ACTA ARITHMETICA
XXIX (1976)

A generalisation of Artin's conjecture for primitive roots

by

K. R. Matthews (Brisbane, Australia)

1. 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 $p \leqslant x$ such that a is a primitive root $\bmod p$, satisfies an asymptotic formula of the form

$$N_a(x) \sim A(a)\frac{x}{\log x} \quad (x \to \infty) \tag{1.1}$$

for some positive constant $A(a)$.

In 1967 Hooley (see [6]) showed that the truth of the Riemann hypothesis for the fields $Q(\sqrt[k]{1}, \sqrt[k]{a})$, k square-free, would imply the truth of Artin's conjecture.

Let $a_1, \ldots, a_n$ be non-zero integers not ± 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 $p \leqslant x$ such that each of $a_1, \ldots, a_n$ is a primitive root $\bmod p$.

An equation

$$N_{a_1,\ldots,a_n}(x) = \frac{x}{\log x} A(a_1, \ldots, a_n) + O\left(\frac{x}{\log^2 x}(\log\log x)^{2^n-1}\right) \tag{1.2}$$

is obtained subject to the truth of the Riemann hypothesis for each of the fields $Q(\sqrt[k]{1}, \sqrt[l_1]{a_1}, \ldots, \sqrt[l_n]{a_n})$ where $k = \langle l_1, \ldots, l_n \rangle$ (the l.c.m. of $l_1, \ldots, l_n$) is square-free. Here

$$A(a_1, \ldots, a_n) = \sum_{k=1}^{\infty} \mu(k) c(k) \tag{1.3}$$

where $c(k)$ is the natural density of the primes $q \equiv 1 (\bmod k)$, $q \nmid a_1 \ldots a_n$, such that for each prime $p \mid k$, at least one of $a_1, \ldots, a_n$ is a pth power residue $\bmod q$.

The constant $A(a_1, \ldots, a_n)$ is then converted to an infinite product, namely

$$(1.4)\quad A(a_1, \ldots, a_n) = \prod_{p>2} (1-c(p)) \sum_{\varepsilon_1=0}^{1} \cdots \sum_{\varepsilon_n=0}^{1} (-1)^{\sum \varepsilon_i} f(|a|)$$
$$a=\varkappa(a_1^{\varepsilon_1}\ldots a_n^{\varepsilon_n})\equiv 1 \pmod 4$$

where

$$(1.5)\quad f(|a|) = \mu(|a|) \prod_{p\mid |a|} \frac{c(p)}{1-c(p)}.$$

(Here $\varkappa(b)$ 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:

$\mathbf{C}_1$: If $a_1^{\varepsilon_1} \ldots a_n^{\varepsilon_n} = b^2$, $b \in \mathbf{Z}$, $\varepsilon_i = 0$ or 1, then $2 \mid \sum \varepsilon_i$.

$\mathbf{C}_2$: If $a_1^{\varepsilon_1} \ldots a_n^{\varepsilon_n} = -3b^2$, $b \in \mathbf{Z}$, $\varepsilon_i = 0$ or 1 and if $2 \mid \sum \varepsilon_i$, then $d'(3)$, the natural density of the primes $q \equiv 1 \pmod 3$, $q \nmid a_1 \ldots a_n$, such that each of $a_1, \ldots, a_n$ is a cubic non-residue mod q, must be positive.

An explicit formula is available for $d'(p)$ $\left(= \dfrac{1}{p-1} - c(p)\right)$ where p is an odd prime. By applying the exclusion principle to Lemma 1 of Schinzel [10], p. 162, we have

$$(1.6)\quad d'(p) = \frac{1}{p^n(p-1)} \sum_{j=0}^{n} (-1)^j p^{n-j} \sigma_j$$

where

$$(1.7)\quad \sigma_j = \sum_{1\leqslant i_1<\ldots<i_j\leqslant n} \sum_{\substack{r_{i_1}=1}}^{p} \cdots \sum_{r_{i_j}=1}^{p} 1$$
$$a_{i_1}^{r_{i_1}}\ldots a_{i_j}^{r_{i_j}} = b^p,\ b\in\mathbf{Z}$$

and $\sigma_0 = 1$.

If $n = 1$, 2 or 3, or if $a_1, \ldots, a_n$ are relatively prime in pairs, an examination of formula (1.6) allows us to replace the condition "$d'(3) > 0$" in $\mathbf{C}_2$ by the statement "none of $a_1, \ldots, a_n$ is a perfect cube". However if $n > 3$ the situation is more complicated.

If $\mathscr{P}$ 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 $\mathbf{C}_1$ and $\mathbf{C}_2$ are each necessary for $\mathscr{P}$ to be infinite.

Finally I would like to acknowledge by 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 $\mathscr{P}$ to be infinite. We prove that the falsity of $\mathbf{C}_1$ or $\mathbf{C}_2$ implies that $\mathscr{P}$ contains at most the element 2. For suppose that $\mathbf{C}_1$ is false. Then we have

$$a_{i_1} \ldots a_{i_j} = b^2, \quad b \in \mathbf{Z}, \quad 1 \leqslant i_1 < \ldots < i_j \leqslant n$$

and j odd. Now if $p \in \mathscr{P}$ 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:

$$1 = \left(\frac{b^2}{p}\right) = \left(\frac{a_{i_1} \ldots a_{i_j}}{p}\right) = \left(\frac{a_{i_1}}{p}\right) \ldots \left(\frac{a_{i_j}}{p}\right) = (-1)^j = -1.$$

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

Now suppose that $\mathbf{C}_2$ is false. Then we have

$$a_{i_1} \ldots a_{i_j} = -3b^2, \quad b \in \mathbf{Z}, \quad 1 \leqslant i_1 < \ldots < i_j \leqslant n$$

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

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

But quadratic reciprocity gives

$$\left(\frac{3}{p}\right) = \left(\frac{p}{3}\right)(-1)^{\frac{p-1}{2}}$$

and consequently $\left(\dfrac{p}{3}\right) = 1$. Hence $p \equiv 1 \pmod 3$. Now as each of $a_1, \ldots, a_n$ is a primitive root mod p, in view of the congruence $p \equiv 1 \pmod 3$ just derived, each of $a_1, \ldots, a_n$ is a cubic non-residue mod p. Hence the condition $d'(3) = 0$ implies that $\mathscr{P}$ 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 $p \in \mathscr{P}$.

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 $\mathfrak{p}$ which

have prescribed pth power residue symbol values at each of n given non-zero rational integers.

DEFINITION 2.1. Let p be a rational prime, α an algebraic integer of $\boldsymbol{Q}(\sqrt[p]{1})$ and $\mathfrak{p}$ a prime ideal of $\boldsymbol{Q}(\sqrt[p]{1})$, $\mathfrak{p}\nmid[p\alpha]$. Then the pth power residue symbol is defined by

$$\left(\frac{\alpha}{\mathfrak{p}}\right)_p \equiv \alpha^{\frac{1}{p}[N(\mathfrak{p})-1]} (\mathrm{mod}\,\mathfrak{p}), \qquad \left(\frac{\alpha}{\mathfrak{p}}\right)_p^p = 1. \tag{2.1}$$

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

We note that

$$\left(\frac{\alpha}{\mathfrak{p}}\right) = 1 \Leftrightarrow \alpha \equiv \beta^p (\mathrm{mod}\,\mathfrak{p}), \qquad \beta \in \boldsymbol{Q}(\sqrt[p]{1}). \tag{2.2}$$

We also have the equation

$$\left(\frac{\alpha_1\alpha_2}{\mathfrak{p}}\right)_p = \left(\frac{\alpha_1}{\mathfrak{p}}\right)_p \left(\frac{\alpha_2}{\mathfrak{p}}\right)_p \tag{2.3}$$

if $\mathfrak{p}\nmid[p\alpha_1\alpha_2]$.

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*

$$N(p,n) = \sum_{\substack{\nu_1=1}}^{p} \cdots \sum_{\substack{\nu_n=1 \\ a_1^{\nu_1}\ldots a_n^{\nu_n}=\beta^p,\ \beta\in\boldsymbol{Q}(\sqrt[p]{1})}}^{p} (\varepsilon_1^{\nu_1}\ldots\varepsilon_n^{\nu_n})^{-1}. \tag{2.4}$$

Let $S(x, p, n)$ $(= S(x, p, n;\ \varepsilon_1, \ldots, \varepsilon_n))$ be the number of prime ideals $\mathfrak{p}$ of the first degree which satisfy $N(\mathfrak{p}) \leqslant x$, and for which the relations

$$\left(\frac{a_j}{\mathfrak{p}}\right)_p = \varepsilon_j \quad (j = 1, \ldots, n) \tag{2.5}$$

are satisfied. Then as $x \to \infty$

$$S(x,p,n) = p^{-n}N(p,n)\pi(x) + O\left(x\exp\left(-A\sqrt{\log x}\right)\right) \tag{2.6}$$

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 $O(\sqrt{x})$.)

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

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

$$N(p,n) = \sum_{\nu_1=1}^{p} \cdots \sum_{\substack{\nu_n=1 \\ a_1^{\nu_1}\ldots a_n^{\nu_n}=\beta^p,\ \beta\in\boldsymbol{Q}(\sqrt[p]{1})}}^{p} 1 \tag{2.7}$$

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

Proof. (i) Let $\mathfrak{p}$ be a prime ideal satisfying conditions (2.5). Then

$$\varepsilon_1^{\nu_1}\ldots\varepsilon_n^{\nu_n} = \left(\frac{a_1}{\mathfrak{p}}\right)_p^{\nu_1} \cdots \left(\frac{a_n}{\mathfrak{p}}\right)_p^{\nu_n} = \left(\frac{a_1^{\nu_1}\ldots a_n^{\nu_n}}{\mathfrak{p}}\right)_p,$$

by (2.3). Hence

$$N(p,n) = \sum_{\substack{\nu_1=1}}^{p} \cdots \sum_{\substack{\nu_n=1 \\ a_1^{\nu_1}\ldots a_n^{\nu_n}=\beta^p,\ \beta\in\boldsymbol{Q}(\sqrt[p]{1})}}^{p} \left(\frac{a_1^{\nu_1}\ldots a_n^{\nu_n}}{\mathfrak{p}}\right)^{-1} = \sum_{\substack{\nu_1=1}}^{p} \cdots \sum_{\substack{\nu_n=1 \\ a_1^{\nu_1}\ldots a_n^{\nu_n}=\beta^p,\ \beta\in\boldsymbol{Q}(\sqrt[p]{1})}}^{p} 1,$$

by (2.2).

(ii) If no prime ideal $\mathfrak{p}$ exists satisfying conditions (2.5) then $S(x, p, n) = 0$ for each x, 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, $q \equiv 1 (\mathrm{mod}\, p)$, $q \nmid a$. Also let $\mathfrak{p}$ be a prime ideal in $\boldsymbol{Q}(\sqrt[p]{1})$, $\mathfrak{p}\mid[q]$. Then the congruence*

$$a \equiv \beta^p (\mathrm{mod}\,\mathfrak{p}), \qquad \beta \in \boldsymbol{Q}(\sqrt[p]{1})$$

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

Proof. The assumption $\mathfrak{p}\mid[q]$ and $q \equiv 1 (\mathrm{mod}\, p)$ implies that $\mathfrak{p}$ is of the first degree. Hence the integers of $\boldsymbol{Q}(\sqrt[p]{1})\,\mathrm{mod}\,\mathfrak{p}$ form a field of q $(= N(\mathfrak{p}))$ elements. Hence if $a \equiv \beta^p (\mathrm{mod}\,\mathfrak{p})$ we have

$$a^{\frac{q-1}{p}} \equiv \beta^{q-1} \equiv 1 (\mathrm{mod}\,\mathfrak{p}),$$

and as $q \in \mathfrak{p}$, it follows that

$$a^{\frac{q-1}{p}} \equiv 1 (\mathrm{mod}\, q).$$

Consequently a is a pth 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 $q \equiv 1 (\mathrm{mod}\, p)$, $q \nmid a_1 \ldots a_n$*

and such that each of $a_1, \dots, a_n$ is a p-th power non-residue mod q. *Then the natural density $d'(p)$, of the set $\mathscr{P}_1$ of such primes q, is positive.*

Proof. The prime ideals in $\boldsymbol{Q}(\sqrt[p]{1})$ of the first degree and not dividing $[pa_1 \dots a_n]$ are grouped naturally together in sets of $p-1$ by the equation

$$[q] = \mathfrak{p}_1 \dots \mathfrak{p}_{p-1}$$

where $q = N(\mathfrak{p}_i) \equiv 1 (\mathrm{mod}\, p)$.

