MATHEMATICS OF COMPUTATION
Volume 66, Number 218, April 1997, Pages 807–822
S 0025-5718(97)00797-7

BOUNDS FOR MULTIPLICATIVE COSETS OVER FIELDS OF PRIME ORDER

COREY POWELL

ABSTRACT. Let m be a positive integer and suppose that p is an odd prime with $p \equiv 1 \bmod m$. Suppose that $a \in (\mathbb{Z}/p\mathbb{Z})^*$ and consider the polynomial $x^m - a$. If this polynomial has any roots in $(\mathbb{Z}/p\mathbb{Z})^*$, where the coset representatives for $\mathbb{Z}/p\mathbb{Z}$ are taken to be all integers u with $|u| < p/2$, then these roots will form a coset of the multiplicative subgroup μ_m of $(\mathbb{Z}/p\mathbb{Z})^*$ consisting of the mth roots of unity mod p. Let C be a coset of μ_m in $(\mathbb{Z}/p\mathbb{Z})^*$, and define $|C| = \max_{u\in C} |u|$. In the paper "Numbers Having m Small mth Roots mod p" (*Mathematics of Computation*, Vol. 61, No. 203 (1993),pp. 393-413), Robinson gives upper bounds for $M_1(m,p) = \min_{C\in(\mathbb{Z}/p\mathbb{Z})^*/\mu_m} |C|$ of the form $M_1(m,p) < K_m p^{1-1/\phi(m)}$, where ϕ is the Euler phi-function. This paper gives lower bounds that are of the same form, and seeks to sharpen the constants in the upper bounds of Robinson. The upper bounds of Robinson are proven to be optimal when m is a power of 2 or when $m = 6$.

1. INTRODUCTION

Let $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ denote the integers, rationals, real numbers, and complex numbers, respectively. Suppose that $m > 1$ is a positive integer and that p is an odd prime with $p \equiv 1 \bmod m$. Take the coset representatives for $\mathbb{Z}/p\mathbb{Z}$ to be all integers u with $|u| < p/2$. The multiplicative group $(\mathbb{Z}/p\mathbb{Z})^*$ has a subgroup μ_m of mth roots of unity mod p, which is generated by a single element t.

If $a \in (\mathbb{Z}/p\mathbb{Z})^*$ has any mth roots mod p, then these roots will form a coset of μ_m in $(\mathbb{Z}/p\mathbb{Z})^*$. Let C be a coset of μ_m. Define $|C| = \max_{u\in C} |u|$ and let $||C|| = \sqrt{\sum_{u\in C} u^2}$. These two measures of the "size" of C are related by the inequality $||C||/\sqrt{m} \le |C| \le ||C||$. Define $M_1(m,p) = \min_{C\in(\mathbb{Z}/p\mathbb{Z})^*/\mu_m} |C|$ and let $M_2(m,p) = \min_{C\in(\mathbb{Z}/p\mathbb{Z})^*/\mu_m} ||C||$.

Let K_m be the infimum of all K's such that $M_1(m,p) \le Kp^{1-1/\phi(m)}$ for all $p \equiv 1 \bmod m$, where ϕ is the Euler phi-function. In [6], Robinson proves that such a K_m exists, and gives the following upper bounds for K_m:

1. $K_m \le 2^\tau$, where τ is the number of distinct odd primes dividing m.
2. $K_m \le 3$ if m is divisible by only one prime greater than 3.
3. $K_m \le 2/\sqrt{3}$ if m is divisible by no prime greater than 3.

Robinson conjectures that there are lower bounds for $M_1(m,p)$ of the form $M_1(m,p) \ge Kp^{1-1/\phi(m)}$, but does not prove this result, and does not establish whether or not the upper bounds he gives for K_m can be improved in general. In [3], Konyagin

Received by the editor May 30, 1995 and, in revised form, January 26, 1996.
1991 *Mathematics Subject Classification.* Primary 11A07, 11A15; Secondary 11N05, 11R18, 11R44.

and Shparlinksi prove the lower bound $M_1(m,p) > (p-1)/2 - p^{3/2}/m$, which is a good bound if p is small compared to m. Section 2 establishes that

$$M_1(m,p) \geq (\sqrt{\phi(m)}(\prod_{\substack{q\ prime\\ q|m}} q^{1/(q-1)})/m)p^{1-1/\phi(m)}$$

if p is sufficiently large compared to m. It follows from the bound above that

$$M_1(m,p) \geq (m^{1/(m-1)-1}\sqrt{\phi(m)})p^{1-1/\phi(m)},$$

since $f(x) = x^{1/(x-1)}$ is a decreasing function of x for $x > 1$. Section 2 proves that:

1. $K_m \leq \prod_{\substack{q \text{ odd prime}\\ q|m}} q^{\frac{1}{2q-2}}$, and
2. $K_m \leq 2/\sqrt{3}$ if m is divisible by no prime greater than 3.

These upper bounds are at least as sharp as Robinson's for all m and p. The first upper bound gives the estimates $K_m < C_\epsilon m^\epsilon$ for any $\epsilon > 0$, where $C_\epsilon = \prod_{\substack{q \text{ odd prime}\\ 1/(2q-2)>\epsilon}} q^{1/(2q-2)-\epsilon}$. Hendrik Lenstra has suggested that $K_m < C\sqrt{\ln m}$ for some constant C, but this bound seems difficult to prove.

Section 9 discusses the possibility of improving these upper and lower bounds.

2. Lower bounds for $M_1(m,p)$ and $M_2(m,p)$

Let ζ_m be a primitive mth root of unity. It is well known from Galois theory that $\mathbb{Q}(\zeta_m)$ is a Galois extension of $\mathbb{Q}$ of degree $\phi(m)$, and that the elements σ_j of the Galois group $\mathrm{Gal}(\mathbb{Q}(\zeta_m)/\mathbb{Q})$ of $\mathbb{Q}(\zeta_m)$ over $\mathbb{Q}$ are uniquely defined by the condition $\sigma_j(\zeta_m) = \zeta_m^j$, where $\gcd(j,m) = 1$. Let $N_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\ )$ and $\mathrm{Tr}_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\ )$ denote the norm and trace maps from $\mathbb{Q}(\zeta_m)$ to $\mathbb{Q}$. It is well known from algebraic number theory that the irreducible polynomial of ζ_m over $\mathbb{Q}$ is the mth cyclotomic polynomial $\Phi_m(X) = \prod_{i\in(\mathbb{Z}/m\mathbb{Z})^*}(X - \zeta_m^i)$, and that the ring of integers of $\mathbb{Q}(\zeta_m)$ is $\mathbb{Z}[\zeta_m]$. The ideal generated by p in $\mathbb{Z}[\zeta_m]$ factors as $p\mathbb{Z}[\zeta_m] = \prod_{i\in(\mathbb{Z}/m\mathbb{Z})^*} P_i$, where

$$P_i = p\mathbb{Z}[\zeta_m] + (\zeta_m - t^i)\mathbb{Z}[\zeta_m].$$

The following theorem will also use the facts that $N_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(P_i) = p$ and that $P_i \cap \mathbb{Z} = p\mathbb{Z}$.

Let l be the largest prime dividing m such that $M_1(m,p) < p/l$, if such a prime exists, and let $l = 1$ otherwise. If $p > (2^\tau \max_{q|m, q \text{ prime}} q)^{\phi(m)}$, then $M_1(m,p) < 2^\tau p^{1-1/\phi(m)} < p/\max_{q|m,q \text{ prime}} q$ by the results of Robinson, and so l will be the largest prime dividing m.

