A problem of Erdős concerning power residue sums

by

P. D. T. A. Elliott (Nottingham)

Let \( k \) be a positive integer. Let \( p \) be a positive rational prime. If \( p \) satisfies \( p \equiv 1 \pmod{k} \), we define \( n_k(p) \) to be the least positive residue which is not a \( k \)-th power \( \pmod{p} \). For other primes we define \( n_k(p) \) to be zero.

Some years ago, in answer to a question of Mirsky, Erdős [3] proved that

\[
\sum_{p \leq x} n_k(p) \sim e \frac{x}{\log x}
\]

as \( x \to \infty \), for a certain constant \( e \). Moreover, he conjectured that a result of this type held for any \( k \).

It is the purpose of the present note to prove that this expectation is justified.

**Theorem 1.** For each integer \( k > 0 \), and constant \( a \) which satisfies \( a < 4e^{-1/2} \), we have as \( x \to \infty \), the asymptotic relation

\[
\sum_{p \leq x} (n_k(p))^a \sim C_{k,a} \frac{x}{\log x},
\]

where \( C_{k,a} \) is a constant. In particular, if \( k \) is an odd prime, we can express \( C_{k,a} \) by

\[
C_{k,a} = \sum_{\ell \text{ odd}} k^{-\ell} \phi^{a-1}.
\]

In this sum \( \phi \) runs over all the rational primes.

The author would like to record his thanks to Professor Heilbronn and Dr. Cassels for their helpful advice. We note that this result has also been stated by Barban ([1]), pp. 61, 62) without proof.

The proof of the theorem falls naturally into three parts. We need various lemmas. Before stating the first of these, we recall that two fields \( E, F \) are said to be *linearly disjoint* over a common subfield \( \theta \), when any
finite set of elements of $E$ which are linearly independent over $G$ remain so over $F$. It is well known (Zariski and Samuel [13], § 16, pp. 169), that this condition is symmetric in $E$ and $F$. We now prove a result, which though of a type well-known in algebraic geometry, seems not to be readily available in the literature.

**Lemma 1.** Let $E, F$ be two extensions of a field $G$, one of which is finite and normal. Then $E$ and $F$ are linearly disjoint over $G$ if and only if their common subfield is $G$.

**Proof.** We can assume, without loss of generality, that $E$ is a finite normal extension of $G$, of degree $n$.

Suppose first that $E$ and $F$ are not linearly disjoint over $G$. Then we can find elements $a_i$, $i = 1, \ldots, m$, of $E$, which are linearly independent over $G$, but linearly dependent over $F$. Thus there are elements $\lambda_i$, $i = 1, \ldots, m$, of $F$, not all zero, so that

$$\sum_{i=1}^{m} \lambda_i a_i = 0.$$

Let $\theta$ be an element generating $E$ over $G$. Then for each value of $i$ satisfying $1 \leq i \leq m$, we can find members $\varrho_i$ in $G$, so that

$$a_i = \sum_{j=1}^{m} \varrho_i \varrho_j^i.$$

Clearly these two relations imply that

$$\sum_{i=1}^{m} \varrho_i \lambda_i = 0.$$

Now not all of the coefficients of the powers of $\varrho$ in this equality are zero. For otherwise, the linear equations

$$\sum_{i=1}^{m} \varrho_i a_i = 0, \quad j = 0, 1, \ldots, n-1,$$

in the variables $a_i$, $i = 1, \ldots, m$, have a non-trivial solution in $G$. Such a solution would imply, by (3) and (4) that $a_1, \ldots, a_m$, were linearly dependent over $G$, contrary to assumption. Thus we see that $\theta$ is the root of an irreducible polynomial $f(x)$, defined over $E$, and of degree at most $n-1$.

Let $g(x)$ be an irreducible polynomial defining $\theta$ over $G$. Since $E$ is a normal extension, $g(x)$ splits completely in $E$ into factors $x - \beta_i$, $i = 1, \ldots, n$, say. The result which we have just proved shows that $f(x)$ divides $g(x)$ in $F(x)$, so that we can assume, without loss of generality, that

$$f(x) = \prod_{i=1}^{s} (x - \beta_i), \quad 1 \leq s \leq n-1.$$

Consider now the elementary symmetric functions of the $\beta_i$, $i = 1, \ldots, s$. These cannot all lie in $G$. For otherwise $f(x)$ would be defined over $G$, and so $g(x)$ would be reducible over $G$, contrary to assumption. Let $\sigma$ be a symmetric function of these $\beta_i$, which does not lie in $G$. Then clearly $\sigma$ lies in both $E$ and $F$, so that

$$G \subseteq \sigma(E) \subseteq E \cap F.$$

This proves one half of the lemma. This result is all we shall need, but for completeness we give a short proof of the remaining half of the lemma.

Suppose now, therefore, that $E$ and $F$ are linearly disjoint over $G$, but that we can find an element $a$, lying in $E$ and $F$, but not in $G$. Then $1, a$ are linearly independent over $G$, but not over $F$. This contradicts our initial hypothesis.

Hence the lemma is proved.

In what follows we shall use $Q$ to denote the field of rational numbers.

**Lemma 2.** Let $l, k$ be positive rational integers. Let $t$ be a rational number which is not a power of a rational number, and for which $-t$ is not a square of a rational number. Then $Q(\sqrt[k]{t})$ can be contained in the cyclotomic field $Q(\sqrt[l]{z})$ only if $l = 1$ or $2$. If $l = 2$ then $t$ must also be made up from squares, and primes which divide $k$.

**Proof.** By considering the prime factors of $l$ it is clear that we need only consider the cases when $l$ does not lie in $Q$, and $l$ is either an odd prime or 1.

