On the Maximal Difference Between an Element and its Inverse in Residue Rings

KEVIN FORD

Department of Mathematics, 1409 West Green Street
University of Illinois at Urbana-Champaign
Urbana, IL 61801, USA
ford@math.uiuc.edu

MIZAN R. KHAN

Department of Mathematics and Computer Science
Eastern Connecticut State University
Willimantic, CT 06226, USA
khanm@easternct.edu

IGOR E. SHPARLINSKI

Department of Computing
Macquarie University
Sydney, NSW 2109, Australia
igor@ics.mq.edu.au

CHRISTIAN L. YANKOV

Department of Mathematics and Computer Science
Eastern Connecticut State University
Willimantic, CT 06226, USA
yankovc@easternct.edu

Abstract

We investigate the distribution of $n - M(n)$ where

$$M(n) = \max\left\{ |a - b| \ : \ 1 \leq a, b \leq n - 1 \text{ and } ab \equiv 1 \pmod{n} \right\}.$$

Exponential sums provide a natural tool to obtaining upper bounds on this quantity. Here we use results about the distribution of integers with a divisor in a given interval to obtain lower bounds on $n - M(n)$. We also present some heuristic arguments showing that these lower bounds are probably tight, and thus our technique can be a more appropriate tool to study $n - M(n)$ than a more traditional way using exponential sums.

Mathematics Subject Classification 2000: 11A07, 11N25

1 Introduction

Let $\mathbb{Z}_n$ denote the residues modulo an integer $n \geq 2$. Throughout this paper we assume these residues to consist of the elements $\{0, 1, 2, ..., n - 1\}$. Also, $\mathbb{Z}_n^*$ denotes the subset of $\mathbb{Z}_n$ consisting of all of the integers between 1 and $n - 1$ that are relatively prime to n. Some years ago the second author [9] considered the arithmetical function $M(n)$ defined by

$$M(n) = \max\left\{ |a - b| \ : \ a, b \in \mathbb{Z}_n^* \text{ and } ab \equiv 1 \pmod{n} \right\}$$

and proved the following by elementary methods.

Proposition 1.
$$n - M(n) \geq \left\lceil 2\sqrt{n - 1} \right\rceil,$$
with equality if and only if $n \in \mathcal{S}$ where

$$\mathcal{S} = \{m^2 + \ell m + 1 \ : \ m, \ell \in \mathbb{Z}, m > 0 \text{ and } 0 \leq \ell < 2\sqrt{m} + 1\};$$

and hence

$$\liminf_{n \to \infty} \frac{n - M(n)}{\sqrt{n}} = 2. \tag{1}$$

2

A variety of results about the distribution of pairs (a, b) of solutions to the congruence $ab \equiv 1 \pmod{n}$ and more general congruences can be found in [1, 3, 5, 10, 11, 12, 13, 14]. In particular, Theorem 4 of [10] implies that

$$n - M(n) \leq n^{3/4 + o(1)}.$$

This upper bound is probably far from optimal, and we believe that

$$n - M(n) \leq n^{1/2 + o(1)}. \tag{2}$$

See Section 4 for a more precise statement of our conjecture.

Here, using some results on the distribution of integers with a divisor in a given interval (see [4, 7]), we obtain a result in the opposite direction. In particular we see that

$$\limsup_{n \to \infty} \frac{n - M(n)}{\sqrt{n}} = \infty. \tag{3}$$

Thus, if (2) is correct, it is tight and one cannot remove $o(1)$ from the exponent. We also consider the extreme behavior of $M(p)$ for primes p, and, using some results on the distribution of shifted primes with a divisor in a given interval (see [4, 8]), prove analogous bounds to (1) and (3).

As a curiosity we note that $2^m - M(2^m) \leq 2^{m/2+3/2}$ for all positive integers m. Indeed, if m is even then this is immediate by Proposition 1, if m is odd then it follows from $(2^{(m+1)/2} - 1)(2^m - 2^{(m+1)/2} - 1) \equiv 1 \pmod{2^m}$.