Theorem 1. *If l is as above, then*

$$M_1(m,p) \geq (\sqrt{\phi(m)}(\prod_{\substack{q \text{ prime}, q\leq l\\ q|m}} q^{\frac{1}{q-1}})/m)p^{1-1/\phi(m)}.$$

The proof of the theorem will follow directly from the following three lemmas together with the fact that $|N_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}\alpha| = N_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\alpha\mathbb{Z}[\zeta_m])$ for any $\alpha \in \mathbb{Z}[\zeta_m]$. Let C be a coset of $(\mathbb{Z}/p\mathbb{Z})^*$, and let $b_0, \ldots, b_{m-1}$ be the elements of C with $b_j \equiv b_0 t^j \bmod p$. Define $\beta_d = \sum_{j=0}^{m-1} b_j \zeta_m^{jd}$, and let $\overline{\beta}$ denote the complex conjugate of β.

Lemma 1.1. *If β_1 is as above, and C is such that $|C| = M_1(m,p)$, then $q^{\frac{\phi(m)}{q-1}} \mid N_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\beta_1)$ for all $q \leq l$.*

Proof. It suffices to show that $\beta_1 \in (\zeta_m^{m/q}-1)\mathbb{Z}[\zeta_m]$, since

$$N_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\zeta_m^{m/q}-1) = N_{\mathbb{Q}(\zeta_q)/\mathbb{Q}}(N_{\mathbb{Q}(\zeta_m)/\mathbb{Q}(\zeta_q)}(\zeta_m^{m/q}-1)) = q^{\frac{\phi(m)}{q-1}}.$$

It is clear that $\zeta_m^i \equiv \zeta_m^{i \operatorname{mod}(m/q)} \bmod (\zeta_m^{m/q}-1)\mathbb{Z}[\zeta_m]$, and hence that

$$\beta_1 \equiv \sum_{j=0}^{\frac{m}{q}-1} (\zeta_m^j \sum_{k=0}^{q-1} b_{j+km/q}) \bmod (\zeta_m^{m/q}-1)\mathbb{Z}[\zeta_m].$$

It follows from the definition of b_i that

$$\begin{aligned} \sum_{k=0}^{q-1} b_{j+km/q} &\equiv b_j \sum_{k=0}^{q-1} t^{km/q} \bmod p \\ &\equiv b_j \frac{1-t^m}{1-t^{m/q}} \bmod p \\ &\equiv 0 \bmod p \end{aligned}$$

for $0 \le j \le \frac{m}{q}-1$. It now follows that $\sum_{k=0}^{q-1} b_{j+km/q} = 0$ for $0 \le j \le \frac{m}{q}-1$ because

$$|\sum_{k=0}^{q-1} b_{j+km/q}| \le \sum_{k=0}^{q-1} |b_{j+km/q}| < pq/l \le p.$$

This proves the lemma.

Lemma 1.2. *If β_1 is as above, then $\beta_1 \ne 0$, and*

$$p^{2(\phi(m)-1)} \mid N_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\beta_1\overline{\beta_1}).$$

Proof. It follows from the definition of P_j that $\zeta_m \equiv t^j \bmod P_j$, where $\gcd(j,m)=1$, and hence $\beta_1 \equiv b_0 \sum_{k=0}^{m-1} t^{k(j+1)} \bmod P_j$. This sum is a geometric series, and so $\beta_1 \equiv b_0(1-t^{m(j+1)})(1-t^{j+1})^{-1} \equiv 0 \bmod P_j$ provided that $j \ne m-1$. It follows that $p^{\phi(m)-1} \mid N_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\beta_1)$. If $j = m-1$, then $\beta_1 \equiv b_0 m \not\equiv 0 \bmod P_j$, which implies that $\beta_1 \notin P_{m-1}$ and hence that $\beta_1 \ne 0$. The lemma now follows since $N_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\beta_1) = N_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\overline{\beta_1})$.

It is a direct consequence of Lemma 1.1 and Lemma 1.2 that

$$|N_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\beta_1\overline{\beta_1})| \ge (\prod_{\substack{q \text{ prime}, q \le l \\ q|m}} q^{2\phi(m)/p-1})p^{2(\phi(m)-1)}$$

if $|C| = M_1(m,p)$. The theorem will now follow from taking the $2\phi(m)$th root of this inequality and combining it with the following inequality.

Lemma 1.3. *If β_1 is as above, then*

$$|C| \ge (\sqrt{\phi(m)}/m)|N_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\beta_1\overline{\beta_1})|^{1/(2\phi(m))}.$$

Proof. It follows from the arithmetic-geometric mean inequality that

$$\mathrm{Tr}_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\beta_1\overline{\beta_1})/m^2 \ge (\phi(m)/m^2)(N_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\beta_1\overline{\beta_1}))^{1/\phi(m)}.$$

The lemma follows by combining this inequality with the following lemma and the inequality $|C|^2 \ge \|C\|^2/m$ and then taking the square root of both sides.

Lemma 1.4. *If β_1 is as above, then*

$$m\|C\|^2 \geq \mathrm{Tr}_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\beta_1\overline{\beta_1}),$$

where equality holds if $\beta_d = 0$ for all d with $\gcd(d,m) \neq 1$.

Proof. The lemma is a consequence of the following computation.

$$\begin{aligned}
\mathrm{Tr}_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\beta_1\overline{\beta_1}) &= \sum_{j\in(\mathbb{Z}/m\mathbb{Z})^*} (\sum_{k=0}^{m-1} b_k\zeta_m^{jk})\overline{(\sum_{k=0}^{m-1} b_k\zeta_m^{jk})} \\
&\leq \sum_{j \bmod m} (\sum_{k=0}^{m-1} b_k\zeta_m^{jk})\overline{(\sum_{k=0}^{m-1} b_k\zeta_m^{jk})} \\
&= \sum_{j \bmod m} \sum_{k=0}^{m-1}\sum_{l=0}^{m-1} b_k b_l \zeta_m^{(k-l)j} \\
&= m\|C\|^2 + \sum_{\substack{0\leq k,l\leq m-1 \\ k\neq l}} b_k b_l \sum_{j \bmod m} \zeta_m^{(k-l)j} \\
&= m\|C\|^2.
\end{aligned}$$

Combining the direct consequence of Lemma 1.1 and Lemma 1.2 with the arithmetic-geometric mean inequality and Lemma 1.4 gives the following lower bound for $M_2(m,p)$.

Theorem 2. *If $M_2(m,p)$ is as previously defined, then*

$$M_2(m,p) \geq (\sqrt{\phi(m)/m} \prod_{\substack{q\ prime, \leq l \\ q|m}} q^{\frac{1}{q-1}})p^{1-1/\phi(m)}.$$

If p is sufficiently large compared to m, then

$$M_2(m,p) \geq (m^{1/(m-1)}\sqrt{\phi(m)/m})p^{1-1/\phi(m)}.$$

An upper bound for this measure will be given in Section 2.

2. Upper bounds for $M_1(m,p)$ and $M_2(m,p)$

The following upper bounds are obtained by using Minkowski's geometry of numbers. The first upper bound below also gives the estimate $M_1(m,p) < C_\epsilon m^\epsilon p^{1-1/\phi(m)}$ for any $\epsilon > 0$, where C_ϵ is as defined in Section 1.

Theorem 3. *If m and p are as above, then*

$$M_1(m,p) \leq \min(p^{1-1/m}, (\prod_{\substack{q \text{ odd prime} \\ q|m}} q^{1/(2q-2)})p^{1-1/\phi(m)}).$$

If 3 is the only odd prime dividing m, then

$$M_1(m,p) \leq \min(p^{1-1/m}, (2/\sqrt{3})p^{1-1/\phi(m)}).$$

If Λ is a lattice of full rank in $\mathbb{R}^n$ and $B = \{v_i\}_{i=1}^n$ is an ordered $\mathbb{Z}$-basis for Λ, then let $d(\Lambda) = |\det(A)|$, where the ith column of A is v_i. This determinant is independent of the choice of ordered basis for Λ. Note that $d(\Lambda) = \sqrt{|\det(M)|}$, where $M_{ij} = \langle v_i, v_j\rangle$ and $\langle\ ,\ \rangle$ is the standard Euclidean inner product. The theorem above is a consequence of the following theorem (see [5], p. 120).