If $l \nmid t$ and so $Q(\sqrt[l]{z})$ lies in the Abelian extension $Q(\sqrt[l]{z})$, it follows from Galois theory that $Q(\sqrt[l]{z})$ must be a normal extension of $Q$. In particular, therefore, the polynomial $x^l - t$, which has a linear factor in that extension field, must split completely into linear factors over it, so that the element $S = \exp(2\pi i/l)$ must also be contained in the field $Q(\sqrt[l]{z})$.

Thus we have the situation,

$$Q \subseteq Q(\sqrt[l]{z}) \subseteq Q(\sqrt[l]{z}).$$

Since we are assuming that $l$ is a prime power, and $t$ is not, then $x^l - t$ must be irreducible over $Q$ (Zariski and Samuel [13], Chapter 2, theorem 7).
Thus the degree of \( l \mathcal{V}(\mathcal{V}^t) \) over \( Q \) is \( t \). Since \( l \mathcal{V} \) is of degree \( g(l) \) over \( Q \), we see that \( g(l) \) divides \( t \), so that \( t \) must be even.

It remains therefore only to deal with the case \( t = 4 \). In this case it is easy to see that \( z = -i \), and \( x^a - y \) splits into \( (x^a - y^i)(x^a + y^i) \) over the field \( Q(i) \), where each quadratic factor is irreducible over \( Q(i) \). Thus, the fields \( Q(i) \) and \( Q(\mathcal{V}^i) \) must coincide, and so therefore do their non-trivial automorphisms defined by \( i \to -i \), and \( \mathcal{V}^i \to -\mathcal{V}^i \). From what we have said there must be rational numbers \( a, b \) so that

\[
\hat{i} = a + b\mathcal{V}^i.
\]

Under the automorphism \( \mathcal{V}^i \to -\mathcal{V}^i \) we obtain that

\[
-i = a - b\mathcal{V}^i,
\]

so that \( i = b\mathcal{V}^i \) and \( t = b^{-2} \). This final result is contrary to hypothesis, and our lemma is therefore proved, save for the final assertion. Before proving it we note that we could have chosen a real value of \( \mathcal{V}^i \), so that \( Q(\mathcal{V}^i) \) would be real and therefore could not contain the complex number \( \tilde{z} \). We have preferred the above proof since it is in some ways more natural, and is in a form suitable for generalization.

For the proof of the final assertion we may clearly assume that \( t \) is a squarefree integer. The discriminant of \( Q(\mathcal{V}^i) \) is then \( 4t \) if \( t = 2 \) or \( 3 \) (mod 4), and \( t \) if \( t = 1 \) (mod 4). The rational primes which divide this discriminant ramify in \( Q(\mathcal{V}^i) \), and so in \( Q(\mathcal{V}^i) \).

**Lemma 3.** In addition to the definitions of Lemma 2 let \( q \) be odd, and let \( q_1, \ldots, q_r \) be rational primes. Then the degree of the field \( Q(\mathcal{V}) ; \mathcal{V}_1, \ldots, \mathcal{V}_r \) over \( Q \) is \( t \hat{v}(q) \).

Proof. We consider the fields

\[
K_i = Q(\mathcal{V}^i_1, \mathcal{V}^i_2, \ldots, \mathcal{V}^i_r), \quad i = 1, \ldots, r,
\]

\[
K_4 = Q(\mathcal{V}^i_1).
\]

Let us suppose, for the moment, that we have shown the result for \( K_i \) at \( 1 \leq i < s \). Then the result holds for \( K_i \) unless, by Lemma 1, \( K_i \) and \( Q(\mathcal{V}^i_1) \) have a common subfield which properly includes \( Q, L \) say.

It is clear from what we have said that the polynomial \( x^a - q \) must then be reducible over \( L \), and so in particular there are integers \( \nu, s, v \), \( 0 \leq v < t \), so that

\[
\zeta^v(\mathcal{V}^i_1)^{v+1}
\]

lies in \( L \), and, therefore, so does \( (\mathcal{V}^i_1)^{v+1} \).

By our hypothesis we can find \( \phi_j, j = 0, 1, \ldots, t-1 \), in \( K_{t-1} \), so that

\[
(\mathcal{V}^i_1)^{v+1} = \sum_{j=0}^{t-1} \phi_j(\mathcal{V}^i_1)^j.
\]

Clearly, the automorphisms of \( K_i \) which leave \( K_{t-1} \) fixed, are given by

\[
\sigma_r : \mathcal{V}^i_1 \to \zeta^r(\mathcal{V}^i_1), \quad r = 0, 1, \ldots, t-1.
\]

Since the polynomial \( x^a - q \) is left invariant by each of these automorphisms \( \sigma_r \), there is an integer \( \mu, 0 \leq \mu < t \), so that

\[
\sigma_1(\mathcal{V}^i_1)^{v+1} = \zeta^{\mu}(\mathcal{V}^i_1)^{v+1}.
\]

Thus, by equating the two expressions for \( \zeta^v(\mathcal{V}^i_1)^{v+1} \), we see that

\[
\zeta^v \left( \sum_{j=0}^{t-1} \phi_j(\mathcal{V}^i_1)^j \right) = \sum_{j=0}^{t-1} \phi_j \zeta^\mu(\mathcal{V}^i_1)^j,
\]

and therefore

\[
(\zeta^v - \zeta^\mu) \phi_j = 0, \quad j = 0, 1, \ldots, t-1.
\]

It is clear from this that \( \phi_j = 0 \) unless \( j = \mu \), and therefore

\[
(\mathcal{V}^i_1)^{v+1} = \phi_\mu(\mathcal{V}^i_1)^v.
\]

It follows from this that we can find integers \( t_{i+1}, t_i = \mu \), both zero, and in absolute value less than \( 1 \), so that

\[
V(\mathcal{V}^i_1)^{v+1} = \mu(\mathcal{V}^i_1)^v
\]

lies in \( K_{t-1} \).