By Lemma 2.3 and (2.2) the conditions

$$\left(\frac{a_j}{\mathfrak{p}}\right)_p \neq 1 \quad (j = 1, \dots, n) \tag{2.8}$$

are either satisfied for all $\mathfrak{p} = \mathfrak{p}_i$ $(i = 1, \dots, p-1)$ or for no $\mathfrak{p}_i$. Hence

$$\sum_{\substack{q \leqslant x \\ q \in \mathscr{P}_1}} 1 = \frac{1}{p-1} \sum\nolimits' 1$$

where $\sum'$ denotes a summation over all prime ideals $\mathfrak{p}$ of $\boldsymbol{Q}(\sqrt[p]{1})$ of the first degree, with $N(\mathfrak{p}) \leqslant x$ and satisfying the conditions (2.8). Hence

$$\sum_{\substack{q \leqslant x \\ q \in \mathscr{P}_1}} 1 = \frac{1}{p-1} \sum_{\varepsilon_1, \dots, \varepsilon_n} S(x, p, n; \varepsilon_1, \dots, \varepsilon_n)$$

where $\varepsilon_1, \dots, \varepsilon_n$ run over all pth roots of unity other than 1. The asymptotic formula (2.6) now gives

$$d'(p) = \lim_{x \to \infty} \frac{1}{\pi(x)} \sum_{\substack{q \leqslant x \\ q \in \mathscr{P}_1}} 1 = \frac{1}{p^n(p-1)} \sum_{\varepsilon_1, \dots, \varepsilon_n} N(p, n; \varepsilon_1, \dots, \varepsilon_n). \tag{2.9}$$

If q_0 is a prime satisfying the hypothesis of Lemma 2.4 and if $\mathfrak{p}_0$ is a prime ideal, $\mathfrak{p}_0 | [q_0]$, then by Lemma 2.3 and (2.2) the conditions

$$\left(\frac{a_j}{\mathfrak{p}_0}\right)_p \neq 1 \quad (j = 1, \dots, n)$$

are satisfied and if $\varepsilon_j' = \left(\frac{a_j}{\mathfrak{p}_0}\right)_p$, (2.9) gives

$$d'(p) \geqslant \frac{1}{p^n(p-1)} N(p, n; \varepsilon_1', \dots, \varepsilon_n')$$

$$> 0, \quad \text{by Lemma 2.2.}$$

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 $p \nmid a$ and for each prime $q | p-1$, a is not a qth power residue mod p. Consequently if $\mathrm{R}(q, p)$ denotes the statement

(3.1) $q | p-1$ and at least one of $a_1, \dots, a_n$ is a qth power residue mod p,

it follows that $N(x)$ $(= N_{a_1, \dots, a_n}(x))$ is the number of primes $p \leqslant x$, $p \nmid a_1 \dots a_n$, such that $\mathrm{R}(q, p)$ is false for all primes q.

Let $N(x, \eta)$ be the number of primes $p \leqslant x$, $p \nmid a_1 \dots a_n$, such that $\mathrm{R}(q, p)$ is false for all primes $q \leqslant \eta$. Then

$$N(x) = N(x, x-1).$$

We let $P(x, k)$ be the number of primes $p \leqslant x$, $p \nmid a_1 \dots a_n$, such that $\mathrm{R}(q, p)$ is true for all primes $q | k$. Then by the exclusion principle

$$N(x, \eta) = \sum\nolimits'_k \mu(k) P(x, k) \tag{3.2}$$

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

Let $\xi_1 = \frac{1}{6} \log x$, $\xi_2 = x^{1/2} \log^{-2} x$, $\xi_3 = x^{1/2} \log x$.

If $\eta_1 < \eta_2$ we let $M(x, \eta_1, \eta_2)$ be the number of primes $p \leqslant x$, $p \nmid a_1 \dots a_n$, such that $\mathrm{R}(q, p)$ is true for at least one prime q, $\eta_1 < q \leqslant \eta_2$. Then with only slight changes to Hooley's argument, the fundamental equation

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

of Hooley follows.

We recall from (3.2) that

$$N(x, \xi_1) = \sum\nolimits'_k \mu(k) P(x, k) \tag{3.4}$$

and note that

$$k \leqslant \prod_{q \leqslant \xi_1} q = e^{\sum_{q \leqslant \xi_1} \log q} \leqslant e^{2\xi_1} = x^{1/3}. \tag{3.5}$$

We also observe that

$$M(x, \xi_1, \xi_2) \leqslant \sum_{\xi_1 < q \leqslant \xi_2} P(x, q). \tag{3.6}$$

4. A formula for $P(x, k)$. We recall from Section 2 that $P(x, k)$ counts the primes $p \leqslant x$, $p \nmid a_1 \dots a_n$, such that for all primes $q | k$ we have $p \equiv 1 (\mathrm{mod}\, q)$ and at least one of $a_1, \dots, a_n$ is a qth power residue mod p. Thus $P(x, k)$ counts the primes $p \leqslant x$, $p \nmid a_1 \dots a_n$, $p \equiv 1 (\mathrm{mod}\, k)$, such that for all primes $q | k$ at least one of $a_1, \dots, a_n$ is a qth power residue mod p.

If $l_1, \dots, l_n$ are divisors of k (square-free) we let $P(x, l_1, \dots, l_n; k)$ denote the number of primes $p \leqslant x$, $p \nmid a_1 \dots a_n$, $p \equiv 1 \pmod k$, such that each of the congruences

$$a_1 \equiv x_1^{l_1} \pmod p, \ \dots, \ a_n \equiv x_n^{l_n} \pmod p$$

is soluble. Then we have the following formula for $P(x, k)$.

LEMMA 4.1.

$$P(x, k) = \mu(k) \sum_{\substack{l_1 \mid k \\ \langle l_1, \dots, l_n \rangle = k}} \dots \sum_{l_n \mid k} \mu(l_1) \dots \mu(l_n) P(x, l_1, \dots, l_n; k). \tag{4.1}$$

Proof. (This is due to Prof. Burgess.) For each i, $1 \leqslant i \leqslant n$, a multiplicative function $f_i(l)$ is defined by

$$f_i(l) = \begin{cases} 1 & \text{if } \ x^l \equiv a_i \pmod p \text{ is soluble,} \\ 0 & \text{otherwise.} \end{cases}$$

Hence if $f(q)$ is defined by

$$f(q) = 1 - \prod_{i=1}^{n} (1 - f_i(q)),$$

we have

$$f(q) = \begin{cases} 1 & \text{if at least one of } x_i^q \equiv a_i \pmod p, \\ & \qquad 1 \leqslant i \leqslant n, \text{ is soluble,} \\ 0 & \text{otherwise.} \end{cases}$$

Hence

$$P(x, k) = \sum_{\substack{p \leqslant x \\ p \equiv 1 \pmod k \\ p \nmid a_1 \dots a_n}} \prod_{q \mid k} f(q). \tag{4.2}$$

But

$$\begin{aligned} \prod_{q \mid k} f(q) &= \prod_{q \mid k} \Big(1 - \prod_{i=1}^{n} (1 - f_i(q))\Big) = \sum_{d \mid k} \mu(d) \prod_{q \mid d} \prod_{i=1}^{n} (1 - f_i(q)) \\ &= \sum_{d \mid k} \mu(d) \prod_{i=1}^{n} \prod_{q \mid d} (1 - f_i(q)) = \sum_{d \mid k} \mu(d) \prod_{i=1}^{n} \sum_{l_i \mid d} \mu(l_i) f_i(l_i) \\ &= \sum_{d \mid k} \mu(d) \sum_{l_1 \mid d} \dots \sum_{l_n \mid d} \mu(l_1) \dots \mu(l_n) \prod_{i=1}^{n} f_i(l_i). \end{aligned}$$

Hence from (4.2)

$$\begin{aligned} P(x, k) &= \sum_{d \mid k} \mu(d) \sum_{l_1 \mid d} \dots \sum_{l_n \mid d} \mu(l_1) \dots \mu(l_n) \sum_{\substack{p \leqslant x \\ p \equiv 1 \pmod k \\ p \nmid a_1 \dots a_n}} \prod_{i=1}^{n} f_i(l_i) \\ &= \sum_{d \mid k} \mu(d) \sum_{l_1 \mid d} \dots \sum_{l_n \mid d} \mu(l_1) \dots \mu(l_n) P(x, l_1, \dots, l_n; k) \\ &= \sum_{l_1 \mid k} \dots \sum_{l_n \mid k} \mu(l_1) \dots \mu(l_n) P(x, l_1, \dots, l_n; k) \sum_{\substack{d \mid k \\ \langle l_1, \dots, l_n \rangle \mid d}} \mu(d). \end{aligned} \tag{4.3}$$

The inner sum of (4.3) simplifies on making the substitution $d = t\langle l_1, \dots, l_n \rangle$. For then

$$\begin{aligned} \sum_{\substack{d \mid k \\ \langle l_1, \dots, l_n \rangle \mid d}} \mu(d) &= \sum_{t \mid k / \langle l_1, \dots, l_n \rangle} \mu(t \langle l_1, \dots, l_n \rangle) = \mu(\langle l_1, \dots, l_n \rangle) \sum_{t \mid k / \langle l_1, \dots, l_n \rangle} \mu(t) \\ &= \begin{cases} \mu(k) & \text{if } \ \langle l_1, \dots, l_n \rangle = k, \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

Consequently (4.3) reduces to (4.1).

5. An asymptotic formula for $P(x, l_1, \dots, l_n; k)$. We recall that $P(x, l_1, \dots, l_n; k)$ counts those primes $p \leqslant x$ which satisfy each of the conditions

$$p \equiv 1 \pmod k, \quad p \nmid a_i, \quad a_i \equiv x_i^{l_i} \pmod p \text{ soluble,} \tag{5.1}$$

where $1 \leqslant i \leqslant n$.

The argument of Hooley [6], Section 4, pp. 212–213, shows that (5.1) is equivalent to the statement

(5.2) $p \nmid k a_i$ and p factorises as a product of first degree prime ideals in $\boldsymbol{Q}(\sqrt[k]{1}, \sqrt[l_i]{a_i})$.

Hence $P(x, l_1, \dots, l_n; k)$ counts the primes $p \leqslant x$ which satisfy the condition

(5.3) $p \nmid k a_1 \dots a_n$ and p factorises as a product of first degree prime ideals in $K = \boldsymbol{Q}(\sqrt[k]{1}, \sqrt[l_1]{a_1}, \dots, \sqrt[l_n]{a_n})$.