We recall that $U(x) \ll V(x)$ denotes the inequality $|U(x)| \leq cV(x)$ for a fixed constant $c > 0$, $U(x) = o(V(x))$ denotes that $\lim_{x \to \infty} U(x)/V(x) = 0$, and $U(x) \asymp V(x)$ denotes the inequality $C_1 V(x) \leq U(x) \leq C_2 V(x)$ for some positive constants C_1, C_2. Also, $\log z$ always denotes the natural logarithm of $z > 0$.

2   Divisors in intervals

For an infinite sequence of positive integers $\mathcal{A} = (a_n)_{n=1}^{\infty}$ with $a_1 < a_2 < \cdots$, define
$$H(x, y, z; \mathcal{A}) = \#\{n \leq x \ : \ \exists d | a_n \text{ with } y < d \leq z\}.$$

When $\mathcal{A} = \mathbb{N}$, the set of natural numbers, the first author has determined in [4] the order of magnitude of $H(x, y, z; \mathbb{N})$ for all x, y, z. Also in [4] are

3

given upper bounds for $H(x, y, z; P_b)$ of the expected order of magnitude, where $P_b = \{p + b \; : \; p \text{ prime}\}$ is a set of so-called shifted primes. For the problem of bounding $M(n)$ and $M(p)$, we need analogous results where n and p are restricted to an arithmetic progression. Specifically, define

$$\mathcal{T}_k = \{nk - 1 \; : \; n \in \mathbb{N}\}, \qquad \mathcal{U}_k = \{pk - 1 \; : \; p \text{ prime}\}.$$

As usual, we use $\varphi(k)$ to denote the Euler function of a positive integer k.

Proposition 2. *Uniformly for* $100 \le y \le x^{0.51}$, $1.1y \le z \le y^{1.1}$, $1 \le k \le \log x$, *we have*

$$H(x, y, z; \mathcal{T}_k) \; \ll \; x\frac{k}{\varphi(k)} u^\delta (\log(1/u))^{-3/2}, \tag{4}$$

$$H(x, y, z; \mathcal{U}_k) \; \ll \; x\frac{k}{\varphi(k)} u^\delta (\log(1/u))^{-3/2}, \tag{5}$$

where $z = y^{1+u}$ *and*

$$\delta = 1 - \frac{1 + \log\log 2}{\log 2} = 0.086071\ldots.$$

Proof. These estimates follow from the proofs in [4] with only slight modifications to the proofs of Lemmas 6.1 and 14.1 there. To successfully estimate $H(x, y, z; \mathcal{A})$ from above in these lemmas, one needs upper bounds of the expected order for the quantity

$$W = \#\{x_1 < n \le x_2 \; : \; \ell | a_n, \ a_n/\ell \text{ is not divisible by any prime } q \le m\}$$

uniformly for

$$\frac{x}{\log^3 x} \le x_2 \le x, \quad \frac{x_1}{\log^{10} x_1} \le x_2 - x_1 \le x_1, \quad \ell m \le x, \quad m > \log^{20} x.$$

In the case $\mathcal{A} = \mathcal{T}_k$, a standard application of the "small" sieve (see, for example, [6]) gives

$$W \ll \frac{x_2 - x_1}{\ell \log m} \cdot \frac{k}{\varphi(k)}.$$

This is a factor $k/\varphi(k)$ larger than the corresponding bound in the case $\mathcal{A} = \mathbb{N}$. Similarly, when $\mathcal{A} = \mathcal{U}_k$, applying the "large" sieve as in Lemma 14.1 of [4] yields an upper bound for W which is a factor $k/\varphi(k)$ larger than the bound given there. The results of §4 of [4] then finish the proof. □

4

Remarks. In both cases $\mathcal{A} = \mathcal{T}_k$ and $\mathcal{A} = \mathcal{U}_k$, we have ignored the fact that $W = 0$ if $(\ell, k) > 1$. It is possible to use this fact to remove the factors $k/\varphi(k)$ from (4) and (5), but this requires a more complicated modification of the proofs in [4]. Since we are averaging over k below, we do not gain anything by this improvement.

