MATHEMATICS OF COMPUTATION
VOLUME 61, NUMBER 203
JULY 1993, PAGES 393–413

# NUMBERS HAVING $m$ SMALL $m$th ROOTS mod $p$

RAPHAEL M. ROBINSON

*In memory of my good friend Derrick H. Lehmer*

ABSTRACT. Here are two typical results about the numbers mentioned in the title: If $p$ is a prime such that $p \equiv 1 \pmod 6$ and $p > 67$, then there are exactly six numbers mod $p$, each of which has six sixth roots less than $2\sqrt{3p}$ in absolute value. If $p$ is a prime such that $p \equiv 1 \pmod 8$, then there is at least one number mod $p$ which has eight eighth roots less than $p^{3/4}$ in absolute value.

## 1. INTRODUCTION

Let $m$ be a positive integer, and let $p$ be an odd prime. In order that the congruence

$$x^m \equiv a \pmod p$$

should ever have $m$ solutions, we need that

$$p \equiv 1 \pmod m.$$

In this case, there are $(p-1)/m$ values of $a \pmod p$ other than 0 for which the congruence is solvable, and there are $m$ solutions in each case. We may represent these solutions by their absolutely least residues, so that all lie in the interval $-p/2 < x < p/2$. Such a set will be called a reduced set of $m$th roots.

We are interested in finding such reduced sets of $m$th roots which lie in as small an interval $-c \le x \le c$ as possible. Some results on this problem were presented at a meeting in 1984 and appear in the abstract [6], but the proofs have not previously been published.

The problem is trivial for $m = 1$ and $m = 2$. There is also a very simple result for $m = 4$. It is known that any prime $p \equiv 1 \pmod 4$ can be written as a sum of two squares, $p = x^2 + y^2$. Since $y^2 \equiv -x^2 \pmod p$, we have $y^4 \equiv x^4$, so that the number $x^4$ has four fourth roots $\pm x, \pm y$, all of which are less than $\sqrt{p}$ in absolute value. If I had thought of this example to begin with, it might have been the starting point for my investigations. But the actual starting point involved sixth roots, and is described at the end of §4.

Notice that if $m$ is odd, and the solutions to $x^m \equiv a \pmod p$ are $x_1, x_2, \ldots, x_m$, then the solutions to $x^m \equiv -a \pmod p$ are $-x_1, -x_2, \ldots, -x_m$, and the solutions to $x^{2m} \equiv a^2 \pmod p$ are $\pm x_1, \pm x_2, \ldots, \pm x_m$. The maximum absolute value is the same in all cases. Thus, each number with all its

Received by the editor April 16, 1992.
1991 *Mathematics Subject Classification.* Primary 11A07, 11A15, 11R18; Secondary 11L10.


0025-5718/93 $1.00 + $.25 per page

$(2m)$th roots in $-c \le x \le c$ corresponds to two numbers with their $m$th roots in the interval. Hence the problem for $m$ is equivalent to the problem for $2m$ when $m$ is odd. In particular, the results for $m = 3$ follow from those for $m = 6$ given in §3.

The $p-1$ numbers with $0 < |x| < p/2$ fall into $(p-1)/m$ sets having equal $m$th powers. If $m$ is even, then the maximum absolute values of the various sets are all different. For $1 \le i \le (p-1)/m$, let $M_i(p)$ be the $i$th smallest maximum. (This also depends on $m$, but we omit this from the notation.) The $i$th set will contain both $M_i(p)$ and $-M_i(p)$. But if $m$ is odd, we will let $M_i(p)$ have the same meaning as for $2m$, so that it will be defined only for $1 \le i \le (p-1)/2m$. There will now be two sets with the maximum absolute value $M_i(p)$, one containing $M_i(p)$ and the other containing $-M_i(p)$.

We showed above that $M_1(p) < \sqrt{p}$ when $m = 4$. Detailed information about $M_i(p)$ for $m = 4$ and $m = 6$ is given in §§2 and 3. A connection with Jacobsthal sums is explained in §4. Upper bounds for $M_1(p)$ for other values of $m$ are discussed in §§5 and 6.

## 2. Fourth roots

We shall study sets of small fourth roots mod $p$, where $p$ is a prime with $p \equiv 1 \pmod 4$, by finding a connection between multiplying complex numbers by $i$ and multiplying integers by a primitive fourth root $t$ of 1 mod $p$.

Let $z$ be an element of the ring $\mathbb{Z}[i]$, and put $z = x + yi$, where $x$ and $y$ are integers. Then the points $i^n z$ for $n = 0, 1, 2, 3$ have the coordinates $(x, y), (-y, x), (-x, -y), (y, -x)$. Notice that the abscissas and ordinates run through the same four numbers $\pm x, \pm y$.

Now suppose that $z$ lies on the circle $|z|^2 = hp$, where $h$ is a positive integer. Then $x^2 + y^2 = hp$. If $x \equiv 0 \pmod p$, then $y \equiv 0 \pmod p$, and hence $h \equiv 0 \pmod p$. If $x \not\equiv 0 \pmod p$, then we can solve

$$tx \equiv -y \pmod p$$

for $t$. This yields

$$0 \equiv x^2 + y^2 \equiv x^2(1 + t^2) \pmod p;$$

hence $t^2 \equiv -1 \pmod p$, so that $t$ is a primitive fourth root of 1 mod $p$. Notice that $ty \equiv -t^2 x \equiv x \pmod p$. Thus,

$$iz \equiv tz \pmod p.$$

The points $i^n z$ form the vertices of a square. The abscissas or ordinates constitute sets of fourth roots mod $p$.

Will every reduced set of fourth roots mod $p$ be equal to such a set of ordinates? Let $y$ be one of the given fourth roots. Since $p \equiv 1 \pmod 4$, we can solve $x^2 \equiv -y^2 \pmod p$, and can make $|x| < p/2$. Then the point $z = x + yi$ will lie on the circle $|z|^2 = hp$ for some integer $h$ with $0 < h < p/2$. Draw the inscribed square with a vertex at $z$. The ordinates of the vertices form a reduced set of fourth roots. Since this set includes $y$, it must be identical to the given set. So every reduced set of fourth roots is a set of ordinates.

We want to know what squares can be inscribed in the circle $|z|^2 = hp$. We first ask about squares in $|z|^2 = h$. If one vertex is at $z = x + yi$, then we need

$$(x + yi)(x - yi) = h.$$

Each prime power in $h$ must be split into two conjugate factors in $\mathbb{Z}[i]$. It is known that primes $q \equiv 3 \pmod 4$ remain prime in $\mathbb{Z}[i]$, that 2 is equivalent to the square of a prime, and that primes $p \equiv 1 \pmod 4$ split into two conjugate prime factors which are not equivalent. Thus, primes $q \equiv 3 \pmod 4$ must appear in $h$ to an even power. The only choices occur for primes $q \equiv 1 \pmod 4$. If the power $q^\nu$ occurs in $h$, then we can decide in how many of the $\nu$ cases we choose the first factor of $q$ as a factor of $x + yi$, and then we must choose the second factor in the remaining cases. This gives $\prod(\nu+1)$ choices for $x+yi$ when units are ignored, hence $\prod(\nu + 1)$ inscribed squares.

If $h$ is not a multiple of $p$, then there will be twice as many squares in $|z|^2 = hp$, or $\prod(\nu + 1)$ pairs of conjugate squares. The early possible values of $h$ and the corresponding number of pairs are as follows:

$$h = 1, 2, 4, 5, 8, 9, 10, 13, 16, 17.$$
$$\text{Pairs} = 1, 1, 1, 2, 1, 1, 2, 2, 1, 2.$$

Conjugate squares produce the same set of ordinates. No other square can produce this set. Indeed, every set of ordinates has the form $\{\pm x, \pm y\}$. To produce this set of ordinates, the abscissa corresponding to the ordinate $y$ can only be $\pm x$.

The maximum ordinate for a square inscribed in $|z|^2 = hp$ lies between $\sqrt{hp/2}$ and $\sqrt{hp}$. These intervals are nonoverlapping for $h = 1, 2, 4$. None of these values of $h$ can be a multiple of $p$. Since each of the three corresponding circles contains one pair of conjugate squares, this tells us that there is exactly one set of ordinates less than $\sqrt{p}$, exactly two sets less than $\sqrt{2p}$, and at least three sets less than $2\sqrt{p}$. In general, the intervals overlap, so that it is not possible to give exact counts. Now $\sqrt{2p} < p/2$ if $p > 8$, and $2\sqrt{p} < p/2$ if $p > 16$, so that the sets of ordinates give reduced sets of fourth roots in these cases. When $p = 13$, all three reduced sets of fourth roots are less than $2\sqrt{p}$. Hence, for $p \equiv 1 \pmod 4$ there is always exactly one reduced set of fourth roots less than $\sqrt{p}$, and when $p > 5$ there are exactly two sets less than $\sqrt{2p}$ and at least three sets less than $2\sqrt{p}$.