Theorem 4. *Let Λ be a lattice (of full rank) in $\mathbb{R}^n$ and let K be a bounded 0-symmetric convex body of volume $\mathrm{vol}(K) > 2^n d(\Lambda)$. Then K contains a point $x \neq 0$ of Λ.*

Let $\Psi_m(X)$ be the $m - \phi(m)$th degree polynomial $(X^m - 1)/\Phi_m(X)$ and define

$$V = \{(b_0, \dots, b_{m-1}) \in \mathbb{R}^m | \sum_{j=0}^{m-1} b_j X^j = \Psi_m(X)\Theta(X), \Theta(X) \in \mathbb{R}[X]\}.$$

The subspace V is $\phi(m)$-dimensional since V is isomorphic to the subspace of $\mathbb{R}[X]$ consisting of polynomials r such that $\Psi_m(X) \mid r$ and $\deg(r) < m$. This subspace has a basis $\beta = \{\Psi_m(X), X\Psi_m(X), \dots, X^{\phi(m)-1}\Psi_m(X)\}$. The vector space V contains the lattice

$$L = \{(b_0, \dots, b_{m-1}) \in \mathbb{Z}^m | \sum_{j=0}^{m-1} b_j X^j = \Psi_m(X)\Theta(X), \Theta(X) \in \mathbb{Z}[X]\}.$$

Define also the lattice

$$\mathcal{C} = \{(b_0, \dots, b_{m-1}) \in \mathbb{Z}^m | b_j \equiv b_0 t^j \bmod p, 0 \leq j \leq m-1\}$$

in $\mathbb{R}^m$, and let

$$S_r = \{(b_0, \dots, b_{m-1}) \in \mathbb{R}^m | \max_{0 \leq j \leq m-1} |b_j| < r\}.$$

If $(b_0, \dots, b_{m-1}) \in \mathcal{C}$ and $b_0 \not\equiv 0 \bmod p$, then there is a coset C such that $C = \{b_i \bmod p | 0 \leq i \leq m-1\}$ and so

$$M_1(m,p) \leq |C| \leq \max_{0 \leq j \leq m-1} |b_j|.$$

If r can be chosen so that $r < p$, then $(b_0, \dots, b_{\phi(m)-1}) \in S_r \cap (\mathcal{C} \cap L)$ will have $b_0 \not\equiv 0 \bmod p$. The following lemma proves the first part of Theorem 3.

Lemma 4.1. *If m and p are as above, then $M_1(m,p) \leq p^{1-1/m}$.*

Proof. Let e_i be the ith standard basis element in $\mathbb{R}^m$. The set $B = \{(1, t, \dots, t^{m-1}), pe_i | 2 \leq i \leq m\}$ forms a $\mathbb{Z}$-basis for $\mathcal{C}$, and hence $d(\mathcal{C}) = p^{m-1}$. It is now clear from Theorem 4 that S_r will contain a point of $\mathcal{C}$ if $(2r)^m > 2^m p^{m-1}$, or if $r > p^{1-1/m}$. The lemma now follows from the earlier remarks.

Now, suppose that

$$p^{1-1/m} > (\prod_{\substack{q \text{ odd prime} \\ q|m}} q^{1/(2q-2)}) p^{1-1/\phi(m)}.$$

To apply Theorem 4, define $d(L) = d(I(L))$, where I is an isometry from V to $\mathbb{R}^{\phi(m)}$. Note that this definition is independent of the choice of I and that $d(L) = \sqrt{|\det(M)|}$, where $M_{ij} = \langle v_i, v_j \rangle$ and $B = \{v_i\}_{i=1}^n$ is a basis for L as a $\mathbb{Z}$-module. If the $\phi(m)$-dimensional volume $\mathrm{vol}(S_r \cap V) > 2^{\phi(m)} d(\mathcal{C} \cap L)$, then Theorem 4 would imply that there is a non-zero point of $\mathcal{C} \cap L$ in $S_r \cap V$. It would then be a consequence of these remarks together with the following theorem that there is a non-zero point of $\mathcal{C} \cap L$ in $S_r \cap V$ if

1. $r > (d(\mathcal{C} \cap L))^{1/\phi(m)}$,
2. $r > (d(\mathcal{C} \cap L))^{1/\phi(m)}/\sqrt{2}$ for m even, and
3. $r > ((2/\sqrt{3})d(\mathcal{C} \cap L)/d(L))^{1/\phi(m)}$ if 3 is the only odd prime dividing m.

If r can be chosen such that $r < p$, then it would follow that $M_1(m,p) \leq r$. The following two theorems will prove Theorem 3. Note that $M_1(m,p) = M_1(2m,p)$ if m is odd, since the roots of $X^{2m} - u \bmod p$ are of the form $\pm x$, where x is a root of $X^m - z$ and $z^2 \equiv u \bmod p$.

Theorem 5. *The following are lower bounds for* $\mathrm{vol}(S_r \cap V)$*:*

1. $\mathrm{vol}(S_r \cap V) \geq (2r)^{\phi(m)}$,
2. $\mathrm{vol}(S_r \cap V) \geq (2\sqrt{2}r)^{\phi(m)}$ *if m is even,*
3. $\mathrm{vol}(S_r \cap V) = (\sqrt{3}r)^{\phi(m)}d(L)$ *if $m = 2^e3^f$, with $e, f > 0$.*

Proof. A result of Vaaler (see [8]) shows that $\mathrm{vol}(S_{1/2} \cap V) \geq 1$. A change of variables then establishes the first lower bound. To prove the second lower bound, define

$$W = \{(b_0, \dots .b_{m-1}) | \sum_{j=0}^{m-1} b_jX^j = (X^{m/2} - 1)\Theta(X), \Theta(X) \in \mathbb{R}[X]\}.$$

The fact that m is even implies that $X^{m/2} - 1 \mid \Psi_m(X)$, and so $V \subset W$. The set $B_W = \{w_i\}_{i=0}^{m/2-1}$ is an ordered orthogonal basis for W, where w_i has -1 in the ith coordinate, 1 in the $(i + m/2)$th coordinate, and 0 in all other coordinates. It follows that $w \in S_r \cap W$ if and only if $|a_i| < r$ for $0 \leq i \leq m/2 - 1$, where a_i is the ith coordinate of w with respect to the basis B_W. Map $S_r \cap W$ isometrically to the box $S_{r\sqrt{2}}$ in $\mathbb{R}^{m/2}$ by taking w_i to $\sqrt{2}e_{i+1}$. Applying the result of Vaaler and a change of variables then shows that

$$\begin{aligned}
\mathrm{vol}(S_r \cap V) &= \mathrm{vol}((S_r \cap W) \cap V) \\
&= \mathrm{vol}(S_{r\sqrt{2}} \cap V') \\
&= (2\sqrt{2}r)^{\phi(m)}\mathrm{vol}(S_{1/2} \cap V') \\
&\geq (2\sqrt{2}r)^{\phi(m)},
\end{aligned}$$

where V' is the image of V in $\mathbb{R}^{m/2}$. This establishes the second lower bound.

If $m = 2^e3^f$ with $e, f > 0$, then

$$X^m - 1 = (X^{m/2} - 1)(X^{m/6} + 1)(X^{m/3} - X^{m/6} + 1),$$

with $\Phi_m(X) = X^{m/3} - X^{m/6} + 1$, and hence $\Psi_m(X) = (X^{m/2} - 1)(X^{m/6} + 1)$. The set $B = \{w_i + w_{i+m/6}\}_{i=0}^{m/3-1}$ is an ordered basis for V, and so $B' = \{\sqrt{2}(e_i + e_{i+m/6})\}_{i=1}^{m/3}$ forms an ordered basis for V'. If a_i denotes the ith coordinate of $v' \in V'$ with respect to the basis B', then the ith coordinate of v' with respect to the standard basis is

1. $\sqrt{2}a_i$ if $1 \leq i \leq m/6$,
2. $\sqrt{2}(a_i + a_{i-m/6})$ if $m/6 < i \leq m/3$, and
3. $\sqrt{2}a_{i-m/6}$ if $m/3 < i \leq m/2$.

