the primes $q \equiv 1 (\text{mod } 3)$, $q \mid a_1 \ldots a_n$, such that each of $a_1, \ldots, a_n$ is a cubic non-residue mod $q$, must be positive,

(iii) that the Kummer hypothesis holds for each of the fields $\mathbb{Q}(V_1, V_2, \ldots, V_n)$, where $k = \langle l_1, \ldots, l_n \rangle$ is square-free.

Also let $R(p)$ be the natural density of the primes $q \equiv 1 (\text{mod } p)$, $q \mid a_1 \ldots a_n$, such that at least one of $a_1, \ldots, a_n$ is a $p$-th power residue mod $q$, and let $\chi(p) = 1 - c(p)$. Also let $O(a_1, \ldots, a_n)$ denote the set of distinct square-free numbers of the form $a = x(a_1^a \cdots a_n^a) \equiv 1 (\text{mod } 4)$, $c_i = 0$ or 1, and finally let

$$o(a) = (-1)^{2^{l-1}} \text{ and } f(|a|) = \mu(|a|) \prod_{p | \omega(a)} \frac{\phi(p)}{1 - \phi(p)}.$$  

Then the following asymptotic formula holds for $N_{a_1, \ldots, a_n}(x)$, the number of primes $p \leq x$ such that each of $a_1, \ldots, a_n$ is a primitive root mod $p$:

$$N_{a_1, \ldots, a_n}(x) = \frac{\omega}{\log x} \sum_{a \in O(a_1, \ldots, a_n)} o(a)f(|a|) + O \left( \frac{x}{\log x} (\log \log x)^{n-1} \right),$$  

as $x \to \infty$, and the coefficient of $\frac{x}{\log x}$ is positive.
Theorem 13.2. Let \( a_1, \ldots, a_n \) be relatively prime in pairs and not \( \pm 1 \) or a perfect square. We also assume

(i) that if \( a_i^2 \cdot a_n = -3b^2, b \in \mathbb{Z}, \) \( e_i = 0 \) or 1, and 2\( \prod \sum e_i, \) then none of \( a_1, \ldots, a_n \) is a perfect cube,

(ii) that the Riemann hypothesis holds for each of the fields \( \mathbb{Q}(\sqrt{a_1}, \sqrt{a_2}, \ldots, \sqrt{a_n}) \) where \( k = \langle l_1, \ldots, l_n \rangle \) is square-free.

Then the conclusions of Theorem 13.1 remain valid.

Proof. Condition (i) of Theorem 13.1 is satisfied. For under the assumption that \( a_1, \ldots, a_n \) are relatively prime in pairs and not perfect squares, an equation \( a_1^2 \cdot a_n = b^2, b \in \mathbb{Z}, e_i = 0 \) or 1, implies that \( a_i = -a_i \) if \( e_i = 1 \). Hence an equation of the form

\[
(-1)^{2e_i} a_i^2 = b^2
\]

results, and consequently \( 2\prod \sum e_i \).

Condition (ii) of Theorem 13.1 is also satisfied as is evidenced from the case \( p = 3 \) of the following

Lemma 13.3. Suppose that \( a_1, \ldots, a_n \) are relatively prime in pairs and that none of \( a_1, \ldots, a_n \) is a \( p \)-th power, \( p \) an odd prime. Then

\[
\delta'(p) = \frac{1}{p-1} \left(1 - \frac{1}{p}\right)^n.
\]

Proof. From (10.4) we have

\[
\delta'(p) = \frac{1}{p-1} - \sigma(p) = \frac{1}{p^n(p-1)} \sum_{j=0}^{n} (-1)^j p^{n-j} \sigma_j,
\]

where

\[
\sigma_j = \sum_{i_1 < i_2 < \ldots < i_j < n} \tau(i_1, \ldots, i_j)
\]

and

\[
\tau(i_1, \ldots, i_j) = \sum_{a_1^{i_1} \cdots a_n^{i_n} = b^2, b \in \mathbb{Z}} \tau(i_1, \ldots, i_j).
\]

But as \( a_1, \ldots, a_n \) are relatively prime in pairs, we have

\[
\tau(i_1, \ldots, i_j) = \prod_{i=1}^{j} \tau(i_i)
\]

and from (13.2),

\[
\delta'(p) = \frac{1}{p^n(p-1)} \prod_{i=1}^{n} (p - \tau(i_i)).
\]

