Differences Between Powers of a Primitive Root

MARIAN VÂJÂITU and ALEXANDRU ZAHARESCU

Received 14 May 2001

We study the set of differences \( \{g^x - g^y \pmod{p} : 1 \leq x, y \leq N \} \) where \( p \) is a large prime number, \( g \) is a primitive root \( \pmod{p} \), and \( p^{2/3} < N < p \).

2000 Mathematics Subject Classification: 11A07.

1. Introduction. Let \( p \) be a large prime number and \( g \) a primitive root \( \pmod{p} \). The distribution of powers \( g^n \pmod{p}, 1 \leq n \leq N \), for a given integer \( N < p \) has been investigated in [1, 2, 4]. In this paper, we use techniques from [4] to study the set of differences

\[
A := \{g^x - g^y \pmod{p} : 1 \leq x, y \leq N \}.
\]

A natural question, attributed to Andrew Odlyzko, asks for which values of \( N \) can we be sure that any residue \( h \pmod{p} \) belongs to \( A? \) He conjectured that one can take \( N \) to be as small as \( p^{1/2+\varepsilon} \), for any fixed \( \varepsilon > 0 \) and \( p \) large enough in terms of \( \varepsilon \). If true, this would be essentially best possible since \( A \) has at most \( N^2 \) elements. For any residue \( a \pmod{p} \), denote

\[
\nu(N, a) = \# \{1 \leq x, y \leq N : g^x - g^y \equiv a \pmod{p} \}.
\]

If \( a \equiv 0 \pmod{p} \) we have the diagonal solutions \( x = y \), thus \( \nu(N, 0) = N \). For \( a \not\equiv 0 \pmod{p} \) it is proved in [4, Theorem 2] that

\[
\nu(N, a) = \frac{N^2}{p} + O(\sqrt{p} \log^2 p).
\]

It follows that we can take \( N = c_0 p^{3/4} \log p \) in Odlyzko’s problem, for some absolute constant \( c_0 \). The exponent 3/4 is a natural barrier in this problem, as well as in other similar ones. An example of another such problem is the following: given a large prime number \( p \), for which values of \( N \) can we be sure that any residue \( h \pmod{p} \) belongs to the set \( \{x^2y \pmod{p} : 1 \leq x, y \leq N \} \)? Again we expect that \( N \) can be taken to be as small as \( p^{1/2+\varepsilon} \). As with the other problem, it is known that we can take \( N = c_1 p^{3/4} \log p \) for some absolute constant \( c_1 \), and this is proved by using Weil’s bounds for Kloosterman sums [5]. If one assumes the well-known H∗ conjecture of Hooley which gives square root cancellation in short exponential sums of the form \( \sum_{1 \leq x \leq N} e(a\bar{x}/p) \), where \( \bar{x} \) denotes the inverse of \( x \) modulo \( p \), then we show that \( N \) can be taken to be as small as \( p^{2/3+\varepsilon} \) in the above problem. We mention, in passing, that this question is also related to the pair correlation problem for sequences of
fractional parts of the form \((\{n^2\alpha\})_{n\in\mathbb{N}}\), which would be completely solved precisely if one could deal with the case when \(N = p^{2/3+\epsilon}\) (see [3] and the references therein).

Returning to the set \(A\), its structure is also relevant to the pair correlation problem for the set \(\{g^n(\mod p), 1 \leq n \leq N\}\). Here one wants an asymptotic formula for

\[
\# \left\{ 1 \leq x \neq y \leq N : g^x - g^y \equiv h(\mod p), \ h \in \frac{p}{N} J \right\}, \tag{1.4}
\]

for any fixed interval \(J \subset \mathbb{R}\). The pair correlation problem is similar to Odlyzko’s problem, but it is more tractable due to the extra average over \(h\). This problem is solved in [4] for \(N > p^{5/7+\epsilon}\), the result being that the pair correlation is Poissonian as \(p \to \infty\) (here we need \(N/p \to 0\)). It is also proved in [4] that under the assumption of the generalized Riemann hypothesis (for Dirichlet \(L\)-functions) the exponent can be reduced from \(5/7+\epsilon\) to \(2/3+\epsilon\). We mention that by assuming square root type cancellation in certain short character sums with polynomials \(\sum_{1 \leq n \leq N} \chi(P(n))\), the exponent \(3/4\) in Odlyzko’s problem can be reduced to \(2/3+\epsilon\) as well. Taking into account the difficulty of the conjectures which would reduce the exponent to \(2/3+\epsilon\) in all these problems, it might be of interest to have some more modest, but unconditional results, valid in the range \(N > p^{2/3+\epsilon}\).