3 Main Results

3.1 Lower Bounds

Theorem 3. *Let $f(x)$ be any positive function tending monotonically to zero as $x \to \infty$. The inequality*

$$n - M(n) \geq n^{1/2} (\log n)^{\delta/2} (\log \log n)^{3/4} f(n)$$

holds:

- *for all positive integers $n \leq x$, except for possibly $o(x)$ of them,*

- *for all prime $n = p \leq x$ except for possibly $o(x/\log x)$ of them.*

Proof. Let x be large and set

$$y = (\log x)^{\delta/2} (\log \log x)^{3/4} f(x/2).$$

It suffices to show $n - M(n) \leq yn^{1/2}$ for $o(x)$ of the integers n between $x/2$ and x. Without loss of generality, suppose $f(x) \geq 1/\log \log x$ for all $x > 10$. We define $\mathcal{J}_k$ to be the set of positive integers $n \in (x/2, x]$ such that $n - M(n) \leq yn^{1/2}$, and for which there are $a, b \in \mathbb{Z}_n^*$, $ab \equiv 1 \pmod{n}$, $M(n) = b - a$ and $a(n - b) = nk - 1$.

By the arithmetic-geometric mean inequality, for every $n \in \mathcal{J}_k$,

$$\frac{n - M(n)}{2} = \frac{n - b + a}{2} \geq \sqrt{a(n - b)} = \sqrt{kn - 1}. \tag{6}$$

Thus $\mathcal{J}_k = \varnothing$ for $k \geq y^2 + 1$. Suppose $1 \leq k < y^2 + 1$, $n \in \mathcal{J}_k$, $ab \equiv 1 \pmod{n}$ and $a(n - b) = kn - 1$. Then

$$\sqrt{kx/2 - 1} \leq \max(a, n - b) \leq y\sqrt{x}.$$

By inequality (4) of Proposition 2,

$$\#\mathcal{J}_k \leq H(x, \sqrt{kx/2 - 1}, y\sqrt{x}; \mathcal{T}_k) \ll \frac{x(\log(3y^2/k))^\delta}{(\log x)^\delta(\log\log x)^{3/2}}\frac{k}{\varphi(k)}.$$

By the elementary estimate $\sum_{n \leq u}(n/\phi(n))^2 = O(u)$ and the Cauchy-Schwarz inequality,

$$\sum_{1 \leq k < y^2 + 1}\#\mathcal{J}_k \ll \frac{xy^2}{(\log x)^\delta(\log\log x)^{3/2}} = o(x).$$

This proves the first part of the theorem.

The second part, concerning $p - M(p)$, is proved by the same argument, using inequality (5) of Proposition 2. □

3.2   Upper Bound for Primes

We now prove an analogue of (1) for the set of primes.

Theorem 4. *For infinitely many primes p, we have*

$$p - M(p) \leq 2\sqrt{p} + \frac{\sqrt{p}}{\log p}.$$

Proof. Let $\varepsilon = 1/(4\log x)$. We show that for sufficiently large x, there is a prime in the interval $((1-\varepsilon)x, x]$ such that $p - 1$ has a divisor d in the interval $((1-2\varepsilon)\sqrt{x}, (1-\varepsilon)\sqrt{x}]$. If we write $p - 1 = df$, then $M(p) \geq p - f - d$. But, if x is so large that $\varepsilon \leq 0.01$, then

$$f + d = \frac{p-1}{d} + d \leq \frac{x}{(1-2\varepsilon)\sqrt{x}} + (1-\varepsilon)\sqrt{x} \leq (2 + 3\varepsilon)\sqrt{p},$$

which implies the desired result.