For $1 \le i \le (p-1)/4$, we let $M_i(p)$ be the $i$th smallest maximum of a reduced set of fourth roots. This will be studied by its relation to $N_i(p)$, defined for all $i \ge 1$ as the $i$th smallest maximum of a set of ordinates. Since the reduced sets of fourth roots are identical with the sets of ordinates which are less than $p/2$, we see that $M_i(p) = N_i(p)$ whenever $M_i(p)$ is defined. We will have $N_i(p) < p/2$ just when $i \le (p-1)/4$.

The prime $p \equiv 1 \pmod 4$ splits in the ring $\mathbb{Z}[i]$ in the form $p = P\overline{P}$, where $P$ is a prime in the ring. The squares inscribed in $|z|^2 = hp$ are obtained from the squares inscribed in $|z|^2 = h$ by multiplying by $P$ or $\overline{P}$. By a basic square, we shall mean a square with center at 0 and vertices in the ring. If we start with all the basic squares, and expand and rotate by the factors $P$ and $\overline{P}$, we obtain all the squares inscribed in circles $|z|^2 = hp$. But we do not need both a square and its conjugate, so it is enough to use the factor $P$. This causes an expansion by the factor $\sqrt{p}$, and a rotation. Hecke [2, §9] proved that there are primes $P$ in every sector. Hence, when we consider all primes $p \equiv 1 \pmod 4$, the possible rotations caused by multiplying by $P$ are everywhere dense.

Let $\alpha_i$ be the smallest positive number so that for some rotation of the

complex plane, there will be $i$ of the basic squares lying in the strip $y^2 \leq \alpha_i$. Let $\beta_i$ be the smallest positive number so that for every rotation of the complex plane there will be $i$ of the basic squares lying in the strip $y^2 \leq \beta_i$. Then it follows from the above construction that

$$\alpha_i \leq N_i(p)^2/p \leq \beta_i,$$

and that the bounds $\alpha_i$ and $\beta_i$ are the best that will hold for all values of $p$.

The bound $\alpha_i$ will be obtained for a rotation of the plane which places more than one vertex of the squares used at the maximum height. For otherwise, a small rotation would reduce the maximum height. The bound $\beta_i$ will be obtained in a similar case, or when there is just one vertex with maximum height, but it lies on the imaginary axis. For if there were just one vertex at maximum height, and it were not on the imaginary axis, a small rotation would increase the maximum height. On the other hand, if there are two vertices at the maximum height, we may not need both squares, since there may be $i+1$ squares in the strip used. Even if a rotation increases the height of one of the vertices, it may decrease the maximum height used.

In writing a program to compute $\alpha_i$ and $\beta_i$, it is more convenient to consider supporting lines with various orientations rather than rotating the plane. We need to know what possible supporting lines must be examined. This depends on knowing what basic squares we need to consider.

Notice that every basic square whose vertices satisfy

$$\max(|x|, |y|) \leq j$$

is contained in every basic square whose vertices satisfy

$$\max(|x|, |y|) \geq 2j.$$

Indeed, the vertices of the latter square will be at a distance $\geq 2j$ from the origin; hence its sides will be at a distance $\geq (\sqrt{2})j$ from the origin.

There are $(2j+1)^2$ lattice points satisfying the first condition, hence $j(j+1)$ basic squares having these points as vertices. Thus, the values of $\alpha_i$ and $\beta_i$ for $i \leq j(j+1)$ will be found without using any squares from the second set. Hence, we need only look at the $2j(2j-1)$ basic squares with

$$\max(|x|, |y|) \leq 2j-1.$$

Table 1 gives the values of $\alpha_i$ and $\beta_i$ for $i \leq 80$, which were computed in this way.

There are two special cases which will be discussed. In the unrotated plane, there are $j(j+1)$ basic squares with ordinates not exceeding $j$. It follows that

$$\alpha_i \leq j^2 \leq \beta_i \quad \text{for } j(j-1) < i \leq j(j+1).$$

Table 1 shows many cases of equality on the right near the beginning of the interval, and on the left near the end. It can be shown that equality holds for at least one value of $i$ at each end:

$$\beta_i = j^2 \quad \text{for } i = j(j-1)+1,$$
$$\alpha_i = j^2 \quad \text{for } i = j(j+1).$$

TABLE 1. Fourth roots

| $i$ | $\alpha_i$ | $\beta_i$ | $i$ | $\alpha_i$ | $\beta_i$ |
|---|---|---|---|---|---|
| 1 | 1/2 | 1 | 41 | 36 | 41 84/101 |
| 2 | 1 | 2 | 42 | 36 | 42 15/17 |
| 3 | 2 | 4 | 43 | 40 1/2 | 49 |
| 4 | 3 1/5 | 4 1/2 | 44 | 40 1/2 | 49 |
| 5 | 4 | 5 | 45 | 40 1/2 | 49 |
| 6 | 4 | 6 2/5 | 46 | 44 1/10 | 50 |
| 7 | 6 3/13 | 9 | 47 | 45 | 50 |
| 8 | 7 2/17 | 9 | 48 | 45 | 50 |
| 9 | 8 | 9 4/5 | 49 | 47 49/61 | 51 1/5 |
| 10 | 8 | 10 6/25 | 50 | 48 1/13 | 51 1/5 |
| 11 | 9 | 12 1/2 | 51 | 49 | 52 9/10 |
| 12 | 9 | 12 4/5 | 52 | 49 | 53 5/41 |
| 13 | 12 1/10 | 16 | 53 | 49 | 54 22/101 |
| 14 | 12 1/2 | 16 | 54 | 49 | 55 7/61 |
| 15 | 12 1/2 | 16 1/5 | 55 | 49 | 57 4/5 |
| 16 | 15 1/17 | 18 | 56 | 49 | 60 1/2 |
| 17 | 16 | 18 | 57 | 55 7/13 | 64 |
| 18 | 16 | 18 8/13 | 58 | 56 1/13 | 64 |
| 19 | 16 | 20 | 59 | 57 3/5 | 64 |
| 20 | 16 | 20 21/41 | 60 | 57 3/5 | 64 |
| 21 | 18 | 25 | 61 | 57 4/5 | 64 |
| 22 | 20 | 25 | 62 | 60 4/17 | 64 4/5 |
| 23 | 21 39/50 | 25 | 63 | 60 1/2 | 64 4/5 |
| 24 | 22 1/2 | 25 | 64 | 60 1/2 | 64 49/50 |
| 25 | 24 1/26 | 25 16/17 | 65 | 60 1/2 | 65 36/53 |
| 26 | 24 1/5 | 26 23/41 | 66 | 60 1/2 | 67 3/5 |
| 27 | 24 1/2 | 28 4/5 | 67 | 64 | 72 |
| 28 | 24 1/2 | 28 9/10 | 68 | 64 | 72 |
| 29 | 25 | 32 | 69 | 64 | 72 1/5 |
| 30 | 25 | 32 | 70 | 64 | 72 1/5 |
| 31 | 28 4/5 | 36 | 71 | 64 | 72 9/10 |
| 32 | 30 10/13 | 36 | 72 | 64 | 73 12/13 |
| 33 | 31 59/82 | 36 | 73 | 71 81/89 | 81 |
| 34 | 32 | 36 | 74 | 72 | 81 |
| 35 | 32 | 36 1/10 | 75 | 72 | 81 |
| 36 | 32 | 37 3/13 | 76 | 72 | 81 |
| 37 | 36 | 40 1/2 | 77 | 72 | 81 |
| 38 | 36 | 40 1/2 | 78 | 72 | 81 |
| 39 | 36 | 40 1/2 | 79 | 77 15/122 | 84 1/2 |
| 40 | 36 | 41 1/41 | 80 | 77 23/26 | 84 1/2 |

Only the second (and easier) of these two formulas will be proved here.

Newman [4] proved that, for any real $s > 0$, a closed square of side $s$ can contain at most $(s+1)^2$ lattice points. The result is an easy consequence of a theorem of Pick [5], which deals with a polygon whose vertices are at lattice points. If the area is $A$, and there are $B$ lattice points on the boundary and $C$ inside, then Pick's theorem states that $A = B/2 + C - 1$; hence $B + C = A + B/2 + 1$. Apply this to the convex hull of the set of lattice points in a closed square of side $s$. Here, $A \le s^2$ and $B \le 4s$, since the perimeter is at most $4s$ and the lattice points have minimum distance 1. Thus, $B + C \le (s+1)^2$. It is easily seen that equality holds if and only if $s$ is an integer, the vertices are at lattice points, and the sides are horizontal and vertical.

In any rotated plane, the square $|x| \le j$, $|y| \le j$ contains at most $(2j+1)^2$ lattice points, or at most $j(j+1)$ basic squares, and equality holds only for a rotation through a multiple of $90°$. No smaller square can contain so many lattice points, so $\alpha_i = j^2$ when $i = j(j+1)$.