Continuing this process, we see that if the lemma does not hold, then we finally arrive at integers \( t_i \) satisfying \( 0 < t_i < t \) for \( i = 1, \ldots, s \), so that \( t_s \not= 0 \), and

\[
V(\mathcal{V}^i_1)^{v+1} = \mu(\mathcal{V}^i_1)^v
\]

lies in \( K_s = Q(\mathcal{V}^i_1) \).

Let \( l = \mathcal{V}^i_1 \), and suppose, as we clearly may, that not all of the integers \( t_i \) are divisible by \( q \). Our construction shows that \( \varepsilon > 1 \) must hold, and clearly, the value which we have made shows that the product

\[
\prod_{i=1}^{s} \phi_i
\]
is not a $q$th power of a rational number. We can therefore apply our previous lemma, and we see that we must have that both $l$ and $q$ have the value 2. This contradicts our initial hypothesis that $q$ is odd, and so the lemma is proved.

If $l$ is a power of 2, then the situation is a little more complicated. Let us consider the particular case

$$L = Q(V_1') \cap Q(V_{a_1}).$$

If $L = Q$, then $L$ is a normal extension of $Q$, and the polynomial $x^2 - q_1$, which is irreducible over $Q$, splits into two conjugate irreducible polynomials over $L$, where $a_1 | l$. Thus, we can find a power of 2, $2^e$, say, with $e \geq 1$, so that $q_1^{2^e}$ lies in $Q(V_{a_1'})$. By Lemma 3 this can only happen if $e = 1$, and $q_1$ is a divisor of $l$.

This case, if it occurs, means that the polynomial $x^2 - q_1$ splits into two conjugate polynomials over $L$, each of degree $\frac{1}{2}$, so that $L$ is a quadratic extension of $Q$. Since $L$ would then contain $V_{a_1'}$, the only possibility for $L = Q(V_{a_1})$. This situation can actually occur, for it is well-known (see for example Weiss [12], 7-3-1, p. 260), that the field $Q(V_{a_1})$ contains the quadratic subfield

$$Q(V_{a_1'}) = Q((-(1/3)^{0.5})^{-1/2}/a_1).$$

Thus we see that the degree of the field $Q(V_{a_1'}, V_{a_1})$ over $Q$, is $2^e(k)$ unless $V_{a_1}$ lies in $Q(V_{a_1})$, when it is $3^{e/2} = k$. In the former case, the algebraic integers

$$\frac{1}{2}(V_{a_1}), \quad j = 0, 1, \ldots, n - 1,$$

are a field basis for $Q(V_{a_1'}, V_{a_1})$ over $Q(V_{a_1})$, with $n = l$. In the latter case we have a similar result with $n = \frac{1}{2}l$.

When considering the field

$$Q(V_{a_1'}, V_{a_1'}) \cap Q(V_{a_1'}) = Q(V_{a_1'})$$

we proceed as in the proof of Lemma 3, and show that either

$$Q(V_{a_1'}, V_{a_1'}) \cap Q(V_{a_1'}) = Q,$$

or we can find integers $a, b, 0 < b < 2^e$ and $e \geq 1$, so that $V_{a_1'}^{b^{2^e}}$ lies in $Q(V_{a_1'})$. Thus, by Lemma 2, $q_1$ must also be a divisor of $l$.

Indeed, proceeding on the lines of Lemma 3, we can prove, by induction on $r$, that the only possible common subfields of $Q(V_{a_1'})$ and

$$Q_{r - 1} = Q(V_{a_1'}, V_{a_2'}, \ldots, V_{a_{r - 1}}')$$

are $Q$ and $Q(V_{a_1'})$. Moreover, if at the $r$th stage this actually occurs, then a field basis for the compositum over $Q_{r - 1}$ is

$$\langle V_{a_1'}^{j}, \quad j = 0, 1, \ldots, \frac{1}{2}l \rangle.$$

We now differentiate between the two possible cases. Firstly we consider the case when $k$ is divisible by 4. By what we have already noted, the field $Q(V_{a_1'})$ contains $Q(V_{a_1})$ when $q_1 | k$, and this field in turn contains $V_{a_1'}^{1/2}$, and so therefore is $V_{a_1'}$. Putting these results together, we see that the degree of $G_r$ (with an obvious definition) over $Q$ is

$$2^{-r}v(k),$$

where $t$ denotes the number of primes $q_i, i = 1, \ldots, r$, which divide $k$.

The second alternative which we mentioned arises when 2 divides $k$ exactly. We can apply the arguments which we have just used, but we need a little more calculation to determine the quadratic subfields of $Q(V_{a_1'})$. In the application we are interested in, we can take $l = 2$ when $k$ is exactly divisible by 2, and we shall here limit ourselves to this case. More especially, we show that the degree over $Q$ of

$$H_r = Q(V_{a_1'}, V_{a_2'}, \ldots, V_{a_r'}) = 0, 1, 2, \ldots,$$

is $2^{-r}v(k)$, where $t$ now represents the number of the primes $q_i$, which divide $k$, and also satisfy $q_i = 1 \mod (4)$. We give a proof by induction on $t$ the number of primes.