Hence $v' \in S_{1/2} \cap V'$ if and only if $|a_i| < \frac{1}{2\sqrt{2}}$ for $1 \leq i \leq m/3$ and $|a_i + a_{i+m/6}| < \frac{1}{2\sqrt{2}}$ for $1 \leq i \leq m/6$. The computation at the end of the proof of the second lower bound proved that

$$\mathrm{vol}(S_r \cap V) = (2\sqrt{2}r)^{\phi(m)}\mathrm{vol}(S_{1/2} \cap V'). \tag{1}$$

Map V' to $\mathbb{R}^{m/3}$ by taking $\sqrt{2}(e_i + e_{i+m/6})$ to e_i in $\mathbb{R}^{m/3}$. The volume of the image of $S_{1/2} \cap V'$ in $\mathbb{R}^{m/3}$ can be found by evaluating the multiple integral

$$\int_{\frac{-1}{2\sqrt{2}}}^{\frac{1}{2\sqrt{2}}} \cdots \int_{\frac{-1}{2\sqrt{2}}}^{\frac{1}{2\sqrt{2}}} \int_{\max(\frac{-1}{2\sqrt{2}}, \frac{-1}{2\sqrt{2}} - x_1)}^{\min(\frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{2}} - x_1)} \cdots \int_{\max(\frac{-1}{2\sqrt{2}}, \frac{-1}{2\sqrt{2}} - x_{m/6})}^{\min(\frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{2}} - x_{m/6})} dx_{m/3} \cdots dx_1,$$

which has the same value as

$$\left(\int_{-1/(2\sqrt{2})}^{1/(2\sqrt{2})} \int_{\max(-1/(2\sqrt{2}), -1/(2\sqrt{2}) - x_1)}^{\min(1/(2\sqrt{2}), 1/(2\sqrt{2}) - x_1)} dx_2 dx_1 \right)^{m/6} .$$

A routine computation shows that the value of this double integral is 3/8. It follows that $\mathrm{vol}(S_{1/2} \cap V') = (3/8)^{m/6} d(L)$ since B is a $\mathbb{Z}$-basis for L which maps isometrically to B'. Substituting this value for $S_{1/2} \cap V'$ into (1) proves the third equation, which finishes the proof of the theorem.

Theorem 6. *If $\mathcal{C}$ and L are as above, then*

$$d(\mathcal{C} \cap L) = (\prod_{\substack{q \text{ prime} \\ q \mid m}} q^{\phi(m)/(2q-2)}) p^{\phi(m)-1}.$$

The theorem will be proven by a sequence of lemmas that reduce the theorem to problems in algebraic number theory. The following lemma reduces finding $d(\mathcal{C} \cap L)$ to finding $d(L)$.

Lemma 6.1. *If $\mathcal{C}$ and L are as above, then $d(\mathcal{C} \cap L) = p^{\phi(m)-1} d(L)$.*

Proof. It suffices to show that $\#(L/(\mathcal{C} \cap L)) = p^{\phi(m)-1}$, since

$$d(\mathcal{C} \cap L)/d(L) = \#(L/(\mathcal{C} \cap L)).$$

There is a homomorphism Ω from L to $\mathbb{Z}[\zeta_m]$ defined by $\Omega(b_0, \dots, b_{m-1}) = \sum_{j=0}^{m-1} b_j \zeta_m^j$, with $\Omega(L) = \Psi_m(\zeta_m)\mathbb{Z}[\zeta_m]$. If $\Omega(l_0, \dots, l_{m-1}) = 0$, then

$$\Phi_m(X)\Psi_m(X) \mid \sum_{j=0}^{m-1} l_j X^j,$$

and so $(l_0, \dots, l_{m-1}) = 0$ since $\Phi_m(X)\Psi_m(X) = X^m - 1$. This shows that Ω is injective. The set

$$\mathcal{B} = \{\zeta_m^j \Psi_m(\zeta_m) |\, 0 \leq j \leq \phi(m) - 1\}$$

forms a basis for $\Psi_m(\zeta_m)\mathbb{Z}[\zeta_m]$ as a $\mathbb{Z}$-module, and so $\Omega^{-1}(\mathcal{B})$ forms a basis for L as a $\mathbb{Z}$-module. To determine $\#(L/(\mathcal{C} \cap L))$, consider the $\mathbb{Z}/p\mathbb{Z}$ vector spaces L/pL and $(\mathcal{C} \cap L)/pL$. The projection of $\Omega^{-1}(\mathcal{B})$ to L/pL will form a basis for L/pL, and so L/pL is $\phi(m)$-dimensional.

If $r \in \mathbb{Z}[X]$, then let $\rho(r)$ denote the polynomial in $\mathbb{Z}/p\mathbb{Z}[X]$ derived by reducing the coefficients of r mod p. The proof of Lemma 1.2 demonstrated that $(X - t^j) \mid \rho(\sum_{k=0}^{m-1} t^k X^k)$ if $j \not\equiv -1 \bmod p$ and hence that $\rho(\Psi_m(X)) \mid \rho(\sum_{k=0}^{m-1} t^k X^k)$. Let $\Upsilon(X) = h(X)\Psi_m(X)$, where

$$\rho(h(X)\Psi_m(X)) = \rho(\sum_{j=0}^{m-1} t^j X^j),$$

and write the coefficients of $\Upsilon(X)$ into a vector $v \in \mathbb{Z}^m$. From the construction of v, it is clear that $v \in L \cap \mathcal{C}$ and that $v \notin pL$. From the definition of $\mathcal{C}$, it follows that v spans $(L \cap \mathcal{C})/pL$ and hence that

$$\#\{L/(L \cap \mathcal{C})\} = \#\{(L/pL)/((L \cap \mathcal{C})/pL)\} = p^{\phi(m)-1}.$$

This proves the lemma.

It is another consequence of the lemma that $d(L) = \sqrt{|\det(M)|}$, where

$$M_{ij} = \langle \Omega^{-1}(\zeta_m^{i-1}\Psi_m(\zeta_m)), \Omega^{-1}(\zeta_m^{j-1}\Psi_m(\zeta_m))\rangle.$$

For $\alpha, \gamma \in \mathbb{Z}[\zeta_m]$, define $\langle \alpha, \gamma\rangle_m = \mathrm{Tr}_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\alpha\overline{\gamma}) \in \mathbb{Q}$. The following lemma will be important in finding $|\det(M)|$.

Lemma 6.2. *If $u, v \in L$, then $\langle u, v\rangle = \langle \Omega(u), \Omega(v)\rangle_m/m$.*

Proof. It suffices to prove the lemma in the case $u = v$, since

$$\langle u, v\rangle = (\langle u+v, u+v\rangle - \langle u-v, u-v\rangle)/4$$

and

$$\langle \Omega(u), \Omega(v)\rangle_m = (\langle \Omega(u+v), \Omega(u+v)\rangle_m - \langle \Omega(u-v), \Omega(u-v)\rangle_m)/4.$$

Let $u = (b_0, \ldots, b_{m-1})$; it follows that $\sum_{k=0}^{m-1} b_k\zeta^{kd} = 0$ for $\gcd(d, m) \neq 1$ because $u \in L$. The lemma now follows from Lemma 1.4.

It is a consequence of Lemma 6.2 that $\sqrt{|\det(M)|} = \sqrt{|\det(D)|}/m^{\phi(m)/2}$, where $D_{jk} = \mathrm{Tr}_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\zeta_m^j\Psi_m(\zeta_m)\overline{\zeta_m^k\Psi_m(\zeta_m)})$. The next lemma gives the latter determinant in terms of the discriminant of $\Psi_m(\zeta_m)\mathbb{Z}[\zeta_m]$.