The next simplest case is rotation through $45°$. Equivalently, we can consider supporting lines slanted at $45°$. There are $j(j+1)/2$ basic squares with vertices in the square $|x| + |y| \le j$, whose sides are at a distance $j/\sqrt{2}$ from the origin. It follows that

$$\alpha_i \le j^2/2 \le \beta_i \quad \text{for } j(j-1)/2 < i \le j(j+1)/2.$$

Again, there is often equality on the right for several values of $i$ at the beginning of the interval, and on the left at the end. However, it is not true here that there is always at least one such equality. More than half of the entries in Table 1 are supplied by the rotations of $0°$ and $45°$.

Can the bounds $\alpha_i$ and $\beta_i$ be attained for some prime $p$? To compute $\alpha_i$, we rotate the plane so as to minimize the maximum ordinate of the vertices of a set of $i$ basic squares. This rotation will produce more than one vertex with the maximum ordinate. If $N_i(p)^2/p = \alpha_i$, then the rotation used is $P/|P|$, which takes $\overline{P}$ to $\sqrt{p}$. Since there is no lattice point of $\mathbb{Z}[i]$ between 0 and $\overline{P}$, in the rotated lattice there will be no lattice point between 0 and $\sqrt{p}$. Thus, the distance between two vertices at the maximum height is at least $\sqrt{p}$. The convex hull of the $i$ squares has four-fold symmetry, and has four sides of length at least $\sqrt{p}$. The closest that they can come to the origin is $\sqrt{p}/2$, when the convex hull is a square of side $\sqrt{p}$. Hence, $\alpha_i \ge p/4$, and so $N_i(p) \ge p/2$.

If $N_i(p)^2/p = \beta_i$, then the rotation $P/|P|$ either takes more than one vertex of the squares used to a maximum height, or else places a vertex with maximum height on the imaginary axis. The first case again leads to $N_i(p) \ge p/2$. In the second case, the height is at least $\sqrt{p}$, so $\beta_i \ge p$ and $N_i(p) \ge p$.

In any case, we conclude that

$$\alpha_i < N_i(p)^2/p < \beta_i$$

at least when $N_i(p) < p/2$, that is, when $i \le (p-1)/4$. But this means that

$$\alpha_i < M_i(p)^2/p < \beta_i$$

whenever $M_i(p)$ is defined.

## 3. Sixth roots

We shall study sets of small sixth roots mod $p$, where $p$ is a prime with $p \equiv 1 \pmod 6$, by finding a connection between multiplying complex numbers

by $\zeta = e^{2\pi i/6}$ and multiplying integers by a primitive sixth root $t$ of 1 mod $p$.

Let $z$ be an element of the ring $\mathbb{Z}[\zeta]$, and put $z = x + y\zeta$, where $x$ and $y$ are integers. Since $\zeta^2 - \zeta + 1 = 0$, we see that $\zeta z = -y + (x + y)\zeta$. Thus, the points $\zeta^n z$ for $n = 0, 1, \ldots, 5$ have the coordinates

$$(x, y), (-y, x + y), (-x - y, x), (-x, -y), (y, -x - y), (x + y, -x)$$

in the oblique coordinate system being used. Notice that the abscissas and ordinates run through the same six numbers $\pm x, \pm y, \pm(x + y)$.

Now suppose that $z$ lies on the circle $|z|^2 = hp$, where $h$ is a positive integer. Then $(x + y\zeta)(x + y\bar{\zeta}) = hp$, so $x^2 + xy + y^2 = hp$. If $x \equiv 0 \pmod{p}$, then $y \equiv 0 \pmod{p}$, and hence $h \equiv 0 \pmod{p}$. In this case, all abscissas and ordinates are multiples of $p$. If $x \not\equiv 0 \pmod{p}$, then we can solve

$$tx \equiv -y \pmod{p}$$

for $t$. This yields

$$0 \equiv x^2 + xy + y^2 \equiv x^2(1 - t + t^2) \pmod{p},$$

and hence $t^2 - t + 1 \equiv 0 \pmod{p}$, so that $t$ is a primitive sixth root of 1 mod $p$. Notice that $ty \equiv -t^2 x \equiv (1 - t)x \equiv x + y \pmod{p}$. Thus,

$$\zeta z \equiv tz \pmod{p}.$$

The points $\zeta^n z$ form the vertices of a regular hexagon. The abscissas or ordinates constitute sets of sixth roots mod $p$.

Will every reduced set of sixth roots mod $p$ be equal, or at least congruent, to such a set of ordinates? Unlike the case of fourth roots, we cannot always obtain equality, but congruence is possible. Let $y$ be one of the given sixth roots. Since $p \equiv 1 \pmod{6}$, we can solve

$$u^2 \equiv -3y^2 \pmod{p}$$

for $u$, and can then solve $2x + y \equiv u \pmod{p}$ for $x$. It follows that $x^2 + xy + y^2 \equiv 0 \pmod{p}$. Thus, the point $z = x + y\zeta$ lies on the circle $|z|^2 = hp$ for some positive integer $h$. Draw the inscribed hexagon with a vertex at $z$. The ordinates obtained are congruent to the given sixth roots.

We shall now show that a congruent value of $z$ can be chosen so that the six ordinates are all less than $2p/3$ in absolute value. One choice of a fundamental region for the group of translations generated by $z' = z + p$ and $z' = z + p\zeta$ is the hexagon whose sides are the perpendicular bisectors of the segments joining 0 to the points $p\zeta^n$. Every element of $\mathbb{Z}[\zeta]$ is congruent mod $p$ to a point in the hexagon. The circumradius of the hexagon is $p/\sqrt{3}$. Hence, if $|z|^2 = hp$ for a point in the hexagon, we will have $h < p/3$. The ordinate $y$ of a point is $2/\sqrt{3}$ times its height, so the maximum $y$ in the hexagon is $2p/3$. The same is true for the maximum $x$. Thus, we only need to choose $z$ in the hexagon.

From this, it follows that a set of sixth roots which are all less than $p/3$ in absolute value is equal to a set of ordinates. For the congruent numbers which are less than $2p/3$ in absolute value will differ from the given numbers by less than $p$, and hence be equal to them.

We want to know what hexagons can be inscribed in the circle $|z|^2 = hp$. We first ask about hexagons in $|z|^2 = h$. If one vertex is at $z = x + y\zeta$, then we need

$$(x + y\zeta)(x + y\overline{\zeta}) = h.$$

Each prime power in $h$ must be split into two conjugate factors in $\mathbb{Z}[\zeta]$. It is known that 2 and primes $q \equiv 5 \pmod 6$ remain prime in $\mathbb{Z}[\zeta]$, that 3 is equivalent to the square of a prime, and that primes $q \equiv 1 \pmod 6$ split into two conjugate prime factors which are not equivalent. Thus, 2 and primes $q \equiv 5 \pmod 6$ must appear in $h$ to an even power. The only choices occur for primes $q \equiv 1 \pmod 6$. If the power $q^\nu$ occurs in $h$, then we can decide in how many of the $\nu$ cases we choose the first factor of $q$ as a factor of $x + y\zeta$, and then we must choose the second factor in the remaining cases. This gives $\prod(\nu + 1)$ choices for $x + y\zeta$ when units are ignored, hence $\prod(\nu + 1)$ inscribed hexagons.

If $h$ is not a multiple of $p$, then there will be twice as many hexagons in $|z|^2 = hp$, or $\prod(\nu + 1)$ pairs of conjugate hexagons. The early possible values of $h$ and the corresponding number of pairs are as follows:

$$h = 1, 3, 4, 7, 9, 12, 13, 16, 19.$$
$$\text{Pairs} = 1, 1, 1, 2, 1, 1, 2, 1, 2.$$

Conjugate hexagons produce the same set of ordinates. No other hexagon can produce this set. Indeed, every set of ordinates has the form $\{\pm x, \pm y, \pm(x + y)\}$. One can check that the numbers $\pm X, \pm y, \pm(X + y)$ will be the numbers $\pm x, \pm y, \pm(x + y)$ in some order only if $X = x$ or $X = -x - y$.

For a hexagon inscribed in the circle $|z|^2 = hp$, the maximum height of a vertex lies between $(\sqrt{3}/2)\sqrt{hp}$ and $\sqrt{hp}$. Hence the maximum ordinate in our oblique system lies between $\sqrt{hp}$ and $(2/\sqrt{3})\sqrt{hp}$. For the intervals corresponding to $h = 1, 3, 4, 7, 9, 12$, the only overlap is for $h = 7$ and $h = 9$. There are usually two pairs of conjugate hexagons for $h = 7$ and one pair in the other cases. But if $h = p = 7$, then there is one pair of conjugate hexagons and one selfconjugate hexagon. However, this still produces two sets of ordinates.