Suppose that we have proved the result for up to $r - 1$ primes $q_i$, the case $r = 0$ being trivially true. Then if

$$H_{r - 1} = Q(V_{a_1'}') = L$$

we have to show that $L$ is $Q(V_{a_1'})$ if $q_r$ divides $k$ and satisfies $q_r = 1 \mod (4)$, and is $Q$ otherwise. Now since 2 divides $k$ exactly, the discriminant of $Q(V_{a_1'})$ is not divisible by 2 (Weiss [12], 7-5-8, p. 266). Hence 2 does not ramify in $Q(V_{a_1'})$. Thus if $Q(V_{a_1'})$ lies in $Q(V_{a_1'})$, then $q_r = 1 \mod (4)$
must hold, since otherwise the discriminant of \( Q(\sqrt[r]{g_i}) \) would be \( 4g_r \), and 2 would ramify in \( Q(\sqrt{1}) \). Conversely, if \( g_r \equiv 1 \pmod{4} \) is satisfied, then 
\[ \left( \frac{g_r}{-1} \right) \] is even, and \( Q(\sqrt{1}) \) contains \( Q(\sqrt[r]{g_i}) \).

We now show that \( L \) is contained in \( Q(\sqrt{1}) \), and the stated result will then be immediate. For, if \( L \neq Q \), and the polynomial \( x^2 - g_r \) is irreducible over \( H_{r-1} \), then 
\[ H_{r-1}(\sqrt[r]{g_i}) \subseteq H_{r-1}(\sqrt[r]{g_i}), \]
and these fields must coincide. By comparing their discriminants we see that we must have \( g_r = g_{r-1} \), and this contradicts our initial hypotheses. Thus \( x^2 - g_r \) is reducible over \( H_{r-1} \), and 
\[ L = H_{r-1} \cap Q(\sqrt[r]{g_i}) \subseteq H_{r-1} \cap Q(\sqrt[r]{g_i}). \]
Stepping down through the primes \( g_i \), we see that \( L \) does in fact lie in \( Q(\sqrt{1}) \).

Summarising our results we can state the following result.

**Lemma 4.** If \( l \) is a power of 2, then the degree of the field \( Q(\sqrt[l]{v_1}; \sqrt[l]{v_2}, \ldots, \sqrt[l]{v_r}) \) over \( Q \) is \( c(k)l^m \) where \( c(k) \) is bounded below by a constant depending only on \( k \). In particular, \( c(k) \) is \( 2^l \) with \( l \) equal to the number of primes \( q_i \) dividing \( k \), when \( 4 \mid k \). If \( 2 \) divides \( k \) exactly and \( l = 2 \) then we get a similar result, with \( t \) which counts those \( q_i \) dividing \( k \) and also satisfying \( q_i \equiv 1 \pmod{4} \).

**Lemma 5.** Let \( l_1, l_2, \ldots, l_r \) be the factorisation of \( k \) into prime powers. For each distinct rational primes \( g_1, g_2, \ldots, g_r \), let \( K_i \) denote the field 
\[ Q(\sqrt[l_1]{v_1}; \sqrt[l_2]{v_2}, \ldots, \sqrt[l_r]{g_r}); \sqrt[l_1]{v_2}, \ldots, \sqrt[l_r]{g_r}); \sqrt[l_1]{v_3}, \ldots, \sqrt[l_r]{g_r}); \ldots; \sqrt[l_1]{v_r}, \ldots, \sqrt[l_r]{g_r}), \]
and let \( n_i \) denote its degree over \( Q \). Then we have the estimate 
\[ n_i = 2^{-t}l^m(c(k)), \]
where \( t \) is zero if \( k \) is odd, and otherwise has the values in Lemma 4.

**Proof.** Let \( K_i^0, i = 1, 2, \ldots, r \), denote the fields corresponding to \( K_i \), but containing only the \( l_{i1} \)th, \( l_{i2} \)th, \ldots, \( l_{ir} \)th roots of the \( q_i \). If we prove at the \( i \)th stage that 
\[ K_i^0 \cap Q(\sqrt{v_1}; \sqrt[l_2]{v_2}, \ldots, \sqrt[l_r]{v_r}) = Q(\sqrt[l]{g_i}) \]
then the result will follow from Lemmas 1 and 4.

To prove this result we first note that the L. H. side clearly contains the R. H. side. Denoting the L. H. side by \( L \), we see that if \( L \) contains \( Q(\sqrt[l]{g_i}) \) properly, then its degree over \( Q(\sqrt[l]{1}) \) divides both the degrees of \( K_i^0 \) and the field 
\[ Q(\sqrt[l]{1}; \sqrt[l_2]{v_2}, \ldots, \sqrt[l_r]{v_r}); \sqrt[l_1]{v_1}, \ldots, \sqrt[l_r]{g_r}) \]
over \( Q(\sqrt[l]{1}) \). This can only occur if \( l_{i1} \) has a factor in common with one of the \( l_{i2}, \ldots, l_{i1} \), and by our choice of the \( l_{ij} \) this cannot happen.

Thus the lemma is proved.

**Lemma 6.** Let \( E, F \) be normal extensions of \( G \). Then a prime ideal splits completely in the compositum of \( E \) and \( F \) if and only if it splits completely in \( E \) and \( F \).

**Proof.** For a proof of this result we refer to Hasse [6], I Ehr. 17, p. 50.

**Lemma 7.** Let \( E, F \) be algebraic number fields, and let \( \theta \) in \( F \) generate \( F \) over \( E \), with \( f(x) = 0 \) as its defining equation. Then with finitely many exceptions, the prime ideals \( p \) of \( E \) split completely in \( F \) if and only if \( f(x) \) is completely reducible when considered in the residue class ring \( F/\theta \).