We remark that such prime ideals are distinct, for $p \nmid k a_1 \dots a_n$ implies that $p \nmid \Delta(K)$, the discriminant of K. (The primes dividing $\Delta(K_1 K_2)$ are those which divide either $\Delta(K_1)$ or $\Delta(K_2)$; also $\Delta\big(\boldsymbol{Q}(\sqrt[k]{1}, \sqrt[l_i]{a_i})\big)$ is formed from primes which divide $k a_i$ – 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 $\mathfrak{p}$ of K with $N(\mathfrak{p}) \leqslant x$. Then

$$\pi(x, K) = \pi_1(x, K) + \pi_2(x, K) \tag{5.4}$$

where $\pi_1(x, K)$ is the contribution to $\pi(x, K)$ from the first degree prime ideals not dividing $k a_1 \dots a_n$, and $\pi_2(x, K)$ is the remaining contribution.

As K is a Galois extension of $\boldsymbol{Q}$, we know that each prime $p \nmid k a_1 \dots a_n$ has either $N(K)$ $(= [K : \boldsymbol{Q}])$ 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)

$$\pi_1(x, K) = N(K) P(x, l_1, \dots, l_n; k) \tag{5.5}$$

and

$$\pi_2(x, K) \leqslant N(K)\,\omega(ka_1 \dots a_n) + N(K) \sum_{p^2 \leqslant x} 1. \tag{5.6}$$

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

$$N(K)P(x, l_1, \dots, l_n; k) = \pi(x, K) + O\big(N(K)\,\omega(k)\big) + O\big(N(K)x^{1/2}\big). \tag{5.7}$$

6. A recursive formula for $[F(\sqrt[l_1]{a_1}, \dots, \sqrt[l_n]{a_n}):F]$. We shall eventually need a lower estimate for $N(K) = [K : \boldsymbol{Q}]$ where $K = \boldsymbol{Q}(\sqrt[k]{1}, \sqrt[l_1]{a_1}, \dots, \sqrt[l_n]{a_n})$, $a_1, \dots, a_n$ are non-zero rational integers and $k = \langle l_1, \dots, l_n \rangle$.

Now $N(K) = [K : \boldsymbol{Q}(\sqrt[k]{1})]\varphi(k)$, and hence it suffices to investigate $[F(\sqrt[l_1]{a_1}, \dots, \sqrt[l_n]{a_n}):F]$ where F is a number field containing all kth roots of unity.

LEMMA 6.1. *Let F be a number field containing all k-th roots of unity, $a_1, \dots, a_n$ are non-zero elements of F and $l_1, \dots, l_n$ are divisors of k, a square-free integer. Positive integers $\lambda_1, \dots, \lambda_n, \lambda_1', \dots, \lambda_n'$ are defined as follows: λ_1' is the product of those primes $p \mid l_1$ such that $a_1 = \beta^p$, $\beta \in F$; $\lambda_1 = l_1/\lambda_1'$.*

If $1 < r \leqslant n$, λ_r' is the product of those primes $p \mid l_r$ for which integers $m_1, \dots, m_{r-1}$ exist satisfying

$$a_r = a_1^{m_1} \dots a_{r-1}^{m_{r-1}} \beta^p, \qquad \beta \in F,$$

where in addition, $p \nmid m_i \Rightarrow p \mid \lambda_i$, for $1 \leqslant i \leqslant r-1$; also $\lambda_r = l_r \mid \lambda_r'$. Then

$$[F(\sqrt[l_1]{a_1}, \dots, \sqrt[l_n]{a_n}):F] = \lambda_1 \dots \lambda_n. \tag{6.1}$$

Proof. Let $J_0 = K_0 = F$ and for $1 \leqslant r \leqslant n$ let

$$J_r = F(\sqrt[l_1]{a_1}, \dots, \sqrt[l_r]{a_r}), \qquad K_r = F(\sqrt[\lambda_1]{a_1}, \dots, \sqrt[\lambda_r]{a_r}).$$

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

(i) $J_r = K_r$ for $0 \leqslant r \leqslant n$,

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

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

$$[K_r : K_{r-1}] = \lambda_r, \qquad 1 \leqslant r \leqslant n$$

and hence

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

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

$$J_r = J_{r-1}(\sqrt[l_r]{a_r}) = K_{r-1}(\sqrt[l_r]{a_r}). \tag{6.2}$$

We now write $l_r = \lambda_r p_1 \dots p_t$, where $p_1, \dots, p_t$ are the prime factors of λ_r'. Then

$$K_{r-1}(\sqrt[l_r]{a_r}) = K_{r-1}(\sqrt[p_1]{a_r}, \dots, \sqrt[p_t]{a_r}, \sqrt[\lambda_r]{a_r}). \tag{6.3}$$

But from construction of λ_r, we have for $p = p_1, \dots, p_t$

$$\sqrt[p]{a_r} = a_1^{m_1/p} \dots a_{r-1}^{m_{r-1}/p} \beta,$$

where $\beta \in F$ and $p \nmid m_i \Rightarrow p \mid \lambda_i$, $1 \leqslant i \leqslant r-1$. Hence for such p we have $\sqrt[p]{a_r} \in K_{r-1}$ and hence $K_{r-1}(\sqrt[p]{a_r}) = K_{r-1}$. From (6.3) it follows that

$$K_{r-1}(\sqrt[l_r]{a_r}) = K_{r-1}(\sqrt[\lambda_r]{a_r}) = K_r,$$

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 Satz 150 of Hasse [5], pp. 220–221.

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 $x^\lambda - a$ is irreducible over F, where $\lambda \mid k$. Then if β is a non-zero element of F with $\beta = \gamma^p$, $\gamma \in F(\sqrt[\lambda]{a})$ and $p \mid k$, we have $\beta = \alpha^p \delta^p$, where $\delta \in F$ and $p \nmid v \Rightarrow p \mid \lambda$.*

Proof of (ii). For $1 \leqslant r \leqslant n$ let P_r denote the statement

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

(b) if β is a non-zero element of F such that $\beta = \gamma^p$, $\gamma \in K_r$, then $\beta = a_1^{m_1} \dots a_r^{m_r} \delta^p$, where $\delta \in F$ and $p \nmid m_i \Rightarrow p \mid \lambda_i$, for $1 \leqslant i \leqslant r$.

We use a well-known criterion for the irreducibility of $x^\lambda - a$ over a field H, stated for example in Lang [9], Theorem 16, p. 221. For square-free λ this states that $x^\lambda - a$ is irreducible over H if $p \mid \lambda \Rightarrow a \neq \beta^p$, $\beta \in H$. (H is assumed to be of characteristic zero or prime to λ.)

We prove by induction on r that P_r holds for $1 \leqslant r \leqslant n$. Our proof is based on that of Elliott [3], Lemma 3, pp. 134–135. When $r = 1$, (a) follows from the construction of λ_1 and the above mentioned criterion for irreducibility, while (b) reduces to the statement of Lemma 6.2.

Hence we assume $1 < r \leqslant n$ and that P_s is valid for $s < r$. We argue indirectly and assume that $x^{\lambda_r} - a_r$ is reducible over K_{r-1}. Then for some $p \mid \lambda_r$ we have

$$a_r = \beta^p, \qquad \beta \in K_{r-1}. \tag{6.4}$$

Equation (6.4) together with the irreducibility of $x^{\lambda_{r-1}}-\alpha_{r-1}$ over K_{r-2} and the fact that $K_{r-1}=K_{r-2}(\sqrt[\lambda_{r-1}]{\alpha_{r-1}})$ allows the application of Lemma 6.2; we obtain

$$\alpha_r=\alpha_{r-1}^{m_r\lambda_{r-1}^{-1}}\delta^p \quad \text{where} \quad \delta\in K_{r-2}, \tag{6.5}$$

and where $p\nmid m_{r-1}\Rightarrow p\mid\lambda_{r-1}$.

From (6.5) we have

$$\alpha_r\alpha_{r-1}^{-m_r\lambda_{r-1}^{-1}}=\delta^p \quad \text{where} \quad \delta\in K_{r-2}. \tag{6.6}$$

Hence if $r>2$ we have $1\leqslant r-2<r$ and the induction hypothesis gives

$$\alpha_r\alpha_{r-1}^{-m_r\lambda_{r-1}^{-1}}=\alpha_1^{m_1}\dots\alpha_{r-2}^{m_{r-2}}\eta^p \quad \text{where} \quad \eta\in F, \tag{6.7}$$

and where

$$p\nmid m_i\Rightarrow p\mid\lambda_i \quad \text{for} \quad 1\leqslant i\leqslant r-2. \tag{6.8}$$

From (6.6), (6.7) and (6.8) we have

$$\alpha_r=\alpha_1^{m_1}\dots\alpha_{r-1}^{m_r\lambda_{r-1}^{-1}}\eta^p \quad \text{where} \quad \eta\in F, \tag{6.9}$$

and where

$$p\nmid m_i\Rightarrow p\mid\lambda_i \quad \text{for} \quad 1\leqslant i\leqslant r-1. \tag{6.10}$$

If $r=2$, (6.9) and (6.10) remain valid by (6.5) and (6.6). However from the definition of λ_r', (6.9) and (6.10) give $p\mid\lambda_r'$ and hence $p\nmid\lambda_r$, contradicting the initial assumption that $p\mid\lambda_r$.

We have now proved that $x^{\lambda_r}-\alpha_r$ is irreducible over K_{r-1}. One proves (b) in exactly the same way that (6.9) and (6.10) were deduced from (6.4). (One uses the irreducibility of $x^{\lambda_r}-\alpha_r$ over K_{r-1}.)

This completes the proof by induction of the validity of P_r for $1\leqslant r\leqslant n$.

7. An upper estimate for $\Delta(K)$. The argument of Hooley [6] (Section 5) dealt with the Dedekind zeta function of the field $\boldsymbol{Q}(\sqrt[k]{1},\sqrt[k]{a})$. To enable the argument of that section to carry over to $K=\boldsymbol{Q}(\sqrt[k]{1},\sqrt[l_1]{a_1},\dots,\sqrt[l_n]{a_n})$ where $k=\langle l_1,\dots,l_n\rangle$, it suffices to verify that

$$|\Delta(K)|\leqslant k^{AN(K)} \tag{7.1}$$

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 $a^k=a$, a non-zero rational integer. Then if x^k-a is irreducible over F, we have*

$$\Delta\big(F(\alpha)\big) \text{ divides } \big(\Delta(F)\big)^k(ka)^{[F(\alpha):\boldsymbol{Q}]}. \tag{7.2}$$

Proof. Let $\overline{F}$ 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

$$\mathfrak{d}\big(F(\alpha)\big)=\big(\mathfrak{d}(F)\big)^kN_{F/Q}\Big(\mathfrak{d}\big(\overline{F(\alpha)}/\overline{F}\big)\Big), \tag{7.3}$$

where $\mathfrak{d}\big(F(\alpha)\big)$ is an "absolute" discriminant and $\mathfrak{d}\big(\overline{F(\alpha)}/\overline{F}\big)$ is a "relative" discriminant. Now

$$\mathfrak{d}(\overline{F}[\alpha]/\overline{F})=N_{F(\alpha)/F}\big(g'(\alpha)\big)\overline{F},$$

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

$$\mathfrak{d}(\overline{F}[\alpha]/\overline{F})=N_{F(\alpha)/F}(ka^{k-1})\overline{F}=k^k\big(N_{F(\alpha)/F}(\alpha)\big)^{k-1}\overline{F}=k^ka^{k-1}\overline{F}.$$

Also $\overline{F}[\alpha]$ is a finitely generated $\overline{F}$ sub-module of $\overline{F(\alpha)}$ and hence $\mathfrak{d}\big(\overline{F(\alpha)}/\overline{F}\big)$ divides $\mathfrak{d}(\overline{F}[\alpha]/\overline{F})$, and hence divides $k^ka^{k-1}\overline{F}$. (See Cassels and Fröhlich [2], Corollary 1, p. 12.) Hence $N_{F/\boldsymbol{Q}}\big(\mathfrak{d}\big(\overline{F(\alpha)}/\overline{F}\big)\big)$ divides $N_{F/\boldsymbol{Q}}(k^ka^{k-1}\overline{F})$, and consequently divides $(k^ka^{k-1})^{[F:\boldsymbol{Q}]}\boldsymbol{Z}$, as $a\in\boldsymbol{Z}$.