Corresponding to $h \leq 1, 3, 4, 9$, we see that there is exactly one set of ordinates $< (2/\sqrt{3})\sqrt{p}$, exactly two sets $< 2\sqrt{p}$, exactly three sets $< (4/\sqrt{3})\sqrt{p}$, and exactly six sets $< (2\sqrt{3})\sqrt{p}$. But $(2/\sqrt{3})\sqrt{hp} < p/3$ when $p > 12h$, so the conditions $p > 12, 36, 48, 108$ guarantee that the ordinates under the bound are identical with the reduced sets or sixth roots there. Looking at numerical results for the early cases, we have the following conclusions for $p \equiv 1 \pmod 6$: There is exactly one reduced set of sixth roots $< (2/\sqrt{3})\sqrt{p}$, exactly two sets $< 2\sqrt{p}$ if $p > 7$, exactly three sets $< (4/\sqrt{3})\sqrt{p}$ if $p > 13$, and exactly six sets $< (2\sqrt{3})\sqrt{p}$ if $p > 67$. There are actually seven sets under the last bound when $p = 43, 61, 67$.

For $1 \leq i \leq (p - 1)/6$, we let $M_i(p)$ be the $i$th smallest possible maximum of a reduced set of sixth roots. This will be studied by its relation to $N_i(p)$, defined for all $i \geq 1$ as the $i$th smallest maximum of a set of ordinates.

Sets of ordinates which are $< p/3$ are identical with reduced sets of sixth roots which are $< p/3$; hence $M_i(p) = N_i(p)$ when $M_i(p) < p/3$ or $N_i(p) < p/3$. More generally, sets of ordinates which are $< p/2$ are among the reduced

sets of sixth roots. Hence, any $N_i(p) < p/2$ is $M_j(p)$ for some $j \geq i$, so that $M_i(p) \leq M_j(p) = N_i(p)$. On the other hand, if $N_i(p) > p/2$ and $i \leq (p-1)/6$, then we will have $M_i(p) < N_i(p)$. Hence, $M_i(p) \leq N_i(p)$ whenever $M_i(p)$ is defined.

The prime $p \equiv 1 \pmod 6$ splits in the ring $\mathbb{Z}[\zeta]$ in the form $p = P\overline{P}$, where $P$ is a prime in the ring. The hexagons inscribed in $|z|^2 = hp$ are obtained from the hexagons inscribed in $|z|^2 = h$ by multiplying by $P$ or $\overline{P}$. By a basic hexagon, we shall mean a hexagon with center at 0 and vertices in the ring. If we start with all the basic hexagons, and expand and rotate by the factors $P$ and $\overline{P}$, we obtain all the hexagons inscribed in circles $|z|^2 = hp$. But we do not need both a hexagon and its conjugate, so it is enough to use the factor $P$. This causes an expansion by the factor $\sqrt{p}$ and a rotation. Hecke [2, §9] proved that there are primes $P$ of the ring in every sector. Hence, when we consider all primes $p \equiv 1 \pmod 6$, the possible rotations caused by multiplying by $P$ are everywhere dense.

Let $\alpha_i$ be the smallest positive number so that for some rotation of the complex plane, there will be $i$ of the basic hexagons lying in the strip $y^2 \leq \alpha_i$. Let $\beta_i$ be the smallest positive number so that for every rotation of the complex plane, there will be $i$ of the basic hexagons lying in the strip $y^2 \leq \beta_i$. Then it follows from the above construction that

$$\alpha_i \leq N_i(p)^2/p \leq \beta_i,$$

and that the bounds $\alpha_i$ and $\beta_i$ are the best that will hold for all values of $p$.

As in §2, the bound $\alpha_i$ will be obtained for a rotation of the plane which places more than one vertex of the hexagons used at the maximum height. The bound $\beta_i$ will be obtained in a similar manner, or when there is just one vertex with maximum height, but it lies on the imaginary axis (which is now not the same as the $y$-axis).

In writing a program to compute $\alpha_i$ and $\beta_i$, it is more convenient to consider supporting lines with various orientations, rather than rotating the plane. We need to know what possible supporting lines must be examined. This depends on knowing what basic hexagons we need to consider.

Notice that every basic hexagon whose vertices satisfy

$$\max(|x|, |y|, |x+y|) \leq j$$

is contained in every basic hexagon whose vertices satisfy

$$\max(|x|, |y|, |x+y|) \geq k,$$

provided that $k \geq 4j/3$. Indeed, the vertices of the latter hexagon will be at a distance $\geq (\sqrt{3}/2)k$ from the origin, and hence its sides will be at a distance $\geq (\sqrt{3}/2)^2 k = 3k/4 \geq j$ from the origin.

There are $3j^2 + 3j + 1$ lattice points satisfying the first condition, and hence $j(j+1)/2$ basic hexagons having these points as vertices. Thus, the values of $\alpha_i$ and $\beta_i$ will be found for $i \leq j(j+1)/2$ without using any hexagons from the second set. Thus we need only look at the $k(k-1)/2$ basic hexagons with

$$\max(|x|, |y|, |x+y|) \leq k-1,$$

where $k \geq 4j/3$. Table 2 gives the values of $\alpha_i$ and $\beta_i$ for $i \leq 80$, which were computed in this way.

TABLE 2. Sixth roots

| $i$ | $\alpha_i$ | $\beta_i$ | $i$ | $\alpha_i$ | $\beta_i$ |
|---|---|---|---|---|---|
| 1 | 1 | 1 1/3 | 41 | 80 1/21 | 85 1/3 |
| 2 | 3 | 4 | 42 | 81 | 85 1/3 |
| 3 | 4 | 5 1/3 | 43 | 81 | 87 7/31 |
| 4 | 7 | 9 | 44 | 81 | 89 17/43 |
| 5 | 8 1/3 | 9 1/3 | 45 | 81 | 90 18/19 |
| 6 | 9 | 12 | 46 | 88 12/13 | 100 |
| 7 | 12 | 16 | 47 | 92 8/73 | 100 |
| 8 | 14 2/7 | 16 1/3 | 48 | 93 28/57 | 100 |
| 9 | 16 | 17 1/3 | 49 | 95 11/43 | 100 |
| 10 | 16 | 21 1/3 | 50 | 96 1/3 | 100 16/21 |
| 11 | 20 4/7 | 25 | 51 | 96 1/3 | 104 1/7 |
| 12 | 21 1/3 | 25 | 52 | 100 | 108 |
| 13 | 24 1/7 | 27 | 53 | 100 | 108 |
| 14 | 25 | 28 | 54 | 100 | 109 28/57 |
| 15 | 25 | 30 10/13 | 55 | 100 | 112 |
| 16 | 30 6/19 | 36 | 56 | 108 | 121 |
| 17 | 32 1/7 | 36 | 57 | 108 | 121 |
| 18 | 33 1/3 | 36 4/7 | 58 | 112 | 121 |
| 19 | 36 | 40 1/3 | 59 | 115 2/19 | 121 |
| 20 | 36 | 41 2/7 | 60 | 117 | 121 5/19 |
| 21 | 36 | 42 6/7 | 61 | 120 1/31 | 123 6/7 |
| 22 | 40 1/3 | 49 | 62 | 120 1/7 | 124 36/37 |
| 23 | 44 5/31 | 49 | 63 | 120 1/3 | 128 4/7 |
| 24 | 45 16/21 | 49 | 64 | 121 | 133 1/3 |
| 25 | 48 | 51 4/7 | 65 | 121 | 133 1/3 |
| 26 | 48 | 52 12/37 | 66 | 121 | 133 1/3 |
| 27 | 49 | 56 1/3 | 67 | 131 11/19 | 144 |
| 28 | 49 | 57 1/7 | 68 | 133 1/3 | 144 |
| 29 | 56 1/13 | 64 | 69 | 133 1/3 | 144 |
| 30 | 56 1/3 | 64 | 70 | 133 1/3 | 144 |
| 31 | 59 20/31 | 65 1/3 | 71 | 137 2/7 | 147 |
| 32 | 61 5/7 | 65 1/3 | 72 | 141 21/43 | 147 |
| 33 | 63 | 66 9/13 | 73 | 142 3/13 | 147 16/19 |
| 34 | 64 | 69 3/13 | 74 | 144 | 149 1/3 |
| 35 | 64 | 71 8/31 | 75 | 144 | 151 14/19 |
| 36 | 64 | 75 | 76 | 144 | 155 4/7 |
| 37 | 72 1/39 | 81 | 77 | 144 | 155 10/13 |
| 38 | 73 12/13 | 81 | 78 | 144 | 161 1/3 |
| 39 | 75 | 81 | 79 | 154 5/7 | 169 |
| 40 | 75 | 82 2/7 | 80 | 155 4/7 | 169 |

There are two special cases which will be discussed. In the unrotated plane, there are $j(j+1)/2$ basic hexagons with ordinates not exceeding $j$. It follows that

$$\alpha_i \le j^2 \le \beta_i \quad \text{for } j(j-1)/2 < i \le j(j+1)/2.$$

Table 2 shows many cases of equality on the right near the beginning of the interval, and on the left near the end. It can be shown that equality holds for at least one value of $i$ at each end, except at the beginning of the first interval:

$$\beta_i = j^2 \quad \text{for } i = j(j-1)/2 + 1, \ j \ge 2,$$
$$\alpha_i = j^2 \quad \text{for } i = j(j+1)/2.$$

Only the second (and easier) of these two formulas will be proved here.