Lemma 6.3. *If D is as above, then $|\det(D)| = |D_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\Psi_m(\zeta_m)\mathbb{Z}[\zeta_m])|$, where $D_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\ )$ denotes the discriminant.*

Proof. If $\mathrm{Gal}(\mathbb{Q}[\zeta_m]/\mathbb{Q}) = \{\sigma_1, \ldots, \sigma_{\phi(m)}\}$, then $D = PP^*$, where $P_{jk} = \sigma_k(\zeta_m^j\Psi_m(\zeta_m))$ and P^* is the conjugate transpose of P. The determinant is a polynomial in its entries, and so the following calculation proves the lemma:

$$\begin{aligned}
|\det(D)| &= |\det(P)\det(P^*)| \\
&= |\det(P)\overline{\det(P)}| \\
&= |(\det(P))^2| \\
&= |D_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\Psi_m(\zeta_m)\mathbb{Z}[\zeta_m])|.
\end{aligned}$$

It is known from algebraic number theory (see [4], p. 66) that

$$D_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\Psi_m(\zeta_m)\mathbb{Z}[\zeta_m]) = (N_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\Psi_m(\zeta_m)))^2 D_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\mathbb{Z}[\zeta_m]).$$

Differentiating the equation $X^m - 1 = \Phi_m(X)\Psi_m(X)$ and substituting ζ_m for X gives $m\zeta_m^{m-1} = \Phi_m'(\zeta_m)\Psi_m(\zeta_m)$. Taking the norm of both sides gives

$$m^{\phi(m)} = |D_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\mathbb{Z}[\zeta_m])N_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\Psi_m(\zeta_m))|.$$

Hence $\det(D) = m^{2\phi(m)}/|D_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\mathbb{Z}[\zeta_m])|$ and so

$$\begin{aligned}
d(L) &= \sqrt{|\det(M)|} \\
&= \sqrt{|\det(D)|}/m^{\phi(m)/2} \\
&= m^{\phi(m)/2}|D_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\mathbb{Z}[\zeta_m])|^{-1/2}.
\end{aligned}$$

It is also known from algebraic number theory (see [1], p. 88) that

$$|D_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\mathbb{Z}[\zeta_m])| = m^{\phi(m)} / \prod_{\substack{q \text{ prime} \\ q|m}} q^{\phi(m)/(q-1)}.$$

This gives $d(L) = \prod_{\substack{q \text{ prime} \\ q|m}} q^{\phi(m)/(2q-2)}$. Theorem 6 follows from Lemma 6.1 together with the above equality.

The inequality $M_2(m,p) \leq (\sqrt{m})M_1(m,p)$ gives an upper bound for $M_2(m,p)$. If p is sufficiently large compared to m, then this bound can be improved by considering spheres. Let $\mathcal{S}_r = \{x \in \mathbb{R}^m | \sqrt{\langle x, x\rangle} < r\}$. In this case, $\text{vol}(\mathcal{S}_r \cap V) = r^{\phi(m)} m^{\phi(m)/2} / \Gamma(\phi(m)/2+1)$, where Γ is the gamma function. If

$$r^{\phi(m)} m^{\phi(m)/2} / \Gamma(\phi(m)/2+1) > (2^{\phi(m)} \prod_{\substack{q \text{ prime} \\ q|m}} q^{\phi(m)/(2q-2)}) p^{\phi(m)-1},$$

then there will be a non-zero point of $\mathcal{C} \cap L$ that is in $\mathcal{S}_r \cap V$ by Theorem 4 and Theorem 6. Solving for r gives

$$r > \left(\frac{2(\Gamma(\phi(m)/2+1))^{1/\phi(m)} \prod_{\substack{q \text{ prime} \\ q|m}} q^{1/(2q-2)}}{\sqrt{m}} \right) p^{1-1/\phi(m)}.$$

If r can be chosen less than p, then $M_2(m,p) < r$. This can be done if

$$\frac{2^{\phi(m)} \Gamma(\phi(m)/2+1) \prod_{\substack{q \text{ prime} \\ q|m}} q^{\phi(m)/(2q-2)}}{m^{\phi(m)/2}} < p.$$

7. A theorem of Hecke

The material on the idele group presented here is taken from ([4], pp. 137–143, 292–293) and ([1], p. 68). Suppose that k is an algebraic number field with $N = [k : \mathbb{Q}]$, and denote the set of prime ideals in the ring of integers of k by $\mathcal{P}$. Let $\mathcal{M}_k$ be the set of absolute values on k, where each absolute value generates a different topology on k and is normalized to induce a standard absolute value on $\mathbb{Q}$. A standard absolute value v on $\mathbb{Q}$ is of the form $v(q) = |q|$ or $v(q) = p^{-o_p(q)}$, where p is prime and $o_p(q)$ is the exponent of p that appears in the prime factorization of q. The set of archimedean absolute values is denoted by S_∞, and the completion of k with respect to an absolute value v is denoted by k_v. The archimedean absolute values v on k are all of the form $v(x) = |\sigma(x)|$, where σ is an embedding of k into $\mathbb{C}$.

The multiplicative group k_v^* is locally compact in the topology generated by the absolute value v on k_v. If v is an absolute value arising from a prime ideal P, then the absolute value will be called *P-adic*. If v is a P-adic absolute value, then the group $\mathcal{O}_v^*$ consisting of all $k \in k_v^*$ with $v(k) = 1$ forms a compact open subgroup of k_v^*. This group will be frequently referred to as the *P-adic units*.

If $j \in \prod_{v \in \mathcal{M}_k} k_v^*$, then let j_v denote the vth component of j. The *idele group* J_k is the set of all j such that j_v is a P-adic unit for all but finitely many P-adic absolute values v. The topology on J_k is that generated by sets of the form $\prod_{v \in \mathcal{M}_k} U_v$, where U_v is open in k_v^* and $U_v = \mathcal{O}_v^*$ for all but finitely many P-adic valuations v. The idele group J_k is a locally compact topological group with respect to this topology.

A number of properties of J_k will become useful later on. First of all, note that $\alpha \in k^*$ is a P-adic unit for all but finitely many P-adic absolute values. This implies that k^* can be embedded in J_k by taking α to $(\alpha, \alpha, \ldots, \alpha, \ldots)$. The quotient J_k/k^* is called the *idele class group* of k, and is a topological group with the quotient topology.

Secondly, define $\|j\| = \prod_{v \in \mathcal{M}_k} v(j_v)$. This product is well-defined, and determines a continuous group homomorphism from J_k to the multiplicative group $\mathbb{R}^+$ of positive real numbers. The kernel of this map is a closed subgroup of J_k denoted by J_k^0. It follows as a consequence of the product formula that $k^* \subset J_k^0$. The projection of J_k^0 to the idele class group gives a compact subgroup J_k^0/k^* (see [4], p. 142).

The multiplicative group $\mathbb{R}^+$ can be embedded in J_k by taking a positive real number t to the idele j whose archimedean components are $t^{1/N}$ and whose P-adic components are 1. This embedding gives a decomposition of J_k as the internal direct product of J_k^0 and $\mathbb{R}^+$. Define J^{S_∞} to be the subgroup of J_k consisting of all ideles whose archimedean components are 1 and whose P-adic components are P-adic units. Let π be the projection from J_k to $J_k/(\mathbb{R}^+ k^* J^{S_\infty})$. The quotient topology induced on $J_k/(\mathbb{R}^+ k^* J^{S_\infty})$ as a quotient of J_k is the same as that induced on $J_k/(\mathbb{R}^+ k^* J^{S_\infty})$ as a quotient of the idele class group, and both π and the projection π_2 from the idele class group are continuous with respect to this topology. It follows from previous remarks that $\pi(J_k^0) = \pi_2(J_k^0/k^*) = J_k/(\mathbb{R}^+ k^* J^{S_\infty})$ is a compact topological group.