Our first objective, in this paper, is to provide a good upper bound for the second moment

\[
M_2(N) := \sum_{a(\mod p)} \left| \nu(N,a) - \frac{N^2}{p} \right|^2 . \tag{1.5}
\]

From (1.3), it follows that \(M_2(N) \ll p^2 \log^4 p\). The following theorem gives a sharper upper bound for \(M_2(N)\).

**Theorem 1.1.** For any prime number \(p\), any primitive root \(g \mod p\), and any positive integer \(N < p\),

\[
M_2(N) \ll pN \log p . \tag{1.6}
\]

Since each residue \(h(\mod p)\) which does not belong to \(A\) contributes an \(N^4/p^2\) in \(M_2(N)\), we obtain the following corollary.

**Corollary 1.2.** For any prime number \(p\), any primitive root \(g \mod p\), and any positive integer \(N < p\),

\[
\# \{ h(\mod p) : h \notin A \} \ll \frac{p^3 \log p}{N^3} . \tag{1.7}
\]

Thus, for \(N > p^{2/3+\epsilon}\), it follows that almost all the residues \(a(\mod p)\) belong to \(A\). Although by its nature the inequality (1.6) does not give any indication on where the possible residues \(h \notin A\) might be located, there is a way of obtaining results as in Corollary 1.2, with \(h\) restricted to a smaller set.

**Theorem 1.3.** For any prime number \(p\), any primitive root \(g \mod p\), and any positive integer \(N < p\),

\[
\# \{ 1 \leq h < \sqrt{p} : h \text{ prime, } h(\mod p) \notin A \} \ll \left( \frac{p^3 \log p}{N^3} \right)^{1/2} . \tag{1.8}
\]
**Corollary 1.4.** For any \( \epsilon > 0 \), any prime number \( p \), and any primitive root \( g \mod p \), almost all the prime numbers \( h < \sqrt{p} \) (in the sense that the exceptional set has \( \ll \epsilon p^{1/2-\epsilon} \) elements) can be represented in the form

\[
h \equiv g^x - g^y \pmod{p}
\]

with \( 1 \leq x, y \leq p^{2/3+\epsilon} \).

Note that a weaker form of Corollary 1.4, with the range \( 1 \leq x, y \leq p^{2/3+\epsilon} \) replaced by the larger range \( 1 \leq x, y \leq p^{5/6+\epsilon} \), follows directly by taking \( N = p^{5/6+\epsilon} \) in Corollary 1.2. The point in Corollary 1.4 is that it gives a result where \( h \) is restricted to belong to a small set, at no cost of increasing the range \( 1 \leq x, y \leq p^{2/3+\epsilon} \).

**2. Proof of Theorem 1.1.** Let \( p \) be a prime number, \( g \) a primitive root \( \mod p \), and \( N \) a positive integer smaller than \( p \). We know that \( a \equiv 0 \pmod{p} \) contributes an \((N - N^2/p)^2 < N^2 \) in \( M_z(N) \). For \( a \not\equiv 0 \pmod{p} \) define a function \( h_a \) on \( \mathbb{Z}/(p-1)\mathbb{Z} \times \mathbb{Z}/(p-1)\mathbb{Z} \) by

\[
h_a(x, y) = \begin{cases} 1, & \text{if } g^x - g^y \equiv a \pmod{p}, \\ 0, & \text{else}. \end{cases}
\]

Thus \( v(N, a) = \sum_{1 \leq x, y \leq N} h_a(x, y) \). Expanding \( h_a \) in a Fourier series on \( \mathbb{Z}/(p-1)\mathbb{Z} \times \mathbb{Z}/(p-1)\mathbb{Z} \) we get

\[
v(N, a) = \sum_{r,s \equiv 0 \pmod{p-1}} \hat{h}_a(r, s) \sum_{1 \leq x, y \leq N} e\left( \frac{rx + sy}{p-1} \right),
\]

where the Fourier coefficients are given by

\[
\hat{h}_a(r, s) = \frac{1}{(p-1)^2} \sum_{x, y \equiv 0 \pmod{p-1}} h_a(x, y) e\left( -\frac{rx + sy}{p-1} \right).
\]