It now suffices to show that

$$\sum_{(1-2\varepsilon)\sqrt{x} < d \leq (1-\varepsilon)\sqrt{x}}[\vartheta(x; d, 1) - \vartheta((1-\varepsilon)x; d, 1)] > 0, \qquad (7)$$

where

$$\vartheta(x; q, a) = \sum_{\substack{p \leq x \\ p \equiv a \pmod q}}\log p.$$

6

To obtain this result, we apply the main theorem of [2]. A particular case of this theorem gives, for any set $\mathcal{I}$ of positive integers in the interval $[1, \sqrt{z}]$,

$$\sum_{q \in \mathcal{I}} \left| \vartheta(z; q, 1) - \frac{z}{\varphi(q)} \right| \ll \frac{z}{(\log z)^{10}} + \frac{z(\log \log z)^2}{\log^2 z} \sum_{q \in \mathcal{I}} \frac{1}{\varphi(q)}.$$

We apply this with $\mathcal{I} = ((1 - 2\varepsilon)\sqrt{x}, (1 - \varepsilon)\sqrt{x}]$ and with $z = x$ and $z = (1 - \varepsilon)x$. We see that the left side of (7) is

$$\left(\varepsilon x + O\left(\frac{x(\log \log x)^2}{\log^2 x} \right) \right) \sum_{d \in \mathcal{I}} \frac{1}{\varphi(d)} + O\left(\frac{x}{\log^{10} x} \right).$$

Since

$$\sum_{d \in \mathcal{I}} \frac{1}{\varphi(d)} \geq \sum_{d \in \mathcal{I}} \frac{1}{d} \geq \frac{\#\mathcal{I}}{\sqrt{x}} \gg \varepsilon,$$

inequality (7) follows and this completes the proof of the theorem.

$\square$

4 Conjectures and Heuristic Arguments

Conjecture 5. *Let $g(x)$ be any positive function tending monotonically to ∞ as $x \to \infty$. The inequality*

$$n - M(n) \leq n^{1/2}(\log n)^{\delta/2}(\log \log n)^{3/4} g(n)$$

holds:

- *for all positive integers $n \leq x$, except for possibly $o(x)$ of them,*

- *for all prime $n = p \leq x$ except for possibly $o(x/\log x)$ of them.*

Conjecture 6. *For all n, $n - M(n) \leq n^{1/2}(\log n)^{\delta/2 + 1/2 + o(1)}$.*

If true, Conjecture 5, together with Theorem 3, would imply that for most n,

$$n - M(n) \approx n^{1/2}(\log n)^{\delta/2}(\log \log n)^{3/4}.$$

It may be the case that

$$G(n) = \frac{n - M(n)}{n^{1/2}(\log n)^{\delta/2}(\log \log n)^{3/4}}$$

7

has a probability distribution function. That is

$$F(z) = \lim_{x \to \infty} \frac{1}{x} \#\{n \le x \ : \ G(n) \le z\}$$

exists for every real $z > 0$. The same may be true (with a different F) if we restrict to prime n.

We now give a heuristic argument for Conjectures 5 and 6. First, using the methods in [4], one can prove under the hypotheses of Proposition 2 that

$$H(x, y, z; \mathcal{T}_k) - H(x/2, y, z; \mathcal{T}_k) \gg xu^\delta (\log(1/u))^{-3/2}.$$

Let $x/2 < n \le x$ and put

$$y = (\log(x/2))^{\delta/2} (\log\log(x/2))^{3/4} g(x/2).$$

Then, uniformly in $y^2/1000 \le k \le y^2/100$, in the notation of the proof of Theorem 3, we obtain

$$\#\mathcal{J}_k \ge H(x, \sqrt{kx}, \tfrac{1}{4}y\sqrt{x}) - H(x/2, \sqrt{kx}, \tfrac{1}{4}y\sqrt{x}) \gg \frac{x}{(\log x)^\delta (\log\log x)^{3/2}}.$$