If G is a compact topological group, then a *character* of G is a continuous group homomorphism from G to the unit circle in the complex plane. The definition of equidistribution is given in full generality in ([4], pp. 315–316), but it will only be stated here in the context of prime ideals of the ring of integers of k. Define $\tau : \mathcal{P} \longrightarrow J_k$ as follows. For each prime ideal P, select an element $\gamma_P \in k_{v_P}^*$ that generates the prime ideal in $\mathcal{O}_{v_P}$, and define $\tau(P)$ to be the idele with γ_p in the v_Pth component and 1 in all other components. Let $\mathcal{P}_r$ denote the set of prime ideals P such that $N_{k/\mathbb{Q}}(P) \leq r$. If λ is a map from J_k to a compact commutative group G, then $\mathcal{P}$ is $\lambda \circ \tau$*-equidistributed* in G if

$$\lim_{r \to \infty} \frac{1}{\#(\mathcal{P}_r)} \sum_{\psi \in \mathcal{P}_r} \chi \circ \lambda \circ \tau(\psi) = \int_G \chi \tag{2}$$

for all characters χ of G. The measure on G is the unique Haar measure μ with $\mu(G) = 1$. The only property of Haar measure that will be used explicitly is that $\mu(gU) = \mu(U)$ for all Borel-measurable sets U and $g \in G$. See [2] for an in-depth exposition of Haar measure.

If $\mathcal{P}$ is $\lambda \circ \tau$-equidistributed in G, then equation (2) holds if χ is replaced by any integrable function on G, where an integrable function is as defined in ([4], p. 316). The next section will take for granted the fact that the characteristic function on an open set is integrable.

The following theorem due to Hecke (see [4], p. 317) gives a criterion for $\mathcal{P}$ to be $\lambda \circ \tau$-equidistributed in G. It follows from this theorem that $\mathcal{P}$ is $\pi \circ \tau$-equidistributed in $J_k/(\mathbb{R}^+ k^* J^{S_\infty})$.

Theorem 7. *If G is a compact commutative group and $\lambda : J_k \rightarrow G$ is a continuous homomorphism such that $\lambda(J_k^0) = G$ and $\lambda(k^*) = \{1\}$, then $\mathcal{P}$ is $\lambda \circ \tau$-equidistributed in G.*

The next section will prove a theorem on the distribution of principal prime ideals in $\mathbb{Z}[\zeta_m]$ where m is a power of 2 or $m = 6$.

8. The distribution of principal prime ideals in cyclotomic fields

The following theorem will be critical in proving that the upper bound derived in Section 2 is optimal if m is a power of 2 or $m = 6$.

Theorem 8. *If m is a power of 2 or $m = 6$, then for all $\epsilon > 0$, there is $\gamma = \sum_{j=0}^{\phi(m)-1} \gamma_j \zeta_m^j \in \mathbb{Z}[\zeta_m]$ such that:*

1. *γ generates a prime ideal in $\mathbb{Z}[\zeta_m]$,*
2. *$N_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\gamma) = p$, where p is prime and $p \equiv 1 \bmod m$, and*
3. *if $\beta = p/\gamma = \sum_{j=0}^{\phi(m)-1} b_j \zeta_m^j$, then $|b_j/b_0| < \epsilon$ for $1 \leq j \leq \phi(m) - 1$.*

There is also a γ satisfying the first two conditions together with the condition that $b_0 > 0$ and $1 - \epsilon < b_j/b_0 < 1 + \epsilon$ for $1 \leq j \leq m/2 - 1$.

The first step in proving the theorem is to find an open set $U_\epsilon \subset J_k/(\mathbb{R}^+ k^* J^{S_\infty})$ with the property that there is a γ satisfying conditions 1 and 3 in the theorem if $\pi \circ \tau(P)$ is in U_ϵ. The equidistribution criterion together with some additional information on the distribution of primes will then show that there is a γ that satisfies all the conditions of the theorem. The proof for the alternate third condition is very similar.

To find U_ϵ if m is a power of 2, let $\sigma_0, \ldots, \sigma_{\phi(m)/2-1}$ be the embeddings of k into $\mathbb{C}$ defined by $\sigma_i(\zeta_m) = \zeta_m^{z_i}$, where $1 = z_0 < \cdots < z_{\phi(m)/2-1} = m/2 - 1$ are relatively prime to m. These embeddings induce all of the archimedean absolute values on k, and give a metric d on $k^{\phi(m)/2}$ defined by $d(c, c') = \sqrt{\sum_{j=0}^{\phi(m)/2-1} |\sigma_j(c_j - c_j')|^2}$, where c_j and c_j' are the jth components of c_j and c_j', respectively. This metric extends to a metric d on $\mathbb{C}^{\phi(m)/2}$, and d generates the topology on $\mathbb{C}^{\phi(m)/2}$ as a subset of J_k. Consider $\mathbb{C}$ as being embedded in $\mathbb{C}^{\phi(m)/2}$ along the diagonal, and suppose that $c = \sum_{i=0}^{\phi(m)-1} q_i \zeta_m^i \in k^*$ with $d(0, c) < \eta/\sqrt{2\phi(m)}$, where $\eta > 0$ is chosen so that $\eta/(1 - \eta) < \epsilon$ and $(1 - \eta)/(1 + \eta) > 1 - \epsilon$. The following lemma will put a bound on $|q_i|$.

Lemma 8.1. *If c is as above, and m is a power of 2 with $m \geq 4$, then $|q_i| < \eta$ for $0 \geq i \geq \phi(m) - 1$.*

Proof. First of all, note that $j + m/2$ is relatively prime to m if j is relatively prime to m. This implies that $\mathrm{Tr}_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\zeta_m) = 0$, and hence that $\mathrm{Tr}_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\zeta_m^i) = 0$ if $i \not\equiv 0 \bmod m/2$, since ζ_m^i is a primitive $m/\gcd(m, i)$th root of unity. It then follows from a straightforward computation that $\mathrm{Tr}_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(c_j \overline{c_j}) = \phi(m) \sum_{i=0}^{\phi(m)-1} q_i^2$. If $\sigma \in \mathrm{Gal}(\mathbb{Q}(\zeta_m)/\mathbb{Q})$, then $\sigma = \sigma_j$ or $\sigma = \overline{\sigma_j}$ for some j. The lemma now follows from the following calculation.

$$\begin{aligned}
|q_i| &\leq \sqrt{\sum_{j=0}^{\phi(m)-1} q_i^2} \\
&= \sqrt{\phi(m)\mathrm{Tr}_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(c_j\overline{c_j})} \\
&= \sqrt{2\phi(m) d^2(0, 1/c)} \\
&< \eta.
\end{aligned}$$

Let $U = \{c \mid d(c,1) < \eta/\sqrt{2\phi(m)}\}$ and let U_ϵ be the open set $\pi(U \times \prod_v \mathcal{O}_v^*)$, where the product ranges over all P-adic absolute values. The process is the same for the alternate third condition, except that $U = \{c \mid d(c, \sum_{j=0}^{\phi(m)-1} \zeta_m^j) < \eta/\sqrt{2}\}$. If $m = 6$, then every $c \in \mathbb{C}$ has a unique representation of the form $r_1 + r_2\zeta_m$ with $r_1, r_2 \in \mathbb{R}$, and so $U = \{c \mid |r_1/r_2| < \epsilon\}$, or $U = \{c \mid |r_1 - 1| < \eta, |r_2 - 1| < \eta\}$ for the alternate third condition.