Equation (7.3) now shows that $\mathfrak{d}\big(F(\alpha)\big)$ divides $\big(\mathfrak{d}(F)\big)^k(k^ka^{k-1})^{[F:\boldsymbol{Q}]}\boldsymbol{Z}$, and hence (7.2) holds.

LEMMA 7.2. *Let $K=F(\sqrt[l_1]{a_1},\dots,\sqrt[l_n]{a_n})$ where F is a number field containing all k-th roots of unity, $a_1,\dots,a_n$ are non-zero rational integers and $l_1,\dots,l_n$ are divisors of k, a square-free integer. Then*

$$\Delta(K) \text{ divides } \big(\Delta(F)\big)^{[K:F]}(l_1\dots l_na_1\dots a_n)^{[K:\boldsymbol{Q}]}. \tag{7.4}$$

Proof. From Lemma 6.1 we recall that certain divisors $\lambda_1,\dots,\lambda_n$ of $l_1,\dots,l_n$, respectively were constructed with the property that $K=F(\sqrt[\lambda_1]{a_1},\dots,\sqrt[\lambda_n]{a_n})$ and $x^{\lambda_i}-a_i$ is irreducible over

$$K_{i-1}=F(\sqrt[\lambda_1]{a_1},\dots,\sqrt[\lambda_{i-1}]{a_{i-1}}) \quad \text{for} \quad 1\leqslant i\leqslant n.$$

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

$$\Delta_i \text{ divides } \Delta_{i-1}^{\lambda_i}(\lambda_ia_i)^{[K_i:\boldsymbol{Q}]} \quad \text{for} \quad 1\leqslant i\leqslant n. \tag{7.5}$$

It follows by induction on i that

$$\Delta_i \text{ divides } \Delta_0^{\lambda_1\dots\lambda_i}(\lambda_1\dots\lambda_ia_1\dots a_i)^{[K_i:\boldsymbol{Q}]} \quad \text{for} \quad 1\leqslant i\leqslant n. \tag{7.6}$$

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

$$\Delta(K)\ (=\Delta_n) \text{ divides } \big(\Delta(F)\big)^{\lambda_1\dots\lambda_n}(\lambda_1\dots\lambda_na_1\dots a_n)^{[K:\boldsymbol{Q}]}$$

and (7.4) follows since $\lambda_i\mid l_i$ for $1\leqslant i\leqslant n$ and $\lambda_1\dots\lambda_n=[K:F]$ by Lemma 6.1.

LEMMA 7.3. *Let $a_1, \dots, a_n$ be non-zero rational integers, $l_1, \dots, l_n$ be divisors of k, a square-free integer, and let*

$$K = Q(\sqrt[k]{1}, \sqrt[l_1]{a_1}, \dots, \sqrt[l_n]{a_n}).$$

Then

(7.7) $\Delta(K)$ *divides* $(kl_1 \dots l_n a_1 \dots a_n)^{N(K)}$, *where* $N(K) = [K:Q]$.

Proof. From Lemma 7.2 we have

(7.8) $\Delta(K)$ divides $\left[\Delta\left(Q(\sqrt[k]{1})\right)\right]^{[K:Q(\sqrt[k]{1})]}(l_1 \dots l_n a_1 \dots a_n)^{N(K)}$.

Now it is well-known (see Cassels and Fröhlich [2], Chapter 3, Section 1, Lemma 6, p. 88) that $\Delta\left(Q(\sqrt[k]{1})\right)$ divides $k^{\varphi(k)}$ when k is square-free. Hence from (7.8)

$$\Delta(K) \text{ divides } k^{[K:Q(\sqrt[k]{1})][Q(\sqrt[k]{1}):Q]}(l_1 \dots l_n a_1 \dots a_n)^{N(K)},$$

which is (7.7).

8. A conditional asymptotic formula for $N(x)$. The assumption is now made that the Riemann hypothesis holds for each of the fields $K = Q(\sqrt[k]{1}, \sqrt[l_1]{a_1}, \dots, \sqrt[l_n]{a_n})$ where $k = \langle l_1, \dots, l_n \rangle$ is square-free. Then inequality (7.1) allows us to deduce with Hooley [6], Section 5, pp. 214–218, that $\pi(x, K)$, the number of prime ideals $\mathfrak{p}$ of K with $N(\mathfrak{p}) \leqslant x$, satisfies

(8.1) $$\pi(x, K) = \operatorname{li} x + O\left(N(K) x^{1/2} \log kx\right).$$

This expansion of $\pi(x, K)$ is substituted into (5.7) to obtain

(8.2) $$P(x, l_1, \dots, l_n; k) = \frac{1}{N(K)} \operatorname{li} x + O(x^{1/2} \log kx).$$

This expansion of $P(x, l_1, \dots, l_n; k)$ is in turn substituted into (4.1) to give

(8.3) $$P(x, k) = \mu(k) \sum_{\substack{l_1 | k \\ \langle l_1, \dots, l_n \rangle = k}} \dots \sum_{l_n | k} \mu(l_1) \dots \mu(l_n) \left(\frac{1}{N(K)} \operatorname{li} x + O(x^{1/2} \log kx) \right)$$

$$= c(k) \operatorname{li} x + O\left(x^{1/2} \log x d^n(k)\right),$$

where

(8.4) $$c(k) = \mu(k) \sum_{\substack{l_1 | k \\ \langle l_1, \dots, l_n \rangle = k}} \dots \sum_{l_n | k} \frac{\mu(l_1) \dots \mu(l_n)}{N(K)}.$$

Hence

(8.5) $$P(x, k) = c(k) \operatorname{li} x + O(x^{1/2+\varepsilon})$$

for each $\varepsilon > 0$.

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 $p \equiv 1 (\mathrm{mod}\, k)$ such that for each prime $q | k$ at least one of $a_1, \dots, a_n$ 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

(8.6) $$N(x, \xi_1) = \operatorname{li} x \sum_k{}' \mu(k) c(k) + O\left(x^{1/2+\varepsilon} \sum_{k \leqslant x^{1/3}} |\mu(k)|\right)$$

$$= \operatorname{li} x \sum_k{}' \mu(k) c(k) + O(x^{5/6+\varepsilon}).$$

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

LEMMA 8.2. *Let $a_1, \dots, a_n$ be non-zero rational integers not ± 1 and let h_i be the largest positive integer such that a_i is a perfect h_i-th power. Also let $k = \langle l_1, \dots, l_n \rangle$ be square-free,*

$$K = Q(\sqrt[k]{1}, \sqrt[l_1]{a_1}, \dots, \sqrt[l_n]{a_n}), \quad N(K) = [K:Q].$$

Then

(8.7) $$\frac{1}{N(K)} \leqslant \frac{\prod_{i=1}^{n} (h_i, l_i)}{k \varphi(k)}.$$

Proof. From Lemma 6.1 we have

$$\frac{1}{N(K)} = \frac{1}{\varphi(k)} \frac{\lambda_1' \dots \lambda_n'}{l_1 \dots l_n}.$$

The construction of $\lambda_1', \dots, \lambda_n'$ reveals the following chain of inequalities:

$$\lambda_1' \leqslant (h_1, l_1),$$

$$\lambda_2' \leqslant (h_2, l_2)(l_2, l_1),$$

$$\lambda_3' \leqslant (h_3, l_3) \langle (l_3, l_1), (l_3, l_2) \rangle,$$

and so on.

We prove by induction that for $1 \leqslant r \leqslant n$

(8.8) $$\frac{\lambda_1' \dots \lambda_r'}{l_1 \dots l_r} \leqslant \frac{\prod_{i=1}^{r} (h_i, l_i)}{\langle l_1, \dots, l_r \rangle}.$$

Inequality (8.8) is clearly true when $r = 1$. Consequently let $1 \leqslant r < n$ and assume that (8.8) holds. Then

$$\frac{\lambda_1' \dots \lambda_{r+1}'}{l_1 \dots l_{r+1}} \leqslant \frac{\prod_{i=1}^{r}(h_i, l_i)}{\langle l_1, \dots, l_r\rangle} \frac{(h_{r+1}, l_{r+1})\langle (l_{r+1}, l_1), \dots, (l_{r+1}, l_r)\rangle}{l_{r+1}}$$

$$= \prod_{i=1}^{r+1}(h_i, l_i) \frac{(l_{r+1}, \langle l_1, \dots, l_r\rangle)}{l_{r+1}\langle l_1, \dots, l_r\rangle} = \frac{\prod_{i=1}^{r+1}(h_i, l_i)}{\langle l_1, \dots, l_{r+1}\rangle}.$$

This completes the induction.

LEMMA 8.3. *Let $a_1, \dots, a_n$ be non-zero rational integers not ± 1 and let*

$$K = Q(\sqrt[k]{1}, \sqrt[l_1]{a_1}, \dots, \sqrt[l_n]{a_n}), \qquad N(K) = [K:Q].$$

Then if $c(k)$ is defined (as in (8.4)) *by*

$$c(k) = \mu(k) \sum_{\substack{l_1 \mid k \\ \langle l_1, \dots, l_n\rangle = k}} \dots \sum_{l_n \mid k} \frac{\mu(l_1) \dots \mu(l_n)}{N(K)},$$

the series $\sum_{k=1}^{\infty} \mu(k)c(k)$ is absolutely convergent.

Proof. From (8.7) we have

$$|c(k)| \leqslant \sum_{\substack{l_1 \mid k \\ \langle l_1, \dots, l_n\rangle = k}} \dots \sum_{l_n \mid k} \frac{1}{k\varphi(k)} = \frac{(2^n - 1)^{\omega(k)}}{k\varphi(k)} \leqslant \frac{d^n(k)}{k\varphi(k)}. \tag{8.9}$$

Now

$$\frac{d^n(k)}{k\varphi(k)} \leqslant \frac{k^{\varepsilon}\log\log k}{k^2} \qquad \text{for each } \varepsilon > 0,$$

and hence

$$\sum_{k=1}^{\infty} \frac{d^n(k)}{k\varphi(k)}$$

converges by comparison with

$$\sum_{k=3}^{\infty} \frac{\log\log k}{k^{2-\varepsilon}}, \qquad 0 < \varepsilon < 1.$$

Consequently $\sum_{k=1}^{\infty} \mu(k)c(k)$ converges absolutely by (8.9).

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

LEMMA 8.4. *Let*

$$S(x) = \sum_{k>x} \frac{d^n(k)}{k\varphi(k)}.$$

Then

$$S(x) \ll x^{-1}(\log x)^{2^n - 1}. \tag{8.10}$$

Proof. (Kindly supplied by Mr. M. Croft.) By a Stieltjes integration

$$S(x) = \int_x^{\infty} t^{-2}\, d\Big(\sum_{x < k \leqslant t} \frac{k}{\varphi(k)} d^n(k)\Big) = 2\int_x^{\infty} t^{-3}\Big(\sum_{x < k \leqslant t} \frac{k}{\varphi(k)} d^n(k)\Big)\, dt. \tag{8.11}$$

Write

$$\Sigma_1 = \sum_{x < k \leqslant t} \frac{k}{\varphi(k)} d^n(k). \tag{8.12}$$

Then

$$\Sigma_1 \leqslant \sum_{k \leqslant t} d^n(k) \prod_{p \mid k}\Big(1 - \frac{1}{p}\Big)^{-1} = \sum_{k \leqslant t} d^n(k) \sum_{l \mid k} \frac{\mu^2(l)}{\varphi(l)}$$

$$= \sum_{l \leqslant t} \frac{\mu^2(l)}{\varphi(l)} \sum_{\substack{k \leqslant t \\ l \mid k}} d^n(k) = \sum_{l \leqslant t} \frac{\mu^2(l)}{\varphi(l)} \sum_{m \leqslant t/l} d^n(ml)$$

$$\leqslant \sum_{l \leqslant t} \frac{\mu^2(l)}{\varphi(l)} d^n(l) \sum_{m \leqslant t/l} d^n(m) \ll t(\log t)^{2^n - 1} \sum_{l \leqslant t} \frac{\mu^2(l)}{l\varphi(l)} d^n(l),$$

by Wilson [12].