**Proof.** For a proof we refer to Weisz [10], §4 9, p. 185.

**Lemma 8.** Let \( S(a; g_1, \ldots, g_r) \) denote the number of rational primes \( p \), not exceeding \( a \), for which \( g_1, \ldots, g_r \) are all \( k \)-th powers residues (mod. \( p \)). Then we have the relations 
\[ S(a; g_1, \ldots, g_r) = \frac{1}{n_0} \sum_{N < a} 1 + O(a^2) \]
where the prime ideals \( p \) are counted in the ring of integers of \( K_r \), as defined earlier in Lemma 5, and \( n_0 \) denotes the degree of \( K_r \) over \( Q \).

**Proof.** If \( p \) is a rational prime satisfying \( p \equiv 1 \pmod{k} \), then \( p \) splits completely in the cyclotomic field \( Q(\sqrt[l]{1}) \) into \( k \)-th power conjugate prime ideals. Let a typical one of these be \( p \). Then if \( p \) is counted in \( S(a; g_1, \ldots, g_r) \) we can find rational integers \( y_i \), so that 
\[ g_j = y_j^l \pmod{p}, \quad j = 1, \ldots, r, \]
from which 
\[ g_j = y_j^l \pmod{p}, \quad j = 1, \ldots, r. \]

Since \( p \) is of degree 1, any primitive root (mod. \( p \)) is also a primitive root (mod. \( p \)) in \( Q(\sqrt[l]{1}) \). It follows immediately from this that, if (8) is satisfied by integers \( y_i \) of \( Q(\sqrt[l]{1}) \), then it is also satisfied by rational in-
tegers $g_i$. Moreover, $g_i$ is a $k$th power $(\text{mod } p)$ if and only if it is an $l$th power $(\text{mod } p)$ for each prime-power $l$ which divides $k$ exactly.

Let $l$ be an odd prime power exactly dividing $k$. Then $V_{q_1}$ generates $Q(V_{q_1}, \sqrt{b})$ over $Q(V_{q_1})$, and is an integer of the former field. Thus, by Lemma 7, with $f(x) = x^{l} - g_i$, $q_i$ is an $l$th power residue (mod $p$) if and only if $p$ splits completely in $Q(V_{q_1}, \sqrt{b})$, save for finitely many prime ideals $p$.

If now $l$ is a power of two, and $V_{q_1}$ does not lie in $Q(V_{q_1})$, the same proof applies. If, on the other hand, $V_{q_1}$ does lie in $Q(V_{q_1})$, then a basis for $Q(V_{q_1}, \sqrt{b})$ over $Q(V_{q_1})$ is given by

$$
(V_{q_1})^j, \quad j = 0, 1, \ldots, \frac{l-1}{2}.
$$

We now take $f(x) = x^{l} - g_i$, in Lemma 7, and see that $q_i$ is an $l$th power (mod $p$) if and only if $x^{l} - g_i$ is completely reducible (mod $p$), and this happens if and only if $p$ splits completely in $Q(V_{q_1}, \sqrt{b})$. Once again we must allow for finitely many exceptions for $p$.

Carrying out these operations for the primes $q_1, \ldots, q_r$, we see from Lemma 6 that $S(x; q_1, \ldots, q_r)$ counts $\varphi(k)$ times essentially all those primes $p$ of $\overline{Q(V_{q_1})}$ which satisfy $Np < x$, and which split completely in the ring $E$. In other words, the rational prime ideals generated by the primes $p$ counted in $S(x; q_1, \ldots, q_r)$ split completely in $E$.

The statement of the lemma is now immediate, the error term allowing for the finitely many primes $p$ from which exceptional prime ideals $p$ of $Q(V_{q_1})$ may arise, and also for the fact that the sum on the R.H.S. of the equation may count ideals of degree exceeding one. For clearly the number of these does not exceed

$$
n_k \left( \sum_{p < x} 1 + \sum_{p^2 < x} 1 + \ldots \right) = O(x^{\delta_k}).
$$

As it is stated the error term is of course not necessarily uniform with respect to the primes $q_i$, $i = 1, \ldots, r$. Such uniformity can be effected at the cost of a little complication provided we introduce some 'small' additional terms into the error term. We shall say a little more concerning this later, but do not need it for our immediate application.

**Problem of Fermat concerning power residue sums**

**Lemma 9.** For any algebraic number field $K$, we have the asymptotic relation

$$
\sum_{p \leq x} \frac{1}{\varphi(p)} \sim \frac{x}{\log x}
$$

as $x \to \infty$.

**Proof.** This result is, of course, the well-known Prime-Ideal Theorem. For a detailed account we refer to Landau [8].

We can apply this result to Lemma 8, and obtain, as $x \to \infty$, the relation

$$
S(x; q_1, \ldots, q_r) \sim \frac{1}{\varphi} \cdot \frac{x}{\log x}.
$$

We shall need this later in the case when $k$ is an odd prime.

This completes what we need for the first section of the proof. For the second we need some further definitions. In what follows we shall denote the principal ideal generated by an element $\mu$ in the appropriate ring, by $[\mu]$.

From now on until further notice we shall assume that $k$ is an odd prime.

Let $\rho = \exp(2\pi i/k)$, and $\lambda = 1 - \rho$. An algebraic integer $\alpha$, of the field $Q(V_{q_1})$, is said to be primary if $[\lambda] \nmid \alpha$, and if we can find a rational integer $w$, so that

$$
\alpha = w \varphi(\mu x^k).
$$

We now recall the definition of the Eisenstein symbol.