If $\pi \circ \tau(P) \in U_\epsilon$, then it is possible to write $\tau(P) = \gamma r j^S u$, where $\gamma \in k^*, r \in \mathbb{R}^+, j^S \in J^{S_\infty}$, and $u \in U$. An examination of the components shows that γ is a prime element in the P-adic completion and a P'-adic unit for all $P' \in \mathcal{P}$ that are different from P. It follows that $\gamma \in \mathbb{Z}[\zeta_m]$ and that γ generates P. Since U is an open set, it is possible to find $q \in \mathbb{Q}$ sufficiently close to r so that $1/\gamma = qu'$ with $u' \in k^* \cap U$. If $u' = \sum_{j=0}^{\phi(m)-1} q_j \zeta_m^j$, then $|q_0 - 1| < \eta$ and $|q_j| < \eta$ for $1 \leq j \leq \phi(m) - 1$ by the previous lemma, so that $|q_j|/|q_0| < \epsilon$ for $1 \leq j \leq \phi(m) - 1$. This shows that there is a γ satisfying conditions 1 and 3 if $\pi \circ \tau(P) \in U_\epsilon$. The proof of the equivalent statement for the alternate third condition is the same. Note that it is possible to assume without loss of generality that the q_j's are positive in this case, since the previous lemma shows that they are close to 1. The same statements follow in a straightforward manner if $m = 6$.

If χ is the characteristic function on U_ϵ, meaning that χ is 1 on U_ϵ and 0 outside of U_ϵ, then equation (2) becomes the following equation:

$$\lim_{r \to \infty} \frac{\#(\mathcal{P}_r \cap U_\epsilon)}{\#(\mathcal{P}_r)} = \mu(U_\epsilon).$$

The following lemma will show that there are infinitely many primes P that are in U_ϵ.

Lemma 8.2. *Suppose G is a compact topological group with the unique Haar measure μ such that $\mu(G) = 1$. If U is a non-empty open subset of G, then $\mu(U) > 0$.*

Proof. If G is a topological group, then $\bigcup_{g \in G} gU$ is an open cover of G, and so G can be written in the form $G = \bigcup_{j=0}^n g_j U$ for some finite set $\{g_0, \ldots, g_n\} \subset G$. It follows that $\mu(G) = 1 \leq n\mu(U)$ since μ is Haar measure, and so $\mu(U) \geq 1/n > 0$.

The next lemma will show that there must be infinitely many primes P such that $P \in U_\epsilon$ and P satisfies the second condition. This will complete the proof of the theorem.

Lemma 8.3. *If $\mathcal{P}^p$ is the set of prime ideals with prime norm, then*

$$\lim_{r \to \infty} \frac{\#(\mathcal{P}_r \cap \mathcal{P}^p)}{\#(\mathcal{P}_r)} = 1.$$

Proof. If $N_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(P) = p$ with p prime, then there are $\phi(m)$ prime ideals lying above p and $p \equiv 1 \bmod m$. In general, $N_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(P) = p^{o(p)}$, and there are $\phi(m)/o(p)$ ideals lying over p, where $P \cap \mathbb{Z} = p\mathbb{Z}$, and $o(p)$ is the multiplicative order of $p \bmod m$. The above limit then becomes the following:

$$\lim_{r \to \infty} \frac{\phi(m)\#\{p \mid p \text{ prime}, p < r, p \equiv 1 \bmod m\}}{\sum_{j \in (\mathbb{Z}/m\mathbb{Z})^*} \frac{\phi(m)}{o(j)} \#\{p \mid p \text{ prime}, p \equiv j \bmod m, p < r^{1/o(j)}\}}.$$

Divide both the numerator and denominator of this fraction by

$$\#\{p \mid p \text{ prime}, p < r\}.$$

By Dirichlet's theorem on the distribution of primes in arithmetic progressions (see [7], p. 31), the numerator and the $1 \bmod m$ component of the denominator tend to 1 as $r \to \infty$, while the other components of the denominator tend to 0. To see this fact for the other components, divide the numerator and denominator of the component by $\#\{p \mid p \text{ prime}, p < r^{1/o(j)}\}$. By Dirichlet's theorem, the numerator tends to $1/o(j)$, and the denominator tends to $r^{1-1/o(j)}/o(j)$ by the Prime Number Theorem. The lemma now follows.

The next section will be devoted to the application of Theorem 8 to the problem of finding optimal possible bounds for $M_1(m,p)$.

9. What are the optimal bounds for $M_1(m,p)$?

If m is a power of 2, then it was proven in Section 2 that $K_m \leq 1$, where K_m is as defined in Section 1. The following theorem shows that equality holds.

Theorem 9. *If m is a power of 2, then $K_m = 1$.*

Suppose without loss of generality that $\gamma\mathbb{Z}[\zeta_m] = P_{m-1}$, where γ is as in Theorem 8 and P_j is as defined in Section 2. It is clear that $\beta = p/\gamma$ is an element of $\mathbb{Z}[\zeta_m]$, and it follows from Theorem 8 that $|b_j/b_0| < \epsilon$ for $1 \leq j \leq \phi(m) - 1$. The next lemma shows that there is a coset C of μ_m in $(\mathbb{Z}/p\mathbb{Z})^*$ with $C = \{\pm b_0, \ldots, \pm b_{m/2-1}\}$.

Lemma 9.1. *If m is a power of 2, $\beta \in \prod_{j=0}^{m/2-2} P_{2j+1}$, and $\beta \notin P_{m-1}$, then $b_i \equiv b_0 t^i \bmod p$ for $1 \leq i \leq m/2-1$, and $b_0 \not\equiv 0 \bmod p$.*

Proof. It is a consequence of the definition of P_j that $\zeta_m \equiv t^j \bmod P_j$, and hence that $\sum_{i=0}^{m/2-1} b_i t^{ij} \equiv 0 \bmod p$ for all odd numbers j with $j \not\equiv -1 \bmod m$. This means that the vector $v = (b_0 \bmod p, \ldots, b_{m/2-1} \bmod p)$ is in the nullspace of an $(m/2-1) \times m/2$ Vandermonde matrix A with $\text{nullity}(A) = 1$. A geometric series computation shows that $w = (1 \bmod p, t \bmod p, \ldots, t^{m/2-1} \bmod p)$ is in the nullspace of A, and so v is a scalar multiple of w over $\mathbb{Z}/p\mathbb{Z}$. It follows that $b_i \equiv b_0 t^i \bmod p$ for $1 \leq i \leq m/2-1$, and $b_0 \not\equiv 0 \bmod p$ since $\beta \notin P_{m-1}$.

Suppose that C' is a coset of μ_m in $(\mathbb{Z}/p\mathbb{Z})^*$. Let $\beta' = \sum_{j=0}^{m/2-1} b'_j \zeta_m^j$, where $b'_0 \in C'$ and $b'_j \equiv b'_0 t^j$ for $1 \leq j \leq m/2-1$. The same argument as in Lemma 1.2 shows that $\beta' \in P_j$ if $j \neq m-1$, and so $\beta' = c\beta$ for some $c = \sum_{j=0}^{m/2-1} c_j \zeta_m^j \in \mathbb{Z}[\zeta_m]$. Suppose that $|c_k| = \max_{0 \leq j \leq m/2-1} |c_j|$. The calculation below gives a lower bound for $|C'|$ in terms of $|b_0|$:

$$\begin{aligned} |C'| &\geq |b'_k| \\ &= \Big|\sum_{j=0}^{k} b_j c_{k-j} - \sum_{j=k+1}^{m/2-1} b_j c_{m/2+k-j}\Big| \\ &\geq |b_0 c_k| - \sum_{j=1}^{k} |b_j c_{k-j}| - \sum_{j=k+1}^{m/2-1} |b_j c_{m/2+k-j}| \\ &\geq |b_0 c_k|(1 - (m/2-1)\epsilon) \\ &\geq |b_0|(1 - (m/2-1)\epsilon). \end{aligned}$$

It also follows that $|b_0| \geq (\frac{1}{1+(m/2-1)\epsilon})p^{1-2/m}$, since

$$\begin{aligned} |N_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\beta)| &= p^{m/2-1} \\ &= \prod_{j=0}^{m/2-1} |\sum_{l=0}^{m/2-1} b_l \zeta_m^{l(2j+1)}| \\ &\leq \prod_{j=0}^{m/2-1} \sum_{l=0}^{m/2-1} |b_l| \\ &\leq (|b_0|(1+(m/2-1)\epsilon))^{m/2}. \end{aligned}$$

Combining the inequalities for $|C'|$ and $|b_0|$ gives

$$M_1(m,p) \geq \left(\frac{1-(m/2-1)\epsilon}{1+(m/2-1)\epsilon}\right) p^{1-2/m}.$$

Letting $\epsilon \to 0$ proves Theorem 9.