The main contribution in (2.2) comes from the terms with \( r \equiv s \equiv 0 \pmod{p-1} \), and this equals \( \hat{h}_a(0, 0)N^2 \). It is easy to see that \( \hat{h}_a(0, 0) = 1/p + O(1/p^2) \). Thus

\[
v(N, a) = \frac{N^2}{p} \left( 1 + O\left( \frac{1}{p} \right) \right) + R(a),
\]

where

\[
R(a) = \sum_{(r, s) \not\equiv (0, 0)} \hat{h}_a(r, s) F_N(r) F_N(s),
\]

\[
F_N(r) = \sum_{1 \leq x \leq N} e\left( \frac{rx}{p-1} \right), \quad F_N(s) = \sum_{1 \leq y \leq N} e\left( \frac{sy}{p-1} \right).
\]
From (2.4) and the definition of $M_2(N)$, it follows that in order to prove Theorem 1.1 it will be enough to show that

$$\sum_{a=1}^{p-1} |R(a)|^2 \ll pN\log p.$$  

(2.7)

From [4, Lemma 7] it follows that

$$\hat{h}_a(r,s) = \chi_s(-1)\tau(\chi^r)\tau(\chi^s)(-1)\tau(\chi^{r+s})\chi_r^m(a),$$

(2.8)

where $\tau(\chi^r)$, $\tau(\chi^s)$, $\tau(\chi^{r+s})$ are Gauss sums associated with the corresponding multiplicative characters $\chi^r$, $\chi^s$, $\chi^{r+s}$ defined mod $p$, and $\chi$ is the unique character mod $p$ which corresponds to our primitive root $g$ by

$$\chi(g^m) = e\left(\frac{m}{p-1}\right),$$

(2.9)

for any integer $m$. From (2.5) and (2.8) we derive

$$R(a) = \sum_{m \text{ (mod } p-1)} b_m\chi^m(a),$$

(2.10)

where

$$b_m = \frac{\tau(\chi^{-m})}{p(p-1)^2} \sum_{(r,s) \neq (0,0) \text{ (mod } p-1) \atop r+s \equiv m \text{ (mod } p-1)} F_N(r)F_N(s)\chi^s(-1)\tau(\chi^r)\tau(\chi^s).$$

(2.11)

Since

$$|\tau(\chi^n)| = \begin{cases} \sqrt{p}, & \text{if } n \not\equiv 0 \text{ (mod } p-1), \\ 1, & \text{if } n \equiv 0 \text{ (mod } p-1) \end{cases}$$

(2.12)

it follows that

$$|b_m| \ll p^{-3/2} \sum_{r+s \equiv m \text{ (mod } p-1)} |F_N(r)F_N(s)|.$$  

(2.13)

Here $F_N(r)$ and $F_N(s)$ are geometric progressions and can be estimated accurately. We allow $r$, $s$, and $m$ to run over the set $\{-1/2, 0, 1/2, \ldots, (p-1)/2\}$. Then

$$|F_N(r)| \ll \min\left\{N, \frac{p}{|r|}\right\},$$

(2.14)

and similarly for $|F_N(s)|$. From (2.13) and (2.14) it follows that

$$|b_m| \ll p^{-3/2} \sum_{r+s \equiv m \text{ (mod } p-1)} \min\left\{N, \frac{p}{|r|}\right\} \min\left\{N, \frac{p}{|s|}\right\}.$$  

(2.15)
DIFFERENCES BETWEEN POWERS OF A PRIMITIVE ROOT

By Cauchy’s inequality we derive

\[ |b_m| \ll p^{-3/2} \left( \sum_{|r| \leq (p-1)/2} \min \left\{ N^2, \frac{p^2}{|r|^2} \right\} \right)^{1/2} \left( \sum_{|s| \leq (p-1)/2} \min \left\{ N^2, \frac{p^2}{|s|^2} \right\} \right)^{1/2} \]

\[ = p^{-3/2} \sum_{|r| \leq (p-1)/2} \min \left\{ N^2, \frac{p^2}{r^2} \right\} \ll p^{-1/2} N. \]

(2.16)