If $p$ is a prime ideal of $Q(V_{q_1})$, and $p \nmid [\lambda]$, then there is a unique rational integer $\tau$, which satisfies $0 < \tau < k$, and

$$
\alpha = \varphi(\tau p), \quad \tau = \frac{1}{k} (Np - 1).
$$

We define the Eisenstein symbol of $\alpha$ (mod $p$) to be

$$
\left( \frac{\alpha}{p} \right) = \epsilon.
$$

Thus, for prime ideals of first degree, when it is defined the symbol has the value $1$ if and only if $\alpha$ is a $k$th power residue (mod $p$). More generally, if $b$ is an ideal of $Q(V_{q_1})$, and $[\lambda]$, $b$ have no proper common ideal factors, we define

$$
\left( \frac{\alpha}{b} \right) = \prod_{p \mid b} \left( \frac{\alpha}{p} \right).
$$
where the product is taken over the prime ideal divisors of \( b \). We need the following result concerning this symbol.

**Lemma 10.** If \( t \neq k \) is a rational prime, and \( a \) in \( Q(\sqrt{t}) \) is primary, so that the ideals \([a], [t]k\) are coprime, then

\[
\left( \frac{t}{[a]} \right)_k = \left( \frac{a}{[t]} \right)_k.
\]

**Proof.** A proof of this reciprocity law, due to Eisenstein, is given in Landau [9], Satz 1032, p. 303.

This is enough for our needs, but it is perhaps worth including the following result, as it enables us to give more complete results later on.

**Lemma 11.** Let \( a \) be an algebraic integer of \( Q(\sqrt{t}) \), and let \( \gamma \) denote the trace of \((k\gamma)^{-1}(a-1)\) taken from \( Q(\sqrt{t}) \) down to \( Q \). Then if \( a \equiv 1 \pmod{[k\gamma]} \), we have that

\[
\left( \frac{k}{[a]} \right)_k = \gamma.
\]

**Proof.** This result is proved by Hasse [7].

We need some more definitions.

Let \( K \) be an algebraic number field. If \( K \) is generated by the element \( \theta \), then \( \theta \) may have some real conjugates \( \theta_i, i = 1, \ldots, d \). Let \( a \) be an integer of \( K \). Clearly, under a mapping \( \theta \rightarrow \theta_i, a \) is taken into a real number.

If this number is positive for each value of \( i \), we say that \( \theta \) is totally positive, and write \( a \succ 0 \).

Let \( f \) be an ideal of \( K \). Two ideals \( a, b \) of \( K \) are said to be equivalent \( \pmod{f} \), if \( (a, f) = (b, f) = [1] \), and if, furthermore, there are two integers \( a, \beta \) of \( K \), which satisfy the conditions,

\[
[a]_a = [\beta]^b, \quad a = \beta = 1 \pmod{f}, \quad a \succ 0, \quad \beta \succ 0.
\]

In such a case we write \( a \sim b \pmod{f} \). As is well known, the above definition divides the ideals of \( K \) into a finite number of equivalence classes, and we shall denote this number by \( h(f) \).

We now need a further result concerning ideal classes, and we use the same terminology.

**Lemma 12.** The number of prime ideals \( p \) of \( K \) satisfying \( Np < x \), and belonging to a particular ideal class \( \pmod{f} \), is, as \( x \rightarrow \infty \),

\[
\frac{1}{h(f)} \log x.
\]

**Proof.** This result, which is not necessarily uniform with respect to \( f \), is proved by Landau [9], Satz LXXXV, p. 112. By taking \( f = [1] \) we see that this lemma includes Lemma 9.

We now apply these results to give a preliminary estimation for \( S(x; q_1, \ldots, q_r) \).

For the time being, let \( \mathcal{N}_r = [kq_1, \ldots, q_r] \) and let no \( q_i \) be \( k \). We first show that the set of values

\[
\left( \frac{q_j}{p} \right)_k, \quad j = 1, \ldots, r,
\]

when they exist, depend only upon the ideal class \( \pmod{\mathcal{N}_r} \) to which \( p \) belongs.

For, if \( p_1 \sim p_2 \pmod{\mathcal{N}_r} \), we can find integers \( \zeta, \delta \) of \( \mathcal{N}_r \) such that

\[
[\zeta]p_1 = [\delta]p_2, \quad \zeta = \delta \pmod{\mathcal{N}_r}, \quad \zeta \succ 0, \quad \delta \succ 0.
\]

Now the integers \( \zeta, \delta \) are clearly primary, and so we may apply Lemma 10 to show that for each \( j, 1 \leq j \leq r \),

\[
\left( \frac{q_j}{[\zeta]} \right)_k = \left( \frac{q_j}{[\delta]} \right)_k = 1,
\]

since \( \zeta = 1 \pmod{[q_j]} \). Similarly we can prove that for each \( j \),

\[
\left( \frac{q_j}{[\delta]} \right)_k = 1.
\]

Thus, we have the relations

\[
\left( \frac{q_j}{p_1} \right)_k = \left( \frac{q_j}{p_2} \right)_k = \left( \frac{q_j}{[\zeta]} \right)_k = \left( \frac{q_j}{[\delta]} \right)_k.
\]

If now, one of the \( q_j \) is \( k \), then we apply Lemma 11 in place of Lemma 10. Thus, in place of the step (10), we see that by taking \( a = \zeta \) in Lemma 11, we have that \( \gamma \) satisfies \( \gamma \equiv 0 \pmod{[k]} \), and

\[
\left( \frac{k}{[\zeta]} \right)_k = 1.
\]

The step (11) can therefore still be made.