We first show that, for any real $s > 0$, a closed regular hexagon of side $s$ can contain at most $3s^2 + 3s + 1$ lattice points. We again apply the theorem of Pick [5], as in §2. It states that if the vertices of a polygon are at lattice points, the area is $A$, there are $B$ lattice points on the boundary and $C$ inside, then $B + C = A + B/2 + 1$. But now we must compute the area $A$ as a multiple of the area of a fundamental parallelogram, which is $\sqrt{3}/2$. In this sense, the area of a regular hexagon of side $s$ is $3s^2$. Apply Pick's theorem to the convex hull of the set of lattice points in a closed hexagon of side $s$. Here, $A \le 3s^2$ and $B \le 6s$, since the perimeter is at most $6s$ and the lattice points have minimum distance 1. Thus, $B + C \le 3s^2 + 3s + 1$. It is easily seen that equality holds if and only if $s$ is an integer, the vertices are at lattice points, and two of the sides are horizontal.

In any rotated plane, the hexagon $|x| \le j$, $|y| \le j$, $|x + y| \le j$ contains at most $3j^2 + 3j + 1$ lattice points, or at most $j(j+1)/2$ basic hexagons, and equality holds only for a rotation through a multiple of $60°$. No smaller hexagon can contain so many lattice points; hence $\alpha_i = j^2$ when $i = j(j+1)/2$.

The next simplest case is rotation through $30°$ or $90°$. Equivalently, we can consider vertical supporting lines. Notice that the vertical lines containing lattice points are spaced at a distance $1/2$ apart. The number of lattice points on successive verticals to the right of the imaginary axis having arguments $\theta$ with $-30° < \theta \le 30°$ is $0, 1, 1, 1, 2, 2, 2, 3, 3, 3, \ldots$. The sum of the first $j$ terms of this sequence is equal to $1/3+2/3+\cdots+j/3$, unless $j \equiv 1 \pmod 3$, when $1/3$ must be subtracted. In any case, its value is $[j(j+1)/6]$. This gives the number of basic hexagons within a distance $j/2$ of the imaginary axis. When the plane is rotated through $90°$, this corresponds to $|y| \le j/\sqrt{3}$. It follows that

$$\alpha_i \le j^2/3 \le \beta_i \quad \text{for } [j(j-1)/6] < i \le [j(j+1)/6].$$

Again, there is often equality on the right for several values of $i$ at the beginning of the interval, and on the left at the end. However, it is not always true here that there is at least one such equality. More than half of the entries in Table 2 are supplied by rotations of $0°$ and $30°$.

Can the bounds $\alpha_i$ and $\beta_i$ be attained for some prime $p$? To compute $\alpha_i$, we rotate the plane so as to minimize the maximum ordinate of the vertices of a set of $i$ basic hexagons. This rotation will produce more than one vertex with the maximum ordinate. If $N_i(p)^2/p = \alpha_i$, then the rotation used is $P/|P|$, which takes $\overline{P}$ to $\sqrt{p}$. Since there is no lattice point of $\mathbb{Z}[\zeta]$ between 0 and

$\overline{P}$, in the rotated lattice, there will be no lattice point between 0 and $\sqrt{p}$. Thus, the distance between two lattice points at the maximum height is at least $\sqrt{p}$. The convex hull of the $i$ hexagons has six-fold symmetry, and has six sides of length at least $\sqrt{p}$. The closest that they can come to the origin is $(\sqrt{3}/2)\sqrt{p}$, when the convex hull is a regular hexagon of side $\sqrt{p}$. This corresponds to an ordinate $\sqrt{p}$ in the rotated plane. Thus, $\alpha_i \geq p$, and so $N_i(p) \geq p$.

If $N_i(p)^2/p = \beta_i$, then the rotation $P/|P|$ either takes more than one vertex of the hexagons used to a maximum height, or else places a vertex with maximum height on the imaginary axis. The first case again leads to $N_i(p) \geq p$. In the second case, the height is at least $\sqrt{3p}$, since $(\sqrt{3p})i$ is a point of the rotated lattice, and there is no lattice point between it and 0. The height $\sqrt{3p}$ corresponds to an ordinate $2\sqrt{p}$. Hence, $\beta_i \geq 4p$ and $N_i(p) \geq 2p$.

In any case, we conclude that

$$\alpha_i < N_i(p)^2/p < \beta_i$$

at least when $N_i(p) < p$. It is not hard to show that $N_i(p) < p$ just when $i \leq (p-1)/2$, but we do not need to know this to draw conclusions about $M_i(p)$, since we always have $N_i(p) < p$ when $M_i(p) = N_i(p)$. If $p > 9\beta_i$, then $N_i(p)^2 \leq p\beta_i < p^2/9$, so that $N_i(p) < p/3$ and hence $M_i(p) = N_i(p)$. This shows that

$$\alpha_i < M_i(p)^2/p < \beta_i$$

when $p > 9\beta_i$. But we do not need such a strong hypothesis to obtain this inequality. Since $M_i(p) \leq N_i(p)$, the upper bound will hold whenever $M_i(p)$ is defined, that is, for $p \geq 6i+1$. The lower bound will hold at least for $p > 9\alpha_i$. For if it failed, we would have $M_i(p)^2 \leq p\alpha_i < p^2/9$, so that $M_i(p) < p/3$ and hence $M_i(p) = N_i(p)$, which gives a contradiction. Notice that we did not conclude from $p > 9\alpha_i$ that $M_i(p) = N_i(p)$.

## 4. Connection with Jacobsthal sums

If $p$ is an odd prime and $a \not\equiv 0 \pmod{p}$, then the Jacobsthal sum $\phi_n(a)$ and a related sum $\psi_n(a)$ are defined by

$$\phi_n(a) = \sum_{x=0}^{p-1} \left(\frac{x}{p}\right)\left(\frac{x^n+a}{p}\right), \qquad \psi_n(a) = \sum_{x=0}^{p-1} \left(\frac{x^n+a}{p}\right).$$

Some basic results about these sums are given by Berndt and Evans [1, pp. 354–355].

If $p \equiv 1 \pmod{n}$, then we see that

$$\begin{aligned}
\psi_n(a) &\equiv \sum_{x=0}^{p-1} (x^n+a)^{(p-1)/2} \\
&\equiv \sum_{k=0}^{(p-1)/2} \binom{(p-1)/2}{k} a^{(p-1)/2-k} \sum_{x=0}^{p-1} x^{nk} \\
&\equiv -\sum_{j=1}^{[n/2]} \binom{(p-1)/2}{j(p-1)/n} a^{(p-1)/2-j(p-1)/n} \pmod{p},
\end{aligned}$$

since

$$\sum_{x=0}^{p-1} x^s \equiv \begin{cases} -1 \pmod{p} & \text{if } s > 0 \text{ and } p-1 \mid s, \\ 0 \pmod{p} & \text{otherwise.} \end{cases}$$

The formula $\phi_n(a) = \psi_{2n}(a) - \psi_n(a)$ may be used to derive a formula for $\phi_n(a)$. If $p \equiv 1 \pmod{2n}$, then we have

$$\psi_{2n}(a) \equiv -\sum_{j=1}^{n} \binom{(p-1)/2}{j(p-1)/2n} a^{(p-1)/2-j(p-1)/2n} \pmod{p}.$$

But $\psi_n(a)$ is given by the sum of the even-numbered terms, so $\phi_n(a)$ is given by the sum of the odd-numbered terms. Similar formulas for $\phi_n(a)$ were given by Whiteman [7] and Lehmer [3].

In particular, if $p \equiv 1 \pmod{4}$, then

$$\phi_2(a) \equiv -\binom{(p-1)/2}{(p-1)/4} a^{(p-1)/4} \pmod{p}.$$

Since $|\phi_2(a)| < 2\sqrt{p}$, this shows that

$$\binom{(p-1)/2}{(p-1)/4}^4$$

is a number having four fourth roots less than $2\sqrt{p}$ in absolute value. But [1, Theorem 4.4] shows that $\phi_2(a)$ is even. So this set of fourth roots is just double the smallest set.

Similarly, if $p \equiv 1 \pmod{6}$, then

$$\psi_3(a) \equiv -\binom{(p-1)/2}{(p-1)/6} a^{(p-1)/6} \pmod{p}.$$

Since $|\psi_3(a)| < 2\sqrt{p}$, this shows that

$$\binom{(p-1)/2}{(p-1)/6}^6$$

is a number having six sixth roots less than $2\sqrt{p}$ in absolute value. From [1, Eq. (4.1)], we see that the maximum of $|\psi_3(a)|$ is greater than $\sqrt{3p}$. If $p > 16$, then $2\sqrt{p} < p/2$, so that $\psi_3(a)$ furnishes a reduced set of sixth roots, and it is seen to be the next to the smallest. When $p = 13$, the values of $\psi_3(a)$ are $\pm 2, \pm 5, \pm 7$. This yields the reduced set $\pm 2, \pm 5, \pm 6$, which is again the next to the smallest. When $p = 7$, there is only one reduced set of sixth roots.