Hence

$$\Sigma_1 \ll t(\log t)^{2^n - 1}. \tag{8.13}$$

Hence from (8.11), (8.12) and (8.13) we have

$$S(x) \ll \int_x^{\infty} t^{-2}(\log t)^{2^n - 1}\, dt \ll x^{-1}(\log x)^{2^n - 1}.$$

From (8.6) and (8.10) we deduce that

$$(8.14)\quad N(x,\xi_1) = \mathrm{li}x \sum_{k=1}^{\infty}\mu(k)c(k) + O\left(\frac{x}{\log x}\sum_{k>\xi_1}\frac{d^n(k)}{k\varphi(k)}\right) + O(x^{5/6+\varepsilon})$$

$$= \mathrm{li}x \sum_{k=1}^{\infty}\mu(k)c(k) + O\left(\frac{x}{\log x}\cdot\frac{(\log\xi_1)^{2^n-1}}{\xi_1}\right) + O(x^{5/6+\varepsilon})$$

$$= \mathrm{li}x \sum_{k=1}^{\infty}\mu(k)c(k) + O\left(\frac{x}{\log^2 x}(\log\log x)^{2^n-1}\right)$$

$$= \frac{x}{\log x}\sum_{k=1}^{\infty}\mu(k)c(k) + O\left(\frac{x}{\log^2 x}(\log\log x)^{2^n-1}\right).$$

From (3.3) it remains to estimate $M(x,\xi_1,\xi_2)$ from above, using (3.6). From (8.3) we have

$$P(x,q) = c(q)\mathrm{li}x + O(x^{1/2}\log x) \ll \frac{1}{q(q-1)}\mathrm{li}x + O(x^{1/2}\log x)$$

by (8.9).

Hence from (3.6), following Hooley [6], end of Section 6, we have (recalling that $\xi_1 = \frac{1}{6}\log x$ and $\xi_2 = x^{1/2}\log^{-2}x$)

$$(8.15)\quad M(x,\xi_1,\xi_2) \ll \sum_{\xi_1<q\leqslant\xi_2}\left(\frac{\mathrm{li}x}{q(q-1)} + O(x^{1/2}\log x)\right)$$

$$= O\left(\frac{x}{\log x}\sum_{q>\xi_1}\frac{1}{q^2}\right) + O\left(x^{1/2}\log x\sum_{q\leqslant\xi_2}1\right)$$

$$= O\left(\frac{x}{\xi_1\log x}\right) + O\left(\frac{x^{1/2}\,\xi_2\log x}{\log\xi_2}\right) = O\left(\frac{x}{\log^2 x}\right),$$

which is the estimate required.

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

$$(8.16)\qquad N(x) = \frac{x}{\log x}\sum_{k=1}^{\infty}\mu(k)c(k) + O\left(\frac{x}{\log^2 x}(\log\log x)^{2^n-1}\right),$$

where $c(k)$ 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 $a_1,\dots,a_n$ is ±1 and that the Riemann hypothesis holds for each of the fields $Q(\sqrt[k]{1}, \sqrt[l_1]{a_1},\dots,\sqrt[l_n]{a_n})$ where $k = \langle l_1,\dots,l_n\rangle$ is square-free.

9. Another formula for $[F(\sqrt[l_1]{a_1},\dots,\sqrt[l_n]{a_n}):F]$.

LEMMA 9.1. *Let F be a number field containing all k-th roots of unity, $a_1,\dots,a_n$ are non-zero elements of F, and $l_1,\dots,l_n$ are divisors of k. Then*

$$(9.1)\qquad [F(\sqrt[l_1]{a_1},\dots,\sqrt[l_n]{a_n}):F] = l_1\dots l_n \Big/ \sum_{\substack{r_1=1}}^{l_1}\dots\sum_{r_n=1}^{l_n} 1.$$
$$a_1^{r_1k/l_1}\dots a_n^{r_nk/l_n}=\beta^k,\ \beta\in F$$

Proof. (9.1) is a consequence of the formula

$$(9.2)\qquad [F(\sqrt[k]{a_1},\dots,\sqrt[k]{a_n}):F] = k^n \sum_{\substack{r_1=1}}^{k}\dots\sum_{r_n=1}^{k} 1.$$
$$a_1^{r_1}\dots a_n^{r_n}=\beta^k,\ \beta\in F$$

For $F(\sqrt[l_1]{a_1},\dots,\sqrt[l_n]{a_n}) = F(\sqrt[k]{a_1^{k/l_1}},\dots,\sqrt[k]{a_n^{k/l_n}})$ and by (9.2) with $a_1,\dots,a_n$ replaced by $a_1^{k/l_1},\dots,a_n^{k/l_n}$, we have

$$[F(\sqrt[l_1]{a_1},\dots,\sqrt[l_n]{a_n}):F] = k^n \Big/ \sum_{\substack{r_1=1}}^{k}\dots\sum_{r_n=1}^{k} 1$$
$$a_1^{r_1k/l_1}\dots a_n^{r_nk/l_n}=\beta^k,\ \beta\in F$$

$$= k^n \Big/ \prod_{i=1}^{n}\left(\frac{k}{l_i}\right) \sum_{\substack{r_1=1}}^{l_1}\dots\sum_{r_n=1}^{l_n} 1,$$
$$a_1^{r_1k/l_1}\dots a_n^{r_nk/l_n}=\beta^k,\ \beta\in F$$

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 kth powers of the elements of F^* respectively. Following Hasse [5], pp. 222–223, we let $\{a_1,\dots,a_n,\gamma^k\}$ denote the multiplicative group generated by $a_1,\dots,a_n$ and F^{*k}. Then the reader is referred to Hasse for a proof of the fact that the Galois group of $F(\sqrt[k]{a_1},\dots,\sqrt[k]{a_n})$ over F is isomorphic to $\{a_1,\dots,a_n,\gamma^k\}/F^{*k}$. (See Hasse [5], Satz 152, p. 223.) In particular

$$(9.3)\qquad [F(\sqrt[k]{a_1},\dots,\sqrt[k]{a_n}):F] = |\{a_1,\dots,a_n,\gamma^k\}/F^{*k}|.$$

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 $(r_1,\dots,r_n)$ of residues mod k which satisfy

$$a_1^{r_1}\dots a_n^{r_n} = \beta^k,\qquad \beta\in F,$$

then S/T is isomorphic to $\{a_1, \ldots, a_n, \gamma^k\}/F^{*k}$. Now

$$|S| = k^n \quad \text{and} \quad |T| = \sum_{\substack{v_1=1}}^{k} \cdots \sum_{v_n=1}^{k} 1, \tag{9.4}$$
$$a_1^{v_1} \ldots a_n^{v_n} = \beta^k, \beta \in F$$

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

A "transcendental" proof in the case when $F = \boldsymbol{Q}(\sqrt[k]{1})$ 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 $\boldsymbol{Q}(\sqrt[k]{1})$.

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

(i) $a = \beta^k, \beta \in \boldsymbol{Q}(\sqrt[k]{1}) \Leftrightarrow \begin{cases} k \text{ is odd}, & a = b^k, \ b \in \boldsymbol{Z}, \text{ or} \\ k \text{ is even}, & a = b^{k/2}, \ b \in \boldsymbol{Z}, \ \sqrt{b} \in \boldsymbol{Q}(\sqrt[k]{1}), \end{cases}$

(ii) *if k is even,*

$$a = \beta^{k/2}, \quad \beta \in \boldsymbol{Q}(\sqrt[k]{1}) \Leftrightarrow a = b^{k/2}, \quad b \in \boldsymbol{Z},$$

(iii) $\sqrt{a} \in \boldsymbol{Q}(\sqrt[k]{1}) \Leftrightarrow \varkappa(a) | k$ *and* $\varkappa(a) \equiv 1 \pmod 4$.

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

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

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

$$c(k) = \frac{\mu(k)}{\varphi(k)} \sum_{\substack{l_1 | k \\ \langle l_1, \ldots, l_n \rangle = k}} \cdots \sum_{l_n | k} \frac{\mu(l_1) \ldots \mu(l_n)}{l_1 \ldots l_n} d(l_1, \ldots, l_n; k) \tag{10.1}$$

where

$$d(l_1, \ldots, l_n; k) = \sum_{v_1=1}^{l_1} \cdots \sum_{v_n=1}^{l_n} 1. \tag{10.2}$$
$$a_1^{v_1 k/l_1} \ldots a_n^{v_n k/l_n} = \beta^k, \beta \in \boldsymbol{Q}(\sqrt[k]{1})$$

We now define $d'(l_1, \ldots, l_n; k)$ and $c'(k)$ similarly:

$$d'(l_1, \ldots, l_n; k) = \sum_{v_1=1}^{l_1} \cdots \sum_{v_n=1}^{l_n} 1 \tag{10.3}$$
$$a_1^{v_1 k/l_1} \ldots a_n^{v_n k/l_n} = b^k, b \in \boldsymbol{Z}$$

and

$$c'(k) = \frac{\mu(k)}{\varphi(k)} \sum_{\substack{l_1 | k \\ \langle l_1, \ldots, l_n \rangle = k}} \cdots \sum_{l_n | k} \frac{\mu(l_1) \ldots \mu(l_n)}{l_1 \ldots l_n} d'(l_1, \ldots, l_n; k). \tag{10.4}$$

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

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

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

(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

$$g(k) = \sum_{\substack{l_1 | k \\ \langle l_1, \ldots, l_n \rangle = k}} \cdots \sum_{l_n | k} f(l_1, \ldots, l_n; k)$$

is multiplicative.

Proof of Lemma 10.3. Let $a_1^{v_1 l/l_1} \ldots a_n^{v_n l/l_n} = b^l$, $b \in \boldsymbol{Z}$ and $a_1^{\mu_1 (m/m_1)} \ldots a_n^{\mu_n (m/m_n)} = c^m$, $c \in \boldsymbol{Z}$. Then

$$a_1^{(\lambda_1 m_1 + \mu_1 l_1) \frac{lm}{l_1 m_1}} \ldots a_n^{(\lambda_n m_n + \mu_n l_n) \frac{lm}{l_n m_n}} = (bc)^{lm}.$$

Now the residues $\lambda_i m_i + \mu_i l_i \bmod l_i m_i$ are in 1-1 correspondence with the ordered pairs (λ_i, μ_i) of residues mod l_i and m_i respectively. Consequently it remains to show that every $(v_1, \ldots, v_n)$ which contributes to $d'(l_1 m_1, \ldots, l_n m_n; lm)$ arises in the above way. Accordingly we assume that

$$a_1^{v_1 (lm/l_1 m_1)} \ldots a_n^{v_n (lm/l_n m_n)} = d^{lm}, \quad d \in \boldsymbol{Z}, \tag{10.5}$$

and define n-tuples $(\lambda_1, \ldots, \lambda_n)$ and $(\mu_1, \ldots, \mu_n)$ of residues by the congruences

$$\lambda_i m_i + \mu_i l_i \equiv v_i \pmod{l_i m_i}, \quad 1 \leqslant i \leqslant n.$$

From (10.5) we deduce that

$$a_1^{(\lambda_1 m_1 + \mu_1 l_1) \frac{lm}{l_1 m_1}} \ldots a_n^{(\lambda_n m_n + \mu_n l_n) \frac{lm}{l_n m_n}} = (d')^{lm}$$

and hence

$$(a_1^{\nu_1(l/l_1)} \dots a_n^{\nu_n(l/l_n)})^m = (d'')^l, \quad d'' \in \mathbf{Z}. \tag{10.6}$$

It follows from (10.6) and $(l, m) = 1$ that $d'' = f^m$, $f \in \mathbf{Z}$. Hence from (10.6)

$$a_1^{\lambda_1(l/l_1)} \dots a_n^{\lambda_n(l/l_n)} = \zeta f^l,$$

where ζ is an mth root of unity, necessarily ± 1. Hence

$$a_1^{\lambda_1(l/l_1)} \dots a_n^{\lambda_n(l/l_n)} = (\pm f)^l$$

and $(\lambda_1, \dots, \lambda_n)$ contributes to $d'(l_1, \dots, l_n; l)$; similarly for $(\mu_1, \dots, \mu_n)$.