Hence if \( k \) is an odd prime, we have for \( S(x; q_1, \ldots, q_r) \) the estimate

\[
S(x; q_1, \ldots, q_r) = \frac{1}{k-1} \sum_1^\infty \sum_1^{x/q_k} 1 + R.
\]

for a certain error term \( R \). Here \( a \) runs through a set of representatives from certain ideal classes \( \pmod{\mathcal{N}_r} \). Let us now estimate \( R \) in detail.
Problem of Erdős concerning power residue sums

During the argument we assumed that the prime ideals \( p \) which we were dealing with did not divide \( h(1) \), \( i = 1, \ldots, r \). Thus we can account for those \( p \) which are so ruled out, by taking a term \( O(r) \) in \( R \). Moreover, in the R. H. double sum of (12), we have possibly included prime ideals which are not of degree 1. In order to allow for these \( R \) must contain a further error of not more than

\[
\frac{1}{k-1} \sum_{k \leq x} \sum_{p < x, \text{prime}} 1 = O(x^{2k}).
\]

Thus, we see that we may take \( R \) to be \( O(r + x^{2k}) \).

We now apply Lemma 12 to (12), and obtain as \( x \to \infty \), the result

\[
S(x; q_1, \ldots, q_r) \sim \frac{1}{k-1} \sum_{k} \frac{1}{h(\mathfrak{N}_k)} \cdot \frac{x}{\log x},
\]

Now we have already shown in (9) that

\[
S(x; q_1, \ldots, q_r) \sim \frac{1}{n_r} \frac{x}{\log x},
\]

where, by Lemma 5, \( n_r = (k-1)k \). Comparing these two estimates we see that

\[
\sum_{k} 1 = k^{-r} h(\mathfrak{N}_k).
\]

The expression (12), along with the estimate for \( R \), and (13), is now in a form suitable for the application of a generalization of the sieve of A. Selberg. All of our estimates up to now have not been necessarily uniform with respect to the prime \( q_i \), \( i = 1, \ldots, r \). The following lemma will enable us to obtain a uniform inequality which will suffice for our purposes.

**Lemma 13.** Let \( \mathfrak{f} \) be an ideal in \( \mathfrak{K} \), with \( h(\mathfrak{f}) \) corresponding ideal classes. Then we can find a positive constant \( g \), depending only upon \( \mathfrak{K} \), so that if \( x \geq 2 \), and \( N \leq x^2 \),

\[
\sum_{\mathfrak{f}, \mathfrak{f} \equiv \mathfrak{f}_0 \mod \mathfrak{N}} 1 \leq \epsilon_1 \frac{x}{h(\mathfrak{f})} \log x.
\]

**Proof.** A proof of this result is given in Rieger [10], Satz 5, p. 161, and Satz 7, p. 164.

Returning to our consideration of \( S(x; q_1, \ldots, q_r) \) when \( k \) is an odd prime, we see that if \( q_1 < q_2 < \ldots < q_r < c_k \log x \), for a small but fixed constant \( c_k \), then by a well-known estimate from elementary number theory,

\[
N(\{kq_1, \ldots, q_r\}) < x^2,
\]

so that by (12), (13) and Lemma 13,

\[
S(x; q_1, \ldots, q_r) \leq \frac{1}{k-1} \sum_{k} \frac{c_k}{h(\mathfrak{N}_k)} \cdot \frac{x}{\log x} + O(x^{2k}).
\]

Thus

\[
S(x; q_1, \ldots, q_r) = O\left( k^{-r} \frac{x}{\log x} + x^{2k} \right).
\]

This completes our considerations of what we need for the second part of the proof.

For the third and final section we need some further lemmas. We shall use some final lemmas \( k \) need not be a prime.

**Lemma 14.** Let \( 1 \leq a_1 < a_2 < \ldots < a_k < N \) be distinct rational integers. For any integer \( m \), and prime \( p \), let \( Z(r; p) \) denote the number of the \( a_i \) which satisfy \( a_i \equiv r \mod p \). Then we have the following inequality:

\[
\sum_{p \leq x^{1/2}} \sum_{r \leq x^{1/2}} \left( Z(r; p) - \frac{N}{p} \right)^2 \leq 2.2N^2 Z^2.
\]

**Proof.** A proof of this example of Linnik's large sieve, we refer to Davenport and Halberstam [4].

**Lemma 15.** Let \( \psi(x, y) \) denote the number of rational integers, not exceeding \( x \), which are made up entirely from primes \( p \leq y \). Then, if for a fixed \( \delta \) satisfying \( 0 < \delta < 1 \) we have that \( x \geq y > (\log x)^{1/2} \), then for any \( x > 0 \),

\[
\psi(x, y) > c(\delta, \varepsilon) x^{1-\varepsilon}.
\]

Here \( c(\delta, \varepsilon) \) is a constant depending upon \( \delta \), \( \varepsilon \) and \( \varepsilon \) only.

**Proof.** This is a sharpened form of the corresponding result by Erdős in his paper, already mentioned. Suppose first that \( y > x \), and define a positive integer \( k \) by \( y^k < y \leq y^{k+1} \). Clearly we have that

\[
\psi(x, y) \geq \frac{1}{h(\mathfrak{f})} \left( \sum_{r \leq x} 1 \right)^k > y^k (2k \log y)^{-k}.
\]

For if \( x \) is large enough our hypotheses guarantee that

\[
(2k \log y)^k \leq (2 \log x)^k \leq \exp \left( \frac{\log x}{\log y} \right) \log (2 \log x) < c \sqrt{x}.
\]

*Acta Arithmetica* 13, 1.
If, on the contrary, \( y \gg x^e \), let \( r = \lfloor 1/e \rfloor \gg 1 \). If \( x \) is, once again, large compared with \( e \), we have,

\[
y(x, y) > \frac{1}{r!} \left( \sum_{p \leq x^e} 1 \right) > c_\varepsilon(x) x^{e-r(r+1)} \gg c_\varepsilon(x) x^{1-x}.
\]

Since \( \varepsilon > 0 \) is arbitrary, the proof of the lemma is complete.