I became aware of the above formula for $\psi_3(a)$ as a result of some correspondence with Ronald J. Evans in January 1984. This showed that, for each prime $p \equiv 1 \pmod{6}$, there is a set of sixth roots mod $p$, all of which are less than $2\sqrt{p}$ in absolute value. I then wrote a computer program to find all such sets of sixth roots for the first 300 such primes. I found that there were exactly two such sets for each $p > 7$, and that the one furnished by $\psi_3(a)$ was the next to the smallest. This got me interested in the problem, and led to the investigations reported on in this paper.

## 5. Roots of higher order

For other values of $m$, we shall consider only the problem of finding upper bounds for $M_1(p)$, although it is possible to find upper bounds for other $M_l(p)$

as well. I do not know any good lower bounds. We shall prove that, for each $m$, there exists a constant $C$ so that

$$M_1(p) \leq Cp^{1-1/\phi(m)}$$

for all primes $p \equiv 1 \pmod{m}$. I believe that the exponent here is the smallest possible, but I do not have a proof.

Let $K_m$ be the smallest possible value for $C$. We shall prove the following estimates for $K_m$.

(A) $K_m \leq 2^{\mu}$, where $\mu$ is the number of distinct odd primes dividing $m$.

(B) $K_m \leq 3$ if $m$ is divisible by only one prime greater than 3.

(C) $K_m \leq 2/\sqrt{3}$ if $m$ is divisible by no prime greater than 3.

As a special case of (A), we see that $K_m \leq 1$ if $m$ is a power of 2. For $m = 8$, this tells us that $M_1(p) < p^{3/4}$. We showed in §2 that $K_4 = 1$, and it is trivial that $K_1 = K_2 = 1$. It seems likely that $K_8 = 1$ also, but I do not know how to prove that. For $p = 4481$, we have $M_1(p) = 536 > 0.978p^{3/4}$, and hence $K_8 > 0.978$.

The situation for eighth roots is not so simple as for fourth roots. It appears that there are usually two or more sets of eighth roots which are less than $p^{3/4}$. Here is a summary for the first 150 primes $p \equiv 1 \pmod{8}$, giving the frequency of various numbers of sets of eighth roots which are less than $p^{3/4}$.

| Number of sets | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| First 50 primes | 7 | 38 | 4 | 0 | 1 |
| Second 50 primes | 8 | 35 | 6 | 1 | 0 |
| Third 50 primes | 11 | 34 | 5 | 0 | 0 |

Let $\Phi_m(x)$ be the cyclotomic polynomial of order $m$. If $\zeta = e^{2\pi i/m}$, then

$$\Phi_m(x) = \prod_r (x - \zeta^r),$$

where $r$ runs through a reduced residue system mod $m$. If $p \equiv 1 \pmod{m}$ and $t$ is a primitive $m$th root of 1 mod $p$, then we also have

$$\Phi_m(x) \equiv \prod_r (x - t^r) \pmod{p}.$$

Since $\Phi_m(\zeta) = 0$, it follows that

$$\prod_r (\zeta - t^r) \equiv 0 \pmod{p}.$$

In the ring $\mathbb{Z}[\zeta]$, we introduce the ideals $P_r = (p, \zeta - t^r)$ for $1 \leq r \leq m$, $(r, m) = 1$. It follows that $p$ divides the product of the $P_r$. Each two of these ideals are relatively prime, since any common factor of $P_r$ and $P_{r'}$ would divide $t^r - t^{r'}$, which is prime to $p$. Thus a number is divisible by $p$ if and only if it is divisible by each $P_r$. This is the result which we use. Since $p$ has $\phi(m)$ prime ideal factors, it follows that these must be the various $P_r$.

Let $z$ be an element of the ring $\mathbb{Z}[\zeta]$. Then

$$z = \sum_{j=0}^{\phi(m)-1} a_j \zeta^j,$$

where the $a_j$ are integers. We want to find values of $z$ such that

$$\zeta z \equiv tz \pmod{p}.$$

Since $P_1 | \zeta - t$, it will be sufficient to have

$$P_r | z \quad \text{for } 2 \leq r \leq m, \ (r, m) = 1.$$

As $\zeta \equiv t^r \pmod{P_r}$, we see that

$$z \equiv \sum_{j=0}^{\phi(m)-1} a_j t^{jr} \pmod{P_r}.$$

Thus, to have $P_r | z$, it will be sufficient to have

$$\sum_{j=0}^{\phi(m)-1} a_j t^{jr} \equiv 0 \pmod{p}.$$

We want this to be satisfied for $2 \leq r \leq m$, $(r, m) = 1$, and we would like the $|a_j|$ to be as small as possible.

We first consider polynomials

$$F(x) = \sum_{j=0}^{\phi(m)-1} c_j x^j,$$

where the $c_j$ are integers with $0 \leq c_j \leq s$. This allows $(s+1)^{\phi(m)}$ choices for the coefficients. We then consider the values of $F(t^r) \pmod{p}$. There are $p^{\phi(m)-1}$ possible sets of residues. Two of the polynomials will yield the same set of residues if $(s+1)^{\phi(m)} > p^{\phi(m)-1}$. This will be true if we take

$$s = [p^{1-1/\phi(m)}].$$

In this way, we obtain two polynomials $F_1(x)$ and $F_2(x)$ with coefficients in $[0, s]$ such that

$$F_1(t^r) \equiv F_2(t^r) \pmod{p} \quad \text{for } 2 \leq r \leq m, \ (r, m) = 1.$$

Let $G(x) = F_1(x) - F_2(x)$. Then

$$G(x) = \sum_{j=0}^{\phi(m)-1} a_j x^j,$$

where the $a_j$ are integers with $|a_j| \leq s$, and

$$G(t^r) \equiv 0 \pmod{p} \quad \text{for } 2 \leq r \leq m, \ (r, m) = 1.$$

If we put

$$z = G(\zeta) = \sum_{j=0}^{\phi(m)-1} a_j \zeta^j,$$

then $P_r | z$, and we conclude that $\zeta z \equiv tz \pmod{p}$. It will follow that $\zeta^k z \equiv t^k z \pmod{p}$ for all $k$.

We may put

$$\zeta^k z = \sum_{l=0}^{\phi(m)-1} a_{kl}\zeta^l,$$

where the $a_{kl}$ are integers. Here, $a_{0l} = a_l$, and every $a_{kl}$ is a linear combination of the $a_j$ with integer coefficients depending only on $m$. These coefficients can be obtained by repeated use of the cyclotomic equation $\Phi_m(\zeta) = 0$. Since each $|a_j| \leq s$, we can conclude that every $|a_{kl}| \leq Cs$, where $C$ is a constant depending on $m$. Since $\zeta^k z \equiv t^k z \pmod{p}$, we see that $a_{kl} \equiv t^k a_l \pmod{p}$. Thus, the $a_{kl}$ with any fixed $l$ and $0 \leq k < m$ will form a set of $m$th roots of some number mod $p$. Recalling the value of $s$, we see that this proves that

$$M_1(p) \leq Cp^{1-1/\phi(m)}.$$

We wish to estimate $K_m$, which is the smallest possible value of $C$. We start by putting

$$\zeta^k = \sum_{l=0}^{\phi(m)-1} b_{kl}\zeta^l,$$

where the $b_{kl}$ are integers. The value of $b_{kl}$ for a given $l$ depends only on $k \pmod{m}$. For $0 \leq k < m$, $0 \leq l < \phi(m)$, the coefficients $b_{kl}$ form a matrix $B_m$ with $m$ rows and $\phi(m)$ columns. The first $\phi(m)$ rows form the identity matrix $I_{\phi(m)}$ of order $\phi(m)$. If $m$ is even, the lower half of the matrix is the negative of the upper half.

We can now determine all of the $a_{kl}$ in terms of the $a_j$ and the $b_{kl}$. Indeed, we see that

$$\sum_{l=0}^{\phi(m)-1} a_{kl}\zeta^l = \zeta^k z = \sum_{j=0}^{\phi(m)-1} a_j\zeta^{j+k} = \sum_{j=0}^{\phi(m)-1} a_j \sum_{l=0}^{\phi(m)-1} b_{j+k,l}\zeta^l,$$

so that

$$a_{kl} = \sum_{j=0}^{\phi(m)-1} a_j b_{j+k,l}.$$

But $|a_j| \leq s$, so

$$|a_{kl}| \leq s \sum_{j=0}^{\phi(m)-1} |b_{j+k,l}|.$$

For each fixed $l$, the numbers $a_{kl}$ form a set of $m$th roots. This leads to a bound $L_m$:

$$K_m \leq L_m = \min_l \max_k \sum_{j=0}^{\phi(m)-1} |b_{j+k,l}|.$$

Notice that the sum involves adding the absolute values of any $\phi(m)$ consecutive elements of a column of $B_m$, considered as a cycle.