This completes the proof of Lemma 10.3.

Proof of Lemma 10.4. Let $(l, m) = 1$. Then

$$g(lm) = \sum_{\substack{k_1 \mid lm \\ \langle k_1, \dots, k_n \rangle = lm}} \dots \sum_{k_n \mid lm} f(k_1, \dots, k_n; lm).$$

We now write $k_i = l_i m_i$ where $l_i \mid l$ and $m_i \mid m$. Then

$$\langle k_1, \dots, k_n \rangle = lm \Leftrightarrow \langle l_1, \dots, l_n \rangle = l \text{ and } \langle m_1, \dots, m_n \rangle = m.$$

Hence

$$\begin{aligned} g(lm) &= \sum_{\substack{l_1 \mid l \\ \langle l_1, \dots, l_n \rangle = l}} \dots \sum_{l_n \mid l} \sum_{\substack{m_1 \mid m \\ \langle m_1, \dots, m_n \rangle = m}} \dots \sum_{m_n \mid m} f(l_1 m_1, \dots, l_n m_n; lm) \\ &= \sum_{\substack{l_1 \mid l \\ \langle l_1, \dots, l_n \rangle = l}} \dots \sum_{l_n \mid l} \sum_{\substack{m_1 \mid m \\ \langle m_1, \dots, m_n \rangle = m}} \dots \sum_{m_n \mid m} f(l_1, \dots, l_n; l) f(m_1, \dots, m_n; m) = g(l) g(m). \end{aligned}$$

We now proceed to express $c(k)$ in terms of $c'(k)$ when k is an even square-free integer.

Lemma 10.5. *Let $l = \langle l_1, \dots, l_n \rangle$ and $2 = \langle m_1, \dots, m_n \rangle$ where l is odd and square-free. Also let $k = 2l$ and*

$$D(m_1, \dots, m_n; k) = \sum_{\mu_1 = 1}^{m_1} \dots \sum_{\mu_n = 1}^{m_n} 1. \tag{10.7}$$
$$a_1^{2(\mu_1/m_1)} \dots a_n^{2(\mu_n/m_n)} = \beta^2, \beta \in \boldsymbol{Q}(\sqrt[k]{1})$$

Then

$$d'(l_1 m_1, \dots, l_n m_n; k) = d'(l_1, \dots, l_n; k/2) D(m_1, \dots, m_n; k). \tag{10.8}$$

Proof. The argument is quite similar to that of Lemma 10.3 until we reach (10.6). We must now exhibit both parts of the "onto" argument.

(i) Assume that

$$(a_1^{\lambda_1(l/l_1)} \dots a_n^{\lambda_n(l/l_n)})^2 = (d'')^{k/2}, \quad d'' \in \boldsymbol{Q}(\sqrt[k]{1}).$$

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

(ii) Assume that

$$(a_1^{2(\mu_1/m_1)} \dots a_n^{2(\mu_n/m_n)})^{k/2} = (d'')^2, \quad d'' \in \boldsymbol{Q}(\sqrt[k]{1}).$$

Then from Lemma 10.1 (iii) we have $\varkappa(a^{k/2}) \equiv 1 \pmod 4$ and $\varkappa(a^{k/2}) \mid k$ where $a = a_1^{2(\mu_1/m_1)} \dots a_n^{2(\mu_n/m_n)}$. But $\varkappa(a^{k/2}) = \varkappa(a)$ and hence by Lemma 10.1 (iii) again, we have $a = f^2$, $f \in \boldsymbol{Q}(\sqrt[k]{1})$. Hence $(\mu_1, \dots, \mu_n)$ contributes to $D(m_1, \dots, m_n; k)$.

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 10.6. *Let k be an even square-free integer and let*

$$c''(k) = -\sum_{\substack{m_1 \mid 2 \\ \langle m_1, \dots, m_n \rangle = 2}} \dots \sum_{m_n \mid 2} \frac{\mu(m_1) \dots \mu(m_n)}{m_1 \dots m_n} D(m_1, \dots, m_n; k) \tag{10.9}$$

where $D(m_1, \dots, m_n; k)$ is defined (as in (10.7)) *by*

$$D(m_1, \dots, m_n; k) = \sum_{\mu_1 = 1}^{m_1} \dots \sum_{\mu_n = 1}^{m_n} 1.$$
$$a_j^{2(\mu_1/m_1)}, \dots a_n^{2(\mu_n/m_n)} = \beta^2, \beta \in \boldsymbol{Q}(\sqrt[k]{1})$$

Then with $c(k)$ defined as in (10.1), *we have*

$$c(k) = c'(k/2) c''(k).$$

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 F be a field and let p be a prime. For $1 \leqslant i_1 < \dots < i_j \leqslant n$ let*

$$\tau(i_1, \dots, i_j) = \sum_{\nu_{i_1} = 1}^{p} \dots \sum_{\nu_{i_j} = 1}^{p} 1. \tag{10.10}$$
$$a_{i_1}^{\nu_{i_1}} \dots a_{i_j}^{\nu_{i_j}} = b^p, b \in F$$

Also let

$$\sigma_j = \sum_{1 \leqslant i_1 < \dots < i_j \leqslant n} \tau(i_1, \dots, i_j), \quad \sigma_0 = 1. \tag{10.11}$$

Then if $\tau^*(i_1, \dots, i_j)$ and σ_j^* are defined similarly, but with none of $v_{i_1}, \dots, v_{i_j}$ divisible by p, we have

$$\sum_{j=0}^{n}(-1)^j p^{n-j}\sigma_j = \sum_{j=0}^{n}(-1)^j (p-1)^{n-j}\sigma_j^*. \tag{10.12}$$

Proof. The following identity holds:

$$\tau(i_1, \dots, i_j) = 1 + \sum_{1\leqslant j_1\leqslant j}\tau^*(i_{j_1}) + \sum_{1\leqslant j_1<j_2\leqslant j}\tau^*(i_{j_1}, i_{j_2}) + \dots + \tau^*(i_1, \dots, i_j).$$

Hence

$$\begin{aligned}\sigma_j &= \sum_{1\leqslant i_1<\dots<i_j\leqslant n}\tau(i_1, \dots, i_j) \\ &= \sum_{1\leqslant i_1<\dots<i_j\leqslant n} 1 + \sum_{r=1}^{j}\sum_{1\leqslant i_1<\dots<i_j\leqslant n}\ \sum_{1\leqslant j_1<\dots<j_r\leqslant j}\tau^*(i_{j_1}, \dots, i_{j_r}) \\ &= \binom{n}{j} + \sum_{r=1}^{j}\sum_{1\leqslant I_1<\dots<I_r\leqslant n}\tau^*(I_1, \dots, I_r)\sum_{\substack{1\leqslant i_1<\dots<i_j\leqslant n\\ 1\leqslant j_1<\dots<j_r\leqslant j\\ i_{j_1}=I_1, \dots, i_{j_r}=I_r}} 1.\end{aligned} \tag{10.13}$$

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 set of r elements. This number is $\binom{n-r}{j-r}$. Hence

$$\sigma_j = \binom{n}{j} + \sum_{r=1}^{j}\binom{n-r}{j-r}\sigma_r^* = \binom{n}{j} + \sum_{r=1}^{n}\binom{n-r}{j-r}\sigma_r^* = \sum_{r=0}^{n}\binom{n-r}{j-r}\sigma_r^*. \tag{10.14}$$

Hence from (10.14) we have

$$\begin{aligned}\sum_{j=0}^{n}(-1)^j p^{n-j}\sigma_j &= \sum_{j=0}^{n}(-1)^j p^{n-j}\sum_{r=0}^{n}\binom{n-r}{j-r}\sigma_r^* \\ &= \sum_{r=0}^{n}\sigma_r^*\sum_{j=0}^{n}(-1)^j p^{n-j}\binom{n-r}{j-r} = \sum_{r=0}^{n}(-1)^r (p-1)^{n-r}\sigma_r^*,\end{aligned}$$

as asserted.

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

$$c''(k) = c'(2) - \frac{1}{2^n}\sum_{j=1}^{n}(-1)^j \sum_{\substack{1\leqslant i_1<\dots<i_j\leqslant n\\ \varkappa(a_{i_1}\dots a_{i_j})\equiv 1 \pmod 4\\ \varkappa(a_{i_1}\dots a_{i_j})\mid k\\ \varkappa(a_{i_1}\dots a_{i_j})\neq 1}} 1. \tag{10.15}$$

Proof.

$$\begin{aligned}c''(k) &= 1 - \sum_{m_1\mid 2}\dots\sum_{m_n\mid 2}\frac{\mu(m_1)\dots\mu(m_n)}{m_1\dots m_n}\sum_{\mu_1=1}^{m_1}\dots\sum_{\mu_n=1}^{m_n}\ \sum_{\substack{a_1^{2(\mu_1/m_1)}\dots a_n^{2(\mu_n/m_n)}=\beta^2\\ \beta\in Q(\sqrt{1})^k}} 1 \\ &= 1 - \frac{1}{2^n}\sum_{j=0}^{n}(-1)^j 2^{n-j}\sum_{1\leqslant i_1<\dots<i_j\leqslant n}\ \sum_{\substack{\nu_{i_1}=1\\ a_{i_1}^{\nu_{i_1}}\dots a_{i_j}^{\nu_{i_j}}=\beta^2,\ \beta\in Q(\sqrt{1})^k}}^{2}\dots\sum_{\nu_{i_j}=1}^{2} 1 \\ &= 1 - \frac{1}{2^n}\sum_{j=0}^{n}(-1)^j \sum_{\substack{1\leqslant i_1<\dots<i_j\leqslant n\\ a_{i_1}\dots a_{i_j}=\beta^2,\ \beta\in Q(\sqrt{1})^k}} 1,\end{aligned}$$

by Lemma 10.7.