Ignoring the two terms \( r = 0, s = m \) and \( r = m, s = 0 \) which contribute in (2.15) at most \( 2p^{-3/2}N^2 \leq 2p^{-1/2}N \), the rest of the sum in (2.15) is less than or equal to

\[ \sum_{r+s \equiv m \pmod{p-1}} \frac{p^2}{|r||s|} = S_1 + S_2, \]

(2.17)

where we denote by \( S_1 \) the sum of the terms with \( |r| \leq |s| \) and by \( S_2 \) the sum of the terms with \( |r| > |s| \). Note that in \( S_1 \) we have \( |s| \geq |m|/2 \) and so

\[ S_1 \ll \sum_{0 < |r| \leq (p-1)/2} \frac{p^2}{|m||r|} \ll \frac{p^2 \log p}{|m|} \]

(2.18)

and similarly for \( S_2 \). From (2.16), (2.17), and (2.18) we conclude that

\[ |b_m| \ll \frac{1}{\sqrt{p}} \min \left\{ N, \frac{p \log p}{|m|} \right\}. \]

(2.19)

We now return to (2.10) and compute

\[ p^{-1} \sum_{a=1}^{p-1} |R(a)|^2 = \sum_{a=1}^{p-1} \sum_{m_1 \equiv a \pmod{p-1}} \sum_{m_2 \equiv a \pmod{p-1}} b_{m_1} \tilde{b}_{m_2} \chi^{m_1-m_2}(a) \]

\[ = \sum_{m_1, m_2 \equiv a \pmod{p-1}} b_{m_1} \tilde{b}_{m_2} \sum_{a=1}^{p-1} \chi^{m_1-m_2}(a). \]

(2.20)

The orthogonality of characters \( \pmod{p} \) shows that the last inner sum is zero unless \( m_1 = m_2 \) when it equals \( p-1 \), hence

\[ p^{-1} \sum_{a=1}^{p-1} |R(a)|^2 = (p-1) \sum_{m \equiv a \pmod{p-1}} |b_m|^2. \]

(2.21)

Using (2.19) in (2.21) we obtain

\[ \sum_{a=1}^{p-1} |R(a)|^2 \ll \sum_{|m| \leq (p-1)/2} \min \left\{ N^2, \frac{p^2 \log p}{|m|^2} \right\} \ll pN \log p. \]

(2.22)

Thus (2.7) holds and Theorem 1.1 is proved. \( \square \)
3. Proof of Theorem 1.3. Let \( p, g, \) and \( N \) be as in the statement of the theorem. We will combine the second moment estimate from Theorem 1.1 with two new ideas. The first idea is to restrict the range of \( x, y \) to \( 1 \leq x, y \leq N_1 = \lfloor N/2 \rfloor \) in the definition of \( A \) in order to increase the number of residues which do not belong to the set. To be precise, we consider the set

\[
A_1 = \{g^x - g^y (\mod p) : 1 \leq x, y \leq N_1\},
\]

(3.1)

and note that, for any residue \( h (\mod p) \) which does not belong to \( A \) and any integer \( 0 \leq n \leq N_1 \), the residue \( hg^{-n} \) will not belong to \( A_1 \). Indeed, if there were integers \( x, y \in \{1, 2, \ldots, N_1\} \) such that \( g^x - g^y \equiv hg^{-n} (\mod p) \), then \( g^{x+n} - g^{y+n} \equiv h (\mod p) \) which is not the case since \( 1 \leq x+n, y+n \leq N \), and \( h \) does not belong to \( A \). Therefore, if \( \mathcal{H} \) is a set of residues \( (\mod p) \) which do not belong to \( A \), no element of the set \( \mathcal{M} = \{hg^{-n} (\mod p) : h \in \mathcal{H}, 0 \leq n \leq N_1\} \) will belong to \( A_1 \). The second idea is captured in the following lemma.