There is a simple relation between the matrices $B_m$ and $B_{qm}$, where $q$ is a prime such that $q|m$. Notice that $\phi(qm) = q\phi(m)$. Let $\eta = e^{2\pi i/qm}$, so that $\zeta = \eta^q$. Then

$$\eta^{kq+r} = \sum_{l=0}^{\phi(m)-1} b_{kl}\eta^{lq+r}$$

for $0 \leq r < q$. This shows that the matrix $B_{qm}$ is obtained from $B_m$ by replacing each element $b_{kl}$ by the matrix $b_{kl}I_q$. It follows that $L_{qm} = L_m$. Thus, $L_m$ depends only on the distinct primes dividing $m$, and not on their multiplicities.

There is also a simple relation between $B_m$ and $B_{2m}$ when $m$ is odd. Here $-\zeta$ is a primitive $(2m)$th root of 1. Since

$$(-\zeta)^k = \sum_{l=0}^{\phi(m)-1} (-1)^{k+l} b_{kl}(-\zeta)^l,$$

it follows that the first $m$ rows of $B_{2m}$ have the elements $(-1)^{k+l}b_{kl}$. The last $m$ rows are the negatives of these. Hence, $L_{2m} = L_m$. Thus, $L_m$ depends only on the distinct odd primes dividing $m$.

Since $L_1 = 1$, it follows that $L_m = 1$ whenever $m$ is a power of 2, and so $K_m \leq 1$ in these cases. We now look at the case $m = q$, where $q$ is an odd prime. Since $\zeta^{q-1} = -1 - \zeta - \zeta^2 - \cdots - \zeta^{q-2}$, the matrix $B_q$ consists of the identity matrix $I_{q-1}$ followed by a row in which all of the elements are $-1$. We can include the elements 1 and $-1$ among $q-1$ consecutive elements of each column, considered as a cycle. Thus, $L_q = 2$, and hence $L_m = 2$ and $K_m \leq 2$ whenever $m$ is divisible by just one odd prime. So we have proved (A) for the cases $\mu = 0$ and $\mu = 1$. A general proof will be postponed to §6.

*Proof of* (B). To show that $K_m \leq 3$ if $m$ is divisible by only one prime greater than 3, it will be sufficient to prove that $L_{3q} = 3$ when $q > 3$ is prime. We need to consider

$$\Phi_{3q}(x) = \frac{(x^{3q}-1)(x-1)}{(x^q-1)(x^3-1)} = \frac{(x^{2q}+x^q+1)(x-1)}{x^3-1}$$
$$= \frac{x^{2q+1} - x^{2q} + x^{q+1} - x^q + x - 1}{x^3 - 1}.$$

In computing the quotient by long division, only the first two terms in the numerator are used until we have passed $x^{q-1}$ in the quotient. Hence, the coefficients $(1, -1, 0)$ repeat in the quotient until this point is reached. But the sequence of coefficients of the cyclotomic polynomial is palindromic, and $x^{q-1}$ is the central term. Hence the remaining coefficients may be obtained by reflection in the center. Thus, the complete sequence of coefficients is as follows:

$$(1, -1, 0)^{(q-1)/3}, 1, (0, -1, 1)^{(q-1)/3} \quad \text{when } q \equiv 1 \pmod 6,$$
$$(1, -1, 0)^{(q-2)/3}, 1, -1, 1, (0, -1, 1)^{(q-2)/3} \quad \text{when } q \equiv 5 \pmod 6.$$

The first $2q-2$ rows of the matrix $B_{3q}$ form the matrix $I_{2q-2}$. Let the matrix formed by the last $q+2$ rows be denoted by $\overline{B}_{3q}$. We look first at the

case $q = 5$. Here,

$$\overline{B}_{15} = \begin{bmatrix} -1 & 1 & 0 & -1 & 1 & -1 & 0 & 1 \\ -1 & 0 & 1 & -1 & 0 & 0 & -1 & 1 \\ -1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & -1 \\ 1 & -1 & 0 & 0 & -1 & 1 & 0 & -1 \\ 1 & 0 & -1 & 1 & -1 & 0 & 1 & -1 \end{bmatrix}.$$

The first and last rows may be read directly from the cyclotomic equation, and the remaining rows may be computed from $\zeta^9 = -\zeta^{-1} - \zeta^4$, $\zeta^{10} = -1 - \zeta^5$, $\zeta^{11} = -\zeta - \zeta^6$, $\zeta^{12} = -\zeta^2 - \zeta^7$, and $\zeta^{13} = -\zeta^3 - \zeta^8$. Notice that the number of elements 1 or $-1$ in the various columns is $(5, 3, 3, 3, 3, 3, 3, 5)$. The maximum sum of the absolute values of eight consecutive elements of a column of $B_{15}$ (considered as a cycle) will be larger if we can include all of the elements counted above and a diagonal element from $I_8$. This is possible only in the first and last columns. Hence, the maxima will be $(6, 3, 3, 3, 3, 3, 3, 6)$. Since $L_{15}$ is the minimum of these, we see that $L_{15} = 3$.

Next, look at the case $q = 7$. The first two rows of $\overline{B}_{21}$ are as follows:

$$\begin{bmatrix} -1 & 1 & 0 & -1 & 1 & 0 & -1 & 0 & 1 & -1 & 0 & 1 \\ -1 & 0 & 1 & -1 & 0 & 1 & -1 & -1 & 1 & 0 & -1 & 1 \end{bmatrix}.$$

The next five rows contain two matrices $-I_5$ separated by two columns of zeros. The last two rows are the second and first rows in reverse order. Notice that the number of elements 1 or $-1$ in the various columns is $(5, 3, 3, 5, 3, 3, 3, 3, 5, 3, 3, 5)$. We can include a diagonal element of $I_{12}$ in the first three and last three columns of $B_{21}$. Thus, the maxima are $(6, 4, 4, 5, 3, 3, 3, 3, 5, 4, 4, 6)$, and so $L_{21} = 3$.

In general, the first two rows and last two rows of $\overline{B}_{21}$ or $\overline{B}_{15}$ appear as the central portions of the first two rows and last two rows of $\overline{B}_{3q}$ when $q \equiv 1$ or $q \equiv 5 \pmod 6$. The rows are completed by repeating the first three elements and last three elements of the central portion. The intermediate rows contain two matrices $-I_{q-2}$ separated by two columns of zeros. In the first $q-4$ and last $q-4$ columns of $B_{3q}$, we can include all the elements 1 or $-1$ of $\overline{B}_{3q}$ and the diagonal element of $I_{2q-2}$ in a block of $2q-2$ consecutive elements of the column, considered as a cycle. This is not possible for the central six columns. Thus, the maxima for the various columns will be

$$(6, 4, 4)^{(q-4)/3}, 5, 3, 3, 3, 3, 5, (4, 4, 6)^{(q-4)/3}$$

when $q \equiv 1 \pmod 6$, and

$$(6, 4, 4)^{(q-5)/3}, 6, 3, 3, 3, 3, 3, 3, 6, (4, 4, 6)^{(q-5)/3}$$

when $q \equiv 5 \pmod 6$. In either case, we see that $L_{3q} = 3$.

## 6. Roots of higher order (continued)

All of the bounds for $K_m$ found in §5 were obtained from the inequality $K_m \leq L_m$ by computing $L_m$. We now use other methods to prove (A) and (C). Note that (A) gives us, for the first time, a bound for $K_m$ which is easy to compute, whereas (C) gives an improved bound for $K_m$ in an interesting case.

*Proof of* (A). This states that $K_m \leq 2^\mu$, where $\mu$ is the number of distinct odd primes dividing $m$. As shown in §5, this is equivalent to $K_m \leq L_m$ when $\mu = 0$ or $\mu = 1$. It is weaker than $K_m \leq L_m$ when $\mu = 2$ and one of the odd primes is 3, since $L_m = 3$ in this case. It is unclear whether it is stronger or weaker than $K_m \leq L_m$ in other cases, since $L_m$ is hard to compute.

We start by proving the following result about roots of unity: If $\zeta = e^{2\pi i/m}$, then there is a set $S_m$ of $\phi(m)$ elements of $\mathbb{Z}[\zeta]$ such that every power of $\zeta$ is a sum of at most $2^\mu$ elements of $S_m$ or their negatives, where $\mu$ is the number of distinct odd primes dividing $m$.

This is trivial for $m = 1$. The proof uses induction from $m$ to $qm$, where $q$ is a prime.

*Case* 1, *when* $q|m$. We obtain the set $S_{qm}$ by multiplying the elements of $S_m$ by the numbers $e^{2\pi ik/qm}$ with $0 \leq k < q$. The new sums are obtained from the old ones using a constant multiplier. Hence, the new value of $\mu$ will be the same as the old one.