Let κ_m be the supremum of all κ such that $M_1(m,p) \geq \kappa p^{1-1/\phi(m)}$ for all primes p. Section 2 gives a lower bound for κ_m, and the following theorem applies Theorem 8 to give an upper bound in the case where m is a power of 2. This theorem proves that the upper bound $M_1(4,p) \geq (\sqrt{2}/2)p^{1/2}$ proven in Section 2 is optimal.

Theorem 10. *If m is a power of 2, then $\kappa_m \leq \frac{1}{2^{1-2/m}}$.*

Proof. Let β satisfy the alternate third condition in Theorem 8. Write $\beta = b_0(\sum_{j=0}^{m/2-1} \zeta_m^j + \sum_{j=0}^{m/2-1} d_j \zeta_m^j)$, where $|d_j| < \epsilon$ for $0 \leq j \leq m/2-1$, and note that $\sum_{j=0}^{m/2-1} \zeta_m^j = 2/(1-\zeta_m)$ with $N_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(2/(1-\zeta_m)) = 2^{m/2-1}$. An upper bound will be placed on $|b_0|$ by using the fact that $p^{m/2-1} = |N_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\beta)|$. The right-hand side of this equality expands as

$$|b_0|^{m/2}| \prod_{i=0}^{m/2-1} 2/(1-\zeta_m^{2i+1}) + \sum_{j=0}^{m/2-1} c_j \zeta_m^{(2i+1)j}|.$$

One term in the product is $N_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(2/(1-\zeta_m))$, and each of the other $2^{m/2}-1$ terms has absolute value that is bounded above by

$$f(m)\epsilon = (1/(1-\cos(2\pi/m)))^{m/2-1} m\epsilon.$$

The cosine term comes from the absolute value of $|1-\zeta_m|$ and the $m\epsilon$ term is an upper bound for $\sum_{j=0}^{m/2-1} c_j \zeta_m^{ij}$. Putting these bounds together gives the inequality

$$p^{m/2-1} \geq |b_0|^{m/2}|2^{m/2-1} - (2^{m/2}-1)f(m)\epsilon|,$$

or

$$|b_0| \leq \left(\frac{1}{|2^{m/2-1} - (2^{m/2}-1)f(m)\epsilon|^{2/m}}\right) p^{1-2/m}.$$

Lemma 9.1 shows that there is a coset C with $C = \{\pm b_0, \ldots, \pm b_{m/2-1}\}$, and so $M_1(m,p) \leq |C| \leq (1+\epsilon)|b_0|$. Letting $\epsilon \to 0$ proves the theorem.

If $m = 2^e 3^f$ with $e, f > 0$, then the upper bound

$$M_1(m,p) \leq (2/\sqrt{3})p^{1-m/3}$$

proven in Section 2 is the same as the bound proven by Robinson. This suggests that perhaps this bound is optimal. In general, this seems difficult to prove, in part because it seems difficult to extend Theorem 8 to this case. Theorem 8 does apply, however, when $m = 6$, and plays a crucial role in the following result, which shows that the upper bound from Section 2 is optimal.

Theorem 11. *If K_m and κ_m are as previously defined, then $K_6 = 2/\sqrt{3}$ and $\kappa_6 \leq 1$.*

Proof. Pick β and γ that satisfy the alternate third condition in Theorem 8, and suppose without loss of generality that $\gamma\mathbb{Z}[\zeta_m] = P_{m-1}$, so that $\beta\mathbb{Z}[\zeta_m] = P_1$. It follows that $b_0 + b_1 t \equiv 0 \bmod p$, and hence that $b_1 \equiv b_0 t^2 \bmod p$. This means that $b_0 + b_1 \equiv b_0(1+t^2) \equiv b_0 t \bmod p$, which gives a coset $C = \{\pm b_0, \pm(b_0+b_1), \pm b_1\}$. Note that

$$N_{\mathbb{Q}(\zeta_m)/\mathbb{Q}}(\beta) = p = b_0 + b_0 b_1 + b_1^2 \leq b_0^2(1 + (1+\epsilon) + (1+\epsilon)^2),$$

which implies that $b_0 + b_1 \geq (2-\epsilon)\left(\frac{p}{1+(1+\epsilon)+(1+\epsilon)^2}\right)^{1/2}$.

Suppose that C' is another coset with $C' = \{\pm b_0', \pm(b_0'+b_1'), \pm b_1'\}$, where $b_1 \equiv b_0 t^2 \bmod p$. It follows that $b_0' + b_1'\zeta_m \in P_1$, and so

$$b_0' + b_1'\zeta_m = (c_0 + c_1\zeta_m)(b_0 + b_1\zeta_m)$$

for some $c_0, c_1 \in \mathbb{Z}$. If $c_1 = 0$, then $|b_0' + b_1'| = |c_0||b_0 + b_1| \geq |b_0 + b_1|$. Suppose now without loss of generality that $c_1 > 0$. If $c_0 = 0$, then $|b_1'| = c_1|b_0 + b_1|$. If $c_0 > 0$, then $|b_1'| = (c_0+c_1)b_1 + c_1 b_0 > b_0 + b_1$, and $c_0 < 0$ implies that $|b_0'| = |c_0 b_0 - c_1 b_1| > b_0 + b_1$. It now follows that

$$|C'| \geq |C| \geq (2-\epsilon)\left(\frac{p}{1+(1+\epsilon)+(1+\epsilon)^2}\right)^{1/2},$$

which proves the first part of the theorem.

To prove the second part of the theorem, let γ, β, and C be as before, except that β satisfies the regular third condition of Theorem 8. Under these circumstances, it follows that $|C| \leq (1+\epsilon)|b_0|$ and that

$$p = b_0^2 + b_0 b_1 + b_1^2 \leq b_0^2(1 + \epsilon + \epsilon^2).$$

Solving for b_0 gives $|C| \leq \frac{1+\epsilon}{\sqrt{1+\epsilon+\epsilon^2}}\sqrt{p}$, which proves the second part of the theorem.

In general, the bounds from Section 2 can be improved by finding the volume of $S_{1/2} \cap V$ exactly instead of using the Vaaler estimate. It is doubtful, however, that this improvement will lead to an optimal bound.

Acknowledgment

The author would like to thank Hendrik Lenstra for his insights and suggestions, and would like to acknowledge the late Rafael Robinson, whose work forms the basis for this paper.

References

1. J.W.S. Cassels and A. Frohlich, *Algebraic Number Theory*, Academic Press Limited, 1967. MR **35:**6500

2. John L. Kelley and T.P. Srinivasan, *Measure and Integral*, Springer-Verlag, 1988. MR **89e:**28001

3. Sergey Konyagin and Igor Shparlinski, *On the Distribution of Residues of Finitely Generated Multiplicative Groups and Some of Their Applications*, to appear.
4. S. Lang, *Algebraic Number Theory*, Springer-Verlag, 1994. MR **95f:**11085
5. C.G. Lekkerkerker, *Geometry of Numbers*, Wolters-Noordhoff and North-Holland Publishing Companies, 1969. MR **42:**5915
6. R.M. Robinson, *Numbers Having m Small mth Roots mod p*, Mathematics of Computation **61** (1993), no. 203, 393–413. MR **93k:**11002
7. P. Stevenhagen and H.W. Lenstra, Jr., *Chebotarev and his density theorem*, Math. Intelligencer **18** (1996), 26–37. CMP 96:14
8. J.E. Vaaler, *A Geometric Inequality with Applications to Linear Forms*, Pacific Journal of Mathematics **83** (1979), no. 2, 543–553. MR **81d:**52007

Department of Mathematics, University of California at Berkeley, Berkeley, California 94720