**Lemma 16.** For any rational prime \( p \), and any \( \varepsilon > 0 \), we have the estimate

\[
v_\varepsilon(p) < c_\varepsilon x p^{1-\varepsilon},
\]

with \( x = \frac{1}{2} x^{1-\varepsilon} \).

**Proof.** This result is obtained by using the method of I. M. Vinogradov [11], in conjunction with the well-known character sum estimate of Burgess [3]. The proof requires only simple changes.

We can now give a proof of the theorem. The integer \( k \) is not assumed to be an odd prime unless stated.

**Proof of the theorem.** As indicated earlier, we divide the sum which we estimate into three parts. We write

\[
L_3 = \sum_{\ell < M} (\varepsilon) = L_1 + L_2 + L_3,
\]

where \( L_1 \) is defined to be the left hand sum with the extra condition \( n_k(p) < M \), where \( M \) is an integer. \( L_2 \) and \( L_3 \) are defined similarly; the extra conditions being respectively

\[
M \leq n_k(p) < \varepsilon \log x, \quad \varepsilon \log x \leq n_k(p) < x.
\]

To estimate \( L_1 \) we apply (9) to the relation

\[
L_1 = \sum_{q \leq M} q^x \left( \sum_{p \leq x^{1/2}} \frac{1}{p} \right) - \sum_{q \leq M} q^x (S(x; q) - S(x; q))
\]

where \( q_1, \ldots, q_r \) are the first \( r \) rational primes, and so obtain that

\[
L_1 = \sum_{q < M} q^x (n_k^{-1} - n_k^{-1}) \frac{x}{\log x} + o_M \left( \frac{x}{\log x} \right),
\]

as \( x \to \infty \). We write \( o_M \) to denote that the error term may not tend to zero uniformly with respect to \( M \). By using the estimate

\[
v_\varepsilon \gg c(k) x^e
\]

of Lemma 5, we see that we may extend to infinity in a natural way the series which is the coefficient of the leading term in this expression.
Such a prime will exist if \( s \) is large enough. Then it is evident from the estimate (14), that if \( r \asymp s \) holds, then when \( k \) has an odd prime divisor,

\[
S(s; g_1, \ldots, g_k) \leq S(s; g_1, \ldots, g_k) = O\left(\frac{x^{1+2^{-s}}}{\log x}\right)
= O\left(\frac{x}{\log x} \exp\left(-c_4\log\log x\right)\right).
\]

This result also holds, as in Erdős [3], if \( k \) is a power of 2. Using this, we obtain for \( L_4 \) the estimate

\[
 L_4 \leq (log x)^{\delta} S(s; g_1, \ldots, g_k) < c_6 x (\log x)^{-1}.
\]

Finally, we consider \( L_4 \). If \( N(p) \geq y \), then any rational integer made up from primes not exceeding \( y \), must be a \( k \)-th power residue (mod \( p \)). For any \( x \geq y \), let us consider the set of such integers not exceeding \( x \).

Let us denote these by \( a_i, i = 1, \ldots, Z \), where \( Z = \psi(x^a, y) \). Let \( p \) be a prime for which \( N(p) \geq y \). Then since the integers \( (a_i) \) belong to at most \( (p-1)/k \) residue classes (mod \( p \)), we see that if \( p \geq 2^{k-1} + 1 \),

\[
 p \sum_{i=1}^{Z-1} \frac{Z(p; p)}{\psi(p; p)} \geq \frac{1}{k-1} \left( 1 - \frac{1}{k-1} \right) \frac{Z^2}{p^2}.
\]

Thus, the number of primes \( p < x \), for which \( N(p) \geq y \), is, by Lemma 8, less than

\[
k^3 - k + 1 + 18s^2 Z^{-1} < k^3 + 18s^2 (\psi(x^a, y))^{-1}.
\]

Taking \( s = x^{-1} \) in Lemma 13, we see that if \( y \geq (log x)^{\delta} \) then

\[
\psi(x^a, y) > c_5 x (\log x)^{\delta-1}.
\]

Hence, since Lemma 16 shows that in any case \( \psi(x^a, y) \leq c_6 x (\log x)^{-2} \) for any fixed \( s > 0 \), we have that

\[
 L_4 < c_7 x (\log x)^{\delta} \max_{\sum_{a_i} \leq \psi(x^a, y) \leq \psi(x^a, y)} g^2 \leq c_8 x (\log x)^{-2}.
\]

since, if \( s \) is small enough,

\[
 2x + e + e < 3x + e = 1.
\]

Collecting the results (16), (17), (18) and (19) we see that

\[
 \sum_{p \leq x} \psi(p; p)^{t} = \psi_{\alpha} (1) + O\left( \exp\left(-c_4 (1/\log x)^{\delta}\right)\right).
\]

By letting first \( x \), and then \( \alpha \) tend to infinity, we see that

\[
 \lim_{\alpha \to \infty} \sum_{\alpha} \psi(p; p)^{t}
\]
exists, and has the value \( \psi_{\alpha} \).

This completes the proof of the theorem.

In particular, \( 4^{1-\delta} \geq 4 \epsilon > 1 \) so that we may take \( a = 1 \), and obtain the analogue of Erdős' theorem.

If \( k \) is odd, then \( \psi_{k} = L_4 L_5 \), so that

\[
 \psi_{k} = \frac{k-1}{q(k)} \sum_{i=1}^{k=1} \frac{k}{i} q_{i},
\]

and for odd primes \( q(k) = k-1 \), giving the value of \( \psi_{k} \) stated in the theorem.

Finally, we note that a sharper error term can be obtained in the theorem by using the Siegel-Brauer theorem, (see Brauer [2], Theorem 2, p. 743), for the fields \( K \) in which we need to apply Lemma 7, are all normal extensions of \( K \).

References