**Lemma 3.1.** Let \( p \) be a prime number, \( g \) a primitive root \( \mod p \), \( \mathcal{H} \) a set of prime numbers smaller than \( \sqrt{p} \), \( N_1 \) an integer larger than \( |\mathcal{H}| \), and denote \( \mathcal{M} = \{hg^{-n} (\mod p) : h \in \mathcal{H}, 0 \leq n \leq N_1\} \). Then

\[
|\mathcal{M}| \geq \frac{|\mathcal{H}|(|\mathcal{H}| + 1)}{2}.
\]

(3.2)

**Proof.** The set \( \mathcal{M} \) becomes larger if one increases \( N_1 \) thus it is enough to deal with the case \( N_1 = |\mathcal{H}| \). Consider the sets

\[
\mathcal{H}_n = \{hg^{-n} (\mod p) : h \in \mathcal{H}\}.
\]

(3.3)

Each of these sets has exactly \( |\mathcal{H}| \) elements and we have

\[
\mathcal{M} = \bigcup_{0 \leq n \leq N_1} \mathcal{H}_n.
\]

(3.4)

We claim that for any \( 1 \leq n_1 \neq n_2 \leq N_1 \), the intersection \( \mathcal{H}_{n_1} \cap \mathcal{H}_{n_2} \) has at most one element. Indeed, assume that for some distinct \( n_1, n_2 \in \{1, 2, \ldots, N_1\} \), the set \( \mathcal{H}_{n_1} \cap \mathcal{H}_{n_2} \) has at least two elements, call them \( a \) and \( b \). There are then prime numbers \( p_1, p_2, p_3, p_4 \in \mathcal{H} \) such that

\[
a \equiv p_1 g^{-n_1} \equiv p_2 g^{-n_2} (\mod p),
b \equiv p_3 g^{-n_1} \equiv p_4 g^{-n_2} (\mod p).
\]

(3.5)

Note that since \( n_1 \neq n_2 (\mod p-1) \) we have \( g^{-n_1} \neq g^{-n_2} (\mod p) \) hence the numbers \( p_1 \) and \( p_2 \) are distinct. Also, \( p_1 \) and \( p_3 \) are distinct because \( a \) and \( b \) are distinct. We have

\[
ab = p_1 p_4 g^{n_1 - n_2} = p_2 p_3 g^{n_1 - n_2} (\mod p),
\]

(3.6)

thus

\[
p_1 p_4 \equiv p_2 p_3 (\mod p).
\]

(3.7)
Now the point is that $p_1 p_4$ and $p_2 p_3$ are positive integers less than $p$, and so the above congruence implies the equality $p_1 p_4 = p_2 p_3$. Since these four factors are prime numbers, $p_1$ coincides with either $p_2$ or $p_3$, which is not the case. This proves the claim. We now count in $M$ all the elements of $\mathcal{H}_0$, all the elements of $\mathcal{H}_1$ with possibly one exception if this was already counted in $\mathcal{H}_0$, from $\mathcal{H}_2$ we count all the elements with at most two exceptions, and so on. Thus

$$|M| \geq |\mathcal{H}| + (|\mathcal{H}| - 1) + \cdots + 1 = \frac{|\mathcal{H}|(|\mathcal{H}| + 1)}{2},$$

(3.8)

which proves the lemma.

We now apply Lemma 3.1 to the set $\mathcal{H}$ of prime numbers $< \sqrt{p}$ which do not belong to $A$, and with $N_1 = \lfloor N/2 \rfloor$. It follows that the corresponding set $M$ has at least $|\mathcal{H}|^2/2$ elements. As we know, none of them belongs to $A_1$. Thus each such element contributes an $N_1^4/p^2$ in $M_2(N_1)$, and combining this with Theorem 1.1 we find that

$$\frac{|\mathcal{H}|^2}{2} \frac{N_1^4}{p^2} \leq M_2(N_1) \ll p N_1 \log p.$$ 

(3.9)

This implies

$$|\mathcal{H}| \ll \left(\frac{p^3 \log p}{N^3}\right)^{1/2},$$

(3.10)

which completes the proof of Theorem 1.3.

**References**


**Marian Vâjîitu:** Institute of Mathematics of the Romanian Academy, P.O. Box 1-764, 70700 Bucharest, Romania

_E-mail address:_ mvajaitu@stoilow.imar.ro

**Alexandru Zaharescu:** Institute of Mathematics of the Romanian Academy, P.O. Box 1-764, 70700 Bucharest, Romania

_Current address:_ Department of Mathematics, University of Illinois at Urbana-Champaign, Altgeld Hall, 1409 W. Green Street, Urbana, IL 61801, USA

_E-mail address:_ zaharesc@math.uiuc.edu