Thus, the "probability" that a "random" integer lies in a particular $\mathcal{J}_k$ is at least $\vartheta = c_0 (\log x)^{-\delta} (\log\log x)^{-3/2}$ for some positive constant c_0. Since $\gcd(jn-1, hn-1)$ is very small for $j, h \le k$, these "random events" are essentially independent. Thus, the probability that n does not lie in any set $\mathcal{J}_k$ for $y^2/1000 \le k \le y^2/100$ should be

$$O\left((1-\vartheta)^{y^2/200}\right) = O\left(e^{-\vartheta y^2/200}\right) = o(1).$$

Finally, if

$$y^2 \ge \frac{1000 \log x}{-\log(1-\vartheta)},$$

which occurs if $y \ge (\log x)^{\delta/2+1/2+\varepsilon}$, then $(1-\vartheta)^{y^2/200} \le x^{-5}$. It thus seems highly unlikely that any n fails to lie in some $\mathcal{J}_k$.

There are several natural open questions related to this work. For example, what are the analogues of our results for polynomials over finite fields? One needs to extend the results of [4, 7, 8] to polynomials and shifted irreducible polynomials having a divisor whose degree is in a given interval. This is an interesting question on its own.

One can also study the distribution of $|a-b|$ for more general congruences $f(a, b) \equiv 0 \pmod{n}$ with polynomials $f(X, Y) \in \mathbb{Z}[X, Y]$. However it seems that our approach does not apply and exponential sums provide the only feasible alternative.

References

[1] J. BECK AND M. R. KHAN, On the Uniform Distribution of Inverses Modulo n, *Periodica Mathematica Hungarica* **44** (2002), 147–155.

[2] E. BOMBIERI, J. FRIEDLANDER AND H. IWANIEC, Primes in Arithmetic Progressions to Large Moduli, III, *J. Amer. Math. Soc.* **2**, No.2 (1989), 215–224.

[3] C. COBELI AND A. ZAHARESCU, On the Distribution of the $\mathbb{F}_p$-points on an Affine Curve in r Dimensions, *Acta Arithmetica* **99** (2001), 321–329.

[4] K. FORD, The Distribution of Integers with a Divisor in a Given Interval, *Preprint*, 2004.

[5] A. GRANVILLE, I. E. SHPARLINSKI AND A. ZAHARESCU, On the Distribution of Rational Functions Along a Curve over $\mathbb{F}_p$ and Residue Races, *Preprint*, 2004.

[6] H. HALBERSTAM AND H.-E. RICHERT, *Sieve Methods*, Academic Press, 1974.

[7] R. HALL AND G. TENENBAUM, *Divisors*, Cambridge Tracts in Mathematics **90**, Cambridge University Press, 1988.

[8] K.-H. INDLEKOFER AND N. M. TIMOFEEV, Divisors of Shifted Primes, *Publ. Math. Debrecen* **60** (2002), 307–345.

[9] M. R. KHAN, Problem 10736: An Optimization with a Modular Constraint, *Amer. Math. Monthly* **108** (2001), 374–375.

[10] M. R. KHAN AND I. E. SHPARLINSKI, On the Maximal Difference between an Element and its Inverse modulo n, *Periodica Mathematica Hungarica* **47** (2003), 111–117.

[11] M. VAJAITU AND A. ZAHARESCU, Distribution of Values of Rational Maps on the $\mathbb{F}_p$-points on an Affine Curve, *Monathsh. Math.* **136** (2002), 81–86.

[12] W. ZHANG, On the Difference between an Integer and Its Inverse Modulo n, *J. Number Theory* **52** (1995), 1–6.

[13] W. ZHANG, On the Distribution of Inverses Modulo n, *J. Number Theory* **61** (1996), 301–310.

[14] Z. ZHENG, The Distribution of Zeros of an Irreducible Curve over a Finite Field, *J. Number Theory* **59** (1996), 106–118.