Similarly

$$c'(2) = 1 - \frac{1}{2^n}\sum_{j=0}^{n}(-1)^j \sum_{\substack{1\leqslant i_1<\dots<i_j\leqslant n\\ a_{i_1}\dots a_{i_j}=b^2,\ b\in\mathbf{Z}}} 1. \tag{10.16}$$

Hence

$$c''(k) = c'(2) - \frac{1}{2^n}\sum_{j=1}^{n}(-1)^j \sum_{\substack{1\leqslant i_1<\dots<i_j\leqslant n\\ a_{i_1}\dots a_{i_j}=\beta^2\\ \beta\in Q(\sqrt{1})^k,\ \beta\notin\mathbf{Z}}} 1,$$

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

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

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

$$c(k) = c'(k) - \frac{1}{2^n}c'(k/2)\sum_{j=1}^{n}(-1)^j \sum_{\substack{1\leqslant i_1<\dots<i_j\leqslant n\\ \varkappa(a_{i_1}\dots a_{i_j})\equiv 1 \pmod 4\\ \varkappa(a_{i_1}\dots a_{i_j})\mid k\\ \varkappa(a_{i_1}\dots a_{i_j})\neq 1}} 1. \tag{10.17}$$

Lemma 10.10. *The series $\sum_{k=1}^{\infty}\mu(k)c'(k)$ is absolutely convergent.*

Proof. First let k be an even square-free integer. Then by (10.17) we have

$$\begin{aligned} c'(k) &= c(k)+O\big(c(k/2)\big) \\ &= O\left(\frac{d^n(k)}{k\varphi(k)}\right)+O\left(\frac{d^n(k/2)}{\frac{1}{2}k\varphi(k/2)}\right) \quad \text{by (8.9)}, \\ &= O\left(\frac{d^n(k)}{k\varphi(k)}\right). \end{aligned}$$

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

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

Lemma 11.1. *For each prime p let*

$$\chi(p) = 1-c'(p) = 1-c(p). \tag{11.1}$$

Then

$$\sum_{k=1}^{\infty}\mu(k)c'(k) = \prod_{p}\chi(p). \tag{11.2}$$

Lemma 11.2.

$$\chi(2) = \frac{1}{2^n}\sum_{\substack{\varepsilon_1=0}}^{1}\cdots\sum_{\varepsilon_n=0}^{1}(-1)^{\Sigma\varepsilon_i}, \quad a_1^{\varepsilon_1}\ldots a_n^{\varepsilon_n}=b^2,\ b\in\mathbf{Z} \tag{11.3}$$

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

$$\chi(2) = \frac{1}{2^n}\sum_{j=0}^{n}(-1)^j \sum_{\substack{1\leqslant i_1<\ldots<i_j\leqslant n \\ a_{i_1}\ldots a_{i_j}=b^2,\ b\in\mathbf{Z}}} 1$$

and this is equivalent to (11.3).

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

$$\sum_{k=1}^{\infty}\mu(k)c(k) = \frac{1}{2^n}\prod_{p>2}\chi(p)\sum_{\varepsilon_1=0}^{1}\cdots\sum_{\varepsilon_n=0}^{1}(-1)^{\Sigma\varepsilon_i}f(|a|) \tag{11.4}$$

$$a=*(a_1^{\varepsilon_1}\ldots a_n^{\varepsilon_n})\equiv 1\pmod 4$$

where

$$f(|a|) = \mu(|a|)\prod_{p\mid|a|}\frac{c'(p)}{1-c'(p)}. \tag{11.5}$$

Proof. From Lemmas 10.2, 10.9 and 11.1 we have

$$\begin{aligned} &\sum_{k=1}^{\infty}\mu(k)c(k) \\ &= \sum_{k=1}^{\infty}\mu(k)c'(k) - \frac{1}{2^n}\sum_{j=1}^{n}(-1)^j \sum_{\substack{1\leqslant i_1<\ldots<i_j\leqslant n \\ a=*(a_{i_1}\ldots a_{i_j})\equiv 1\pmod 4 \\ a\neq 1}} \sum_{\substack{k=1 \\ 2a\mid k}}^{\infty}\mu(k)c'\left(\frac{k}{2}\right) \\ &= \prod_{p}\chi(p) - \frac{1}{2^n}\sum_{j=1}^{n}(-1)^j \sum_{\substack{1\leqslant i_1<\ldots<i_j\leqslant n \\ a=*(a_{i_1}\ldots a_{i_j})\equiv 1\pmod 4 \\ a\neq 1}} \mu(|a|)c'(|a|)\sum_{\substack{k=1 \\ (k,2a)=1}}^{\infty}\mu(k)c'(k) \\ &= \frac{1}{2^n}\prod_{p>2}\chi(p)\left(2^n\chi(2)+\sum_{j=1}^{n}(-1)^j \sum_{\substack{1\leqslant i_1<\ldots<i_j\leqslant n \\ a=*(a_{i_1}\ldots a_{i_j})\equiv 1\pmod 4 \\ a\neq 1}} f(|a|)\right) \\ &= \frac{1}{2^n}\prod_{p>2}\chi(p)\sum_{\varepsilon_1=0}^{1}\cdots\sum_{\varepsilon_n=0}^{1}(-1)^{\Sigma\varepsilon_i}f(|a|) \\ &\qquad a=*(a_1^{\varepsilon_1}\ldots a_n^{\varepsilon_n})\equiv 1\pmod 4 \end{aligned}$$

by Lemma 11.2.

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

$$0\leqslant c(p)\leqslant\frac{1}{p-1} \tag{11.6}$$

and hence

$$\chi(p)\geqslant 1-\frac{1}{p-1}>0 \quad \text{if} \quad p>2.$$

Consequently by (11.4) the vanishing of $\sum_{k=1}^{\infty}\mu(k)c(k)$ 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 $\mathbf{C}_1$ is false. If $\mathbf{C}_1$ holds, some simplification of the finite sum is possible.

It is convenient to state

LEMMA 11.4. *Let $S(a_1, \dots, a_n)$ be the set of n-tuples $(\varepsilon_1, \dots, \varepsilon_n)$, $\varepsilon_i = 0$ or 1, satisfying $a_1^{\varepsilon_1} \dots a_n^{\varepsilon_n} = b^2$, $b \in \mathbf{Z}$. Then*

(i) $\mathbf{C}_1$ *implies that* $\chi(2) = \dfrac{1}{2^n}|S(a_1, \dots, a_n)|$,

(ii) $\mathbf{C}_1$ *implies that the number of solutions* $(\varepsilon_1, \dots, \varepsilon_n)$, $\varepsilon_i = 0$ *or* 1 *of* $\varkappa(a_1^{\varepsilon_1} \dots a_n^{\varepsilon_n}) = \varkappa(a_1^{\eta_1} \dots a_n^{\eta_n})$ *is* $|S(a_1, \dots, a_n)|$,

(iii) *if* $\mathbf{C}_1$, *then* $\varkappa(a_1^{\varepsilon_1} \dots a_n^{\varepsilon_n}) = \varkappa(a_1^{\eta_1} \dots a_n^{\eta_n})$ *implies that* $\sum \varepsilon_i \equiv \sum \eta_i \pmod 2$.

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

Let $G(a_1, \dots, a_n)$ be the set of integers (square-free) of the form $a = \varkappa(a_1^{\varepsilon_1} \dots a_n^{\varepsilon_n}) \equiv 1 \pmod 4$, $\varepsilon_i = 0$ or 1. Then from Lemma 11.4 (iii), assuming $\mathbf{C}_1$, the expression $(-1)^{\sum \varepsilon_i}$ depends on a only; and we may define unambiguously a function $\omega(a)$ on $G(a_1, \dots, a_n)$ by the formula

$$\omega(a) = (-1)^{\sum \varepsilon_i}. \tag{11.7}$$

It follows from Lemma 11.4 (ii) that

$$\sum_{\substack{\varepsilon_1=0}}^{1} \dots \sum_{\substack{\varepsilon_n=0 \\ a=\varkappa(a_1^{\varepsilon_1} \dots a_n^{\varepsilon_n})\equiv 1 \pmod 4}}^{1} (-1)^{\sum \varepsilon_i} f(|a|) \tag{11.8}$$

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

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

LEMMA 11.5. *On the assumption that $\mathbf{C}_1$ holds*

$$\sum_{k=1}^{\infty} \mu(k) c(k) = \prod_p \chi(p) \sum_{a \in G(a_1, \dots, a_n)} \omega(a) f(|a|),$$

where $f(|a|)$ is defined by (11.5) *and where $\prod_p \chi(p) > 0$.*

We remark that $G(a_1, \dots, a_n)$ is closed under the associative operation $a \otimes b = \varkappa(ab)$; also $a \otimes a = 1$. Hence (see Ledermann [10], Lemma 3, p. 47) $G(a_1, \dots, a_n)$ is isomorphic to a group $C_2 \times \dots \times C_2$, the direct product of t cyclic groups of order 2. We observe also that

$$\omega(a \otimes b) = \omega(a)\omega(b) \tag{11.9}$$

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

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

$$h(a) = \omega(a) f(|a|). \tag{12.1}$$

In this section we prove that $\sum_{a \in G(a_1, \dots, a_n)} h(a)$ is positive if $\mathbf{C}_2$ holds. But first we need some information about $h(a)$.

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

(i) $|h(a)| \leqslant 1$, *and*

(ii) $\mathbf{C}_2$ *implies that* $h(a) \neq -1$.

Proof. (i) follows from the inequalities

$$|h(a)| = \prod_{p||a|} \frac{c(p)}{1 - c(p)} \leqslant \prod_{p||a|} \frac{1}{p-2} \leqslant 1. \tag{12.2}$$

(See inequalities (11.6).)

To prove (ii) we assume that $h(a) = -1$ for some $a \in G(a_1, \dots, a_n)$. Then inequalities (12.2) imply that $a = -3$ and $c(3) = \frac{1}{2}$. Also as $f(3) = -1$ we have $\omega(-3) = 1$. Consequently from (11.7), $-3 = \varkappa(a_1^{\varepsilon_1} \dots a_n^{\varepsilon_n})$ where $2 | \sum \varepsilon_i$. Hence $a_1^{\varepsilon_1} \dots a_n^{\varepsilon_n} = -3b^2$, $b \in \mathbf{Z}$ where $2 | \sum \varepsilon_i$; this, together with $c(3) = \frac{1}{2}$, implies that $\mathbf{C}_2$ is false.

The reader may be amused to prove that if $\mathbf{C}_2$ fails to hold, then $\sum_{a \in G(a_1, \dots, a_n)} h(a)$ vanishes (as it must by Section 2).

Let $\alpha_1, \dots, \alpha_t$ be a basis for the group $G(a_1, \dots, a_n)$. Then the elements of $G(a_1, \dots, a_n)$ are the numbers $a = \alpha_1^{\varepsilon_1} \otimes \dots \otimes \alpha_t^{\varepsilon_t}$, $\varepsilon_i = 0$ or 1, and hence by (11.9) we have

$$\sum_{a \in G(a_1, \dots, a_n)} h(a) = \sum_{\varepsilon_1=0}^{1} \dots \sum_{\varepsilon_t=0}^{1} (\omega(\alpha_1))^{\varepsilon_1} \dots (\omega(\alpha_t))^{\varepsilon_t} f(|\alpha_1|^{\varepsilon_1} \otimes \dots \otimes |\alpha_t|^{\varepsilon_t}).$$

It remains to express $|\alpha_1|^{\varepsilon_1} \otimes \dots \otimes |\alpha_t|^{\varepsilon_t}$ as a square-free integer as follows. Write each $|\alpha_i|$ as a product of square-free integers

$$|\alpha_i| = \prod_{\lambda_i} A_{\lambda_i}$$

where λ_i runs through all t-tuples $(\eta_1, \dots, \eta_t)$, $\eta_j = 0$ or 1, with $\eta_i = 1$, and where $A_{(\eta_1, \dots, \eta_t)}$ is the product (possibly empty) of those primes which divide each $|\alpha_j|$ where $\eta_j = 1$. Then if $(\varepsilon_1, \dots, \varepsilon_t) \neq (0, \dots, 0)$, the canonical factorisation of $|a| = |\alpha_1|^{\varepsilon_1} \otimes \dots \otimes |\alpha_t|^{\varepsilon_t}$ is found by replacing each $A_{\lambda_i}^2$ by 1 in the product

$$\Big(\prod_{\lambda_1} A_{\lambda_1}\Big)^{\varepsilon_1} \dots \Big(\prod_{\lambda_t} A_{\lambda_t}\Big)^{\varepsilon_t}.$$