However, as none of \( a_1, \ldots, a_n \) is a \( p \)-th power, we have \( \tau(i_i) = 1 \) and hence

\[
\delta'(p) = \frac{1}{p^n(p-1)} = \frac{1}{p-1} \left(1 - \frac{1}{p}\right)^n.
\]

Finally we state

Theorem 13.4. Let \( a_1 \) and \( a_2 \) be non-zero rational integers not \( \pm 1 \) or perfect squares, and assume

(i) that if \( a_1 a_2 = -3b^2, b \in \mathbb{Z}, \) then neither \( a_1 \) nor \( a_2 \) is a perfect cube,

(ii) that the Riemann hypothesis holds for each of the fields \( \mathbb{Q}(\sqrt{a_1}, \sqrt{a_2}, \ldots, \sqrt{a_n}) \) where \( k = \langle l_1, l_2 \rangle \) is square-free.

Then the conclusions of Theorem 13.1 (with \( n = 2 \)) remain valid. There is a similar theorem when \( n = 3 \).

Both results depend on the case \( p = 3 \) of the following result:

Lemma 13.5. (i) If \( n = 2 \) and neither \( a_1 \) nor \( a_2 \) is a \( p \)-th power, then

\[
\delta'(p) > 0.
\]

(ii) If \( n = 3 \) and none of \( a_1, a_2, \) or \( a_3 \) is a \( p \)-th power, \( p \) odd, then

\[
\delta'(p) > 0.
\]

Proof. We use (13.2).

(i) follows from

\[
p^3(p-1) \delta'(p) = \sum_{j=0}^{n} (-1)^j p^{3-j} \sigma_j = p^3 - 2p + \tau(1, 2) \geq (p-1)^2.
\]

(ii) follows from

\[
p^3(p-1) \delta'(p) = \sum_{j=0}^{n} (-1)^j p^{3-j} \sigma_j
\]

\[
= p^3 - 3p^2 + (p \tau(1, 2) + \tau(1, 2, 3)) + p \tau(2, 3) + \tau(1, 3))
\]

\[
\geq p^3 - 3p^2 + 2p = p(p-2)(p-1).
\]

References

Conjugate algebraic numbers on circles

by

Veikko Ennola and C. J. Smyth (Turku)

1. Introduction. In 1969, R. M. Robinson [4] posed the following question:

(I) Which circles \(|x - \gamma| = \mathcal{E}\) contain infinitely many sets of conjugate algebraic integers?

In order to answer this question, we have asked, more generally:

(II) Which algebraic numbers have all their conjugates lying on a circle?

In this paper we give a complete answer to the second question (Theorems 2 and 3). We also find all circles which contain infinitely many sets of conjugate algebraic numbers. This enables us to show, towards answering question (I), that the following holds:

**Theorem 1.** For every \(n \geq 1\) there are algebraic numbers \(\gamma\) of degree \(n\) such that there is a circle of centre \(\gamma\) containing infinitely many sets of conjugate algebraic integers.

There is a method which should, in principle, enable one, from Theorem 3, to give a complete answer to (I), but so far we have only worked out the details when \(\gamma\) is of degree at most 4.

Previous partial answers to (I) and (II) have been as follows: Robinson [4] answered (I), under the assumption that \(\gamma\) is rational. Question (II) is very easy when the centre \(\gamma\) is rational — see [3], Theorem 3. In [1] the first author answered both (I) and (II) when \(\gamma\) is totally real, and in [2] we did the same for \(\gamma\) not totally real and of degree 3 or 4.

When considering (I) and (II), we can, because of the above results, consider only circles with irrational centre. Hence, since any rational or quadratic \(\beta\) lies, with its other conjugate (if any), on a circle of rational centre (of course they lie on many circles), such \(\beta\) can be excluded from consideration in answering question (II). Further, these \(\beta\) are clearly of no interest to question (I). We can therefore confine our attention to the set \(\mathcal{S}\) of all algebraic numbers \(\beta\), of degree at least 3 over the rationals \(\mathbb{Q}\), whose conjugates (including \(\beta\)) all lie on a circle with irrational centre \(\gamma(\beta)\). It is easy to see that \(\gamma(\beta)\) must be a real algebraic number.