*Case* 2, *when* $q \nmid m$. If $q = 2$, then we may take $S_{qm} = S_m$, since $-\zeta$ is a primitive $(2m)$th root of 1, so $\mu$ is unchanged. Now suppose that $q > 2$. If we formed a set $S'$ by multiplying the elements of $S_m$ by $e^{2\pi ik/q}$ with $0 \leq k < q$, then every power of $e^{2\pi ik/qm}$ could be obtained as a sum of $2^\mu$ elements of $S'$ or their negatives. But $S'$ has too many elements. Instead, we use the sums

$$\sigma_k = \sum_{j=0}^{k} e^{2\pi ij/q},$$

and form $S_{qm}$ by multiplying the elements of $S_m$ by the numbers $\sigma_k$ with $0 \leq k < q-1$. Since $e^{2\pi ik/q}$ is $\sigma_0$ when $k = 0$, $\sigma_k - \sigma_{k-1}$ when $0 < k < q-1$, and $-\sigma_{q-2}$ when $k = q-1$, we can obtain every power of $e^{2\pi i/qm}$ by adding at most $2^{\mu+1}$ elements of $S_{qm}$ or their negatives. Here the new $\mu$ is one unit more than the old one. This completes the proof of the result about roots of unity.

As in §5, we consider polynomials

$$F(x) = \sum_{j=0}^{\phi(m)-1} c_j x^j,$$

where the $c_j$ are integers, but instead of supposing that $0 \leq c_j \leq s$, we restrict other integers, which determine the $c_j$, to the interval $[0, s]$. This gives us the same number of polynomials as before, so that the same value of $s$ may be used.

With $\zeta = e^{2\pi i/m}$, we may put

$$\zeta^k F(\zeta) = \sum_{l=0}^{\phi(m)-1} c_{kl}\zeta^l,$$

where the $c_{kl}$ are integers. Here, $c_{0l} = c_l$, and in general the $c_{kl}$ are linear combinations of the $c_j$, where all linear combinations which we consider are to have integer coefficients. The different values of $\zeta^k$ are obtained for $0 \leq k < m$. If we use other values of $k$, they may be reduced mod $m$. Thus, the $c_{kl}$

form a matrix with $m$ rows and $\phi(m)$ columns. It is clear that any row of the matrix determines the entire matrix.

We shall show that the first column of the matrix also determines the entire matrix. If we start with the equation

$$F(\zeta) = c_0 + c_1\zeta + c_2\zeta^2 + \cdots + c_{\phi(m)-1}\zeta^{\phi(m)-1},$$

and multiply by $\zeta^{-j}$, where $0 \leq j < \phi(m)$, we see that the constant term of the product has the form

$$c_{-j,0} = c_j + \text{linear combination of } c_0, c_1, \ldots, c_{j-1}.$$

We can therefore solve for the $c_j$ as linear combinations of the numbers $c_{-j,0}$ with $0 \leq j < \phi(m)$, so these elements of the first column determine the entire matrix.

Any linear relation among the powers $\zeta^k$ with integer coefficients will also hold for the corresponding $\zeta^k F(\zeta)$ and hence for the $c_{k0}$. Thus, the result about roots of unity shows us that there is a set $T_m$ of $\phi(m)$ linear combinations of the $c_{k0}$ such that every $c_{k0}$ is a sum of at most $2^\mu$ elements of $T_m$ or their negatives. The elements of $T_m$ are linear combinations of the original coefficients $c_j$, and conversely. If we restrict each element of $T_m$ to integers in the interval $[0, s]$, then we obtain the required number of polynomials.

We then choose $z$ and the $a_{kl}$ as in §5. The $a_{k0}$ will be a set of $m$th roots of some number. But the $a_{k0}$ are the differences of two choices for the $c_{k0}$. Hence, there will be a set $U_m$ of $\phi(m)$ integers in $[-s, s]$, obtained as the differences of the corresponding elements of the two sets $T_m$, such that every $a_{k0}$ is a sum of at most $2^\mu$ elements of $U_m$ or their negatives. It follows that $K_m \leq 2^\mu$.

*Proof of* (C). This states that $K_m \leq 2/\sqrt{3}$ if $m$ is divisible by no prime greater than 3. Previously, we knew only that $K_m \leq 2$.

Let $m = 2^\alpha 3^\beta$. We may suppose that $\alpha > 0$ and $\beta > 0$, since we know a stronger result when $\beta = 0$, and the case $\alpha = 0$ is equivalent to the case $\alpha = 1$. Thus, we may put $m = 6n$, and we see that $\phi(m) = 2n$ and $\Phi_m(x) = x^{2n} - x^n + 1$.

As in §5, we consider polynomials

$$F(x) = \sum_{j=0}^{2n-1} c_j x^j,$$

where the $c_j$ are integers with $0 \leq c_j \leq s$, but now we make the additional assumption that $[s/2] \leq c_j + c_{n+j} \leq s + [s/2]$ for $0 \leq j < n$. Since we have imposed additional conditions, we will need a larger value of $s$ than that used in §5.

For each $j$, the pair $(c_j, c_{n+j})$ is confined to a square with two triangles removed. The square contains $(s+1)^2$ lattice points. If $s = 2u$, the two triangles together contain $u(u+1)$ lattice points, whereas if $s = 2u+1$, they contain $(u+1)^2$ lattice points. In either case, the two triangles contain at most $(s+1)^2/4$ lattice points, so that there are at least $3(s+1)^2/4$ choices for $c_j$ and $c_{n+j}$, or at least $\{3(s+1)^2/4\}^n$ choices for all of the coefficients.

As in §5, we consider the values of $F(t^r) \pmod p$, where $t$ is a primitive $m$th root of 1 mod $p$, and $2 \leq r \leq m$, $(r, m) = 1$. There are $p^{2n-1}$ possible

sets of residues. Two of the polynomials will yield the same set of residues if $\{3(s+1)^2/4\}^n > p^{2n-1}$. This will be true if we take $s = [(2/\sqrt{3})p^{1-1/2n}]$. In this way, we obtain two polynomials $F_1(x)$ and $F_2(x)$ such that

$$F_1(t^r) \equiv F_2(t^r) \pmod{p} \quad \text{for } 2 \le r \le m,\ (r, m) = 1.$$

Let $G(x) = F_1(x) - F_2(x)$. Then

$$G(x) = \sum_{j=0}^{2n-1} a_j x^j,$$

where the $a_j$ are integers, not all 0, satisfying not only $|a_j| \le s$ for all $j$, but also $|a_j + a_{n+j}| \le s$ for $0 \le j < n$.

As in §5, we put $z = G(\zeta)$ and conclude that $\zeta^k z \equiv t^k z \pmod{p}$ for all $k$. Hence, if we write the products $\zeta^k z$ as polynomials in $\zeta$ of degree less than $2n$, the coefficients of a fixed power $\zeta^l$ will form a set of $m$th roots. It will be sufficient to look at the constant terms of the polynomials.

Since the cyclotomic equation shifts exponents by multiples of $n$, we need only notice that $\zeta^{-n} = 1 - \zeta^n$ and $\zeta^{-2n} = -\zeta^n$ in order to check that the constant terms of $\zeta^{-j} z$ for $0 \le j < 3n$ are

$$a_0, a_1, a_2, \ldots, a_{n-1},$$
$$a_0 + a_n, a_1 + a_{n+1}, a_2 + a_{n+2}, \ldots, a_{n-1} + a_{2n-1},$$
$$a_n, a_{n+1}, a_{n+2}, \ldots, a_{2n-1}.$$

The remaining $3n$ constant terms are the negatives of these.

Thus, the numbers $\pm a_j$ for $0 \le j < 2n$ and $\pm(a_j + a_{n+j})$ for $0 \le j < n$ form a set of $m$th roots mod $p$. Since none of these exceed $s$ in absolute value, we see that $M_1(p) \le s$. Recalling the value of $s$, we see that this shows that $K_m \le 2/\sqrt{3}$.

The bound is known to be sharp for $m = 3$ and $m = 6$. Numerical evidence suggests that it is also sharp for $m = 12$. For $p = 757$, we have $M_1(p) = 165 > 1.1433p^{3/4}$; hence $K_{12} > 1.1433 > 0.99(2/\sqrt{3})$.

Improved bounds can also be obtained in some other cases. However, the other bounds which I can prove do not appear to be sharp. The proofs are more complicated than for (C), so they will not be included.

## Bibliography


1. Bruce C. Berndt and Ronald J. Evans, *Sums of Gauss, Jacobi, and Jacobsthal*, J. Number Theory **11** (1979), 349–398.
2. E. Hecke, *Eine neue Art von Zetafunktionen und ihre Beziehungen zur Verteilung der Primzahlen*. 2, Math. Z. **6** (1920), 11–51.
3. Emma Lehmer, *On Euler's criterion*, J. Austral. Math. Soc. **1** (1959), 64–70.
4. D. J. Newman, *Problem* E1954, Amer. Math. Monthly **75** (1968), 545–546.
5. Georg Pick, *Geometrisches zur Zahlentheorie*, Sitzungsber. Deutschen Natur.-med. Vereines Böhmen "Lotos" in Prag (2) **19** (1899), 311–319.
6. Raphael M. Robinson, *Numbers having m small mth roots* mod $p$, Abstracts Amer. Math. Soc. **5** (1984), 297–298.
7. Albert Leon Whiteman, *Cyclotomy and Jacobsthal sums*, Amer. J. Math. **74** (1952), 89–99.


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