We get a factorisation of the type $a = \prod_\lambda A_\lambda$ where λ runs over certain 2^{t-1} t-tuples. It is important to notice that the term $A_{(1,0,\dots,0)}^{e_1} \dots A_{(0,\dots,0,1)}^{e_t}$ is present in the above factorisation of $|a|$. For this factorisation we have

$$h(a) = (\omega(a_1))^{e_1} \dots (\omega(a_t))^{e_t} \prod_\lambda f(A_\lambda) = \prod_\lambda g(A_\lambda)$$

where $g(A_\lambda) = \omega(a_i) f(A_{e_i})$ if $\lambda = e_i = (0, \dots, 1, \dots, 0)$ for some i, and where $g(A_\lambda) = f(A_\lambda)$ otherwise. We observe that Lemma 12.1 implies that

$$|g(A_\lambda)| \leqslant 1 \tag{12.3}$$

and that on the assumption that $\mathbf{C}_2$ holds

$$\prod_\lambda g(A_\lambda) \neq -1. \tag{12.4}$$

We can now prove

LEMMA 12.2. *If* $\mathbf{C}_2$ *holds then*

$$\sum_{a \in G(a_1, \dots, a_n)} h(a) > 0.$$

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

LEMMA 12.3. *Let* p_t *be the polynomial*

$$p_t = \prod_{i=1}^{t} \Big(1 + \prod_{\lambda_i} x_{\lambda_i}\Big) \quad .$$

where λ_i *runs through all* t*-tuples* $(\eta_1, \dots, \eta_t)$ *with* $\eta_j = 0$ *or* 1 *and* $\eta_i = 1$. *A related polynomial* q_t *is defined by replacing* $x_{\lambda_i}^2$ *by* 1 *in all monomials in* p_t *other than the monomial* 1, *which are formed by multiplying out the* t *terms in* p_t. *We write*

$$q_t = 1 + m_1 + \dots + m_{2^t-1} \tag{12.5}$$

where $m_1, \dots, m_{2^t-1}$ *are monomials (each of degree* 2^{t-1} *in fact). Then if all variables* x_λ *take on real values satisfying* $|x_\lambda| \leqslant 1$, *with the added restriction that* $m_j \neq -1$ *for* $j = 1, \dots, 2^t - 1$, *we have* $q_t > 0$.

Proof. From the construction of q_t we observe that $p_t = q_t$ if every $x_\lambda = \pm 1$.

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

$$l = q_t = p_t = \prod_{i=1}^{t} \Big(1 + \prod_{\lambda_i} x_{\lambda_i}\Big) \geqslant 0.$$

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

We write for each x_λ, $q_t = f + g x_\lambda$ where f and g are independent of x_λ. Then if $q_t = 0$ and $-1 < x_\lambda < 1$, we must have $g = 0$, otherwise q_t could be made to assume negative values, contrary to (i). We then replace each x_λ by zero. Hence if $q_t = 0$ we may assume that each x_λ is either zero or ± 1. Consequently from (12.5) we have

$$0 = q_t = 1 + m_1 + \dots + m_{2^t-1}$$

where each m_j is either zero or ± 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 $\sum_{a \in G(a_1, \dots, a_n)} h(a) = q(t)$, where the variables in q_t satisfy $x_\lambda = g(A_\lambda)$. Conditions (12.3) and (12.4) are then precisely those of Lemma 12.3, which in turn gives $q_t > 0$.

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

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

(i) *that if* $a_1^{\varepsilon_1} \dots a_n^{\varepsilon_n} = b^2$, $b \in \mathbf{Z}$, $\varepsilon_i = 0$ *or* 1, *then* $2 | \sum \varepsilon_i$,

(ii) *that if* $a_1^{\varepsilon_1} \dots a_n^{\varepsilon_n} = -3b^2$, $b \in \mathbf{Z}$, $\varepsilon_i = 0$ *or* 1, *and if* $2 | \sum \varepsilon_i$, *then* $d'(3)$, *the natural density of the primes* $q \equiv 1 \pmod 3$, $q \nmid a_1 \dots a_n$, *such that each of* $a_1, \dots, a_n$ *is a cubic non-residue* $\bmod q$, *must be positive,*

(iii) *that the Riemann hypothesis holds for each of the fields* $\boldsymbol{Q}(\sqrt[k]{1}, \sqrt[l_1]{a_1}, \dots, \sqrt[l_n]{a_n})$, *where* $k = \langle l_1, \dots, l_n \rangle$ *is square-free.*

Also let $c(p)$ *be the natural density of the primes* $q \equiv 1 \pmod p$, $q \nmid a_1 \dots a_n$, *such that at least one of* $a_1, \dots, a_n$ *is a* p*-th power residue* $\bmod q$, *and let* $\chi(p) = 1 - c(p)$. *Also let* $G(a_1, \dots, a_n)$ *denote the set of distinct square-free numbers of the form* $a = \varkappa(a_1^{\varepsilon_1} \dots a_n^{\varepsilon_n}) \equiv 1 \pmod 4$, $\varepsilon_i = 0$ *or* 1, *and finally let*

$$\omega(a) = (-1)^{\sum \varepsilon_i} \quad \text{and} \quad f(|a|) = \mu(|a|) \prod_{p \| a|} \frac{c(p)}{1 - c(p)}.$$

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

$$N_{a_1, \dots, a_n}(x) = \frac{x}{\log x} \prod_p \chi(p) \sum_{a \in G(a_1, \dots, a_n)} \omega(a) f(|a|) + O\Big(\frac{x}{\log^2 x} (\log\log x)^{2^n - 1}\Big)$$

as $x \to \infty$, *and the coefficient of* $\dfrac{x}{\log x}$ *is positive.*

THEOREM 13.2. *Let $a_1, \dots, a_n$ be relatively prime in pairs and not ± 1 or a perfect square. We also assume*

(i) *that if $a_1^{\varepsilon_1} \dots a_n^{\varepsilon_n} = -3b^2$, $b \in \mathbf{Z}$, $\varepsilon_i = 0$ or 1, and $2 | \sum \varepsilon_i$, then none of $a_1, \dots, a_n$ is a perfect cube,*

(ii) *that the Riemann hypothesis holds for each of the fields $\boldsymbol{Q}(\sqrt[k]{1}, \sqrt[l_1]{a_1}, \dots, \sqrt[l_n]{a_n})$ where $k = \langle l_1, \dots, l_n \rangle$ 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 $a_1, \dots, a_n$ are relatively prime in pairs and not perfect squares, an equation $a_1^{\varepsilon_1} \dots a_n^{\varepsilon_n} = b^2$, $b \in \mathbf{Z}$, $\varepsilon_i = 0$ or 1, implies that $a_i = -x_i^2$ if $\varepsilon_i = 1$. Hence an equation of the form

$$(-1)^{\sum \varepsilon_i} x^2 = b^2$$

results, and consequently $2 | \sum \varepsilon_i$.

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 $a_1, \dots, a_n$ are relatively prime in pairs and that none of $a_1, \dots, a_n$ is a p-th power, p an odd prime. Then*

$$d'(p) = \frac{1}{p-1}\left(1 - \frac{1}{p}\right)^n. \tag{13.1}$$

Proof. From (10.4) we have

$$d'(p) = \frac{1}{p-1} - c(p) = \frac{1}{p^n(p-1)} \sum_{j=0}^{n} (-1)^j p^{n-j} \sigma_j, \tag{13.2}$$

where

$$\sigma_j = \sum_{1 \leqslant i_1 < \dots < i_j \leqslant n} \tau(i_1, \dots, i_j)$$

and

$$\tau(i_1, \dots, i_j) = \sum_{\substack{\nu_{i_1}=1 \\ a_{i_1}^{\nu_{i_1}} \dots a_{i_j}^{\nu_{i_j}} = b^p,\ b \in \mathbf{Z}}}^{p} \dots \sum_{\nu_{i_j}=1}^{p} 1.$$

But as $a_1, \dots, a_n$ are relatively prime in pairs, we have

$$\tau(i_1, \dots, i_j) = \prod_{l=1}^{j} \tau(i_l)$$

and from (13.2),

$$d'(p) = \frac{1}{p^n(p-1)} \prod_{i=1}^{n} (p - \tau(i)).$$

However as none of $a_1, \dots, a_n$ is a pth power, we have $\tau(i) = 1$ and hence

$$d'(p) = \frac{(p-1)^n}{p^n(p-1)} = \frac{1}{p-1}\left(1 - \frac{1}{p}\right)^n.$$

Finally we state

THEOREM 13.4. *Let a_1 and a_2 be non-zero rational integers not ± 1 or perfect squares, and assume*

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

(ii) *that the Riemann hypothesis holds for each of the fields $\boldsymbol{Q}(\sqrt[k]{1}, \sqrt[l_1]{a_1}, \sqrt[l_2]{a_2})$ 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 $d'(p) > 0$.*

(ii) *If $n = 3$ and none of a_1, a_2, or a_3 is a p-th power, p odd, then $d'(p) > 0$.*

Proof. We use (13.2).

(i) follows from

$$p^2(p-1)d'(p) = \sum_{j=0}^{2} (-1)^j p^{2-j} \sigma_j = p^2 - 2p + \tau(1,2) \geqslant (p-1)^2.$$

(ii) follows from

$$\begin{aligned} p^3(p-1)d'(p) &= \sum_{j=0}^{3} (-1)^j p^{3-j} \sigma_j \\ &= p^3 - 3p^2 + \big(p\tau(1,2) - \tau(1,2,3)\big) + p\big(\tau(2,3) + \tau(1,3)\big) \\ &\geqslant p^3 - 3p^2 + 2p = p(p-2)(p-1). \end{aligned}$$

References

[1] E. Artin, *Collected papers*, edited by S. Lang and J. T. Tate, Addison–Wesley, 1965.

[2] J. S. Cassels and A. Frölich (editors), *Algebraic Number Theory*, Academic Press, 1967.

[3] P. D. T. A. Elliott, *A problem of Erdös concerning power residue sums*, Acta Arith. 13 (1967), pp. 131–149.

[4] — *The distribution of power residues and certain related results*, ibid. 17 (1970), pp. 141–159.

[5] H. Hasse, *Klassenkörpertheorie*, Physica-Verlag, Würzburg 1967.

[6] C. Hooley, *On Artin's conjecture*, J. für Math., Band 225, 1967, pp. 209–220.
[7] E. Landau, *Einführung in die elementare und analytische Theorie der algebraischen Zahlen und der Ideale*, B. G. Teubner, Leipzig und Berlin 1918.
[8] — *Vorlesungen über Zahlentheorie*, Dritter Band, Hirzel, Leipzig 1927.
[9] S. Lang, *Algebra*, Addison–Wesley, 1967.
[10] W. Ledermann, *The Theory of Finite Groups*, Oliver and Boyd, 1957.
[11] A. Schinzel, *A refinement of a theorem of Gerst on power residues*, Acta Arith. 17 (1970), pp. 161–168.
[12] B. M. Wilson, *Proofs of some formulae enunciated by Ramanujan*, Proc. London Math. Soc. (2) 21 (1922), pp. 235–255.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF QUEENSLAND
Brisbane, Queensland, Australia

Received on 21. 2. 1974 (537)

ACTA ARITHMETICA
XXIX (1976)

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 $|z-\gamma| = R$ 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 $n \geqslant 1$ there are algebraic numbers γ of degree n such that there is a circle of centre γ 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 γ 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 γ is rational. Question (II) is very easy when the centre γ is rational — see [2], Theorem 3. In [1] the first author answered both (I) and (II) when γ is totally real, and in [2] we did the same for γ 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 β lies, with its other conjugate (if any), on a circle of rational centre (of course they lie on many circles), such β can be excluded from consideration in answering question (II). Further, these β are clearly of no interest to question (I). We can therefore confine our attention to the set $\mathscr{B}$ of all algebraic numbers β, of degree at least 3 over the rationals Q, whose conjugates (including β) all lie on a circle with irrational centre $\gamma(\beta)$. It is easy to see that $\gamma(\beta)$ must be a real algebraic number.