EXPONENTIAL SUMS OVER SUBGROUPS GENERATED BY 2

Rafał Bystrzycki
Dept. of Discrete Mathematics, Adam Mickiewicz University, Poznań, Poland
rafbys@amu.edu.pl

Received: 11/4/16, Revised: 1/29/18, Accepted: 3/16/18, Published: 3/23/18

Abstract

We investigate upper bounds for the absolute value of the sum $s(a/q) = \sum_{r=1}^{\tau} e(\frac{a2^r}{q})$ (where τ is the multiplicative order of 2 modulo q), concentrating primarily on the case of small τ, i.e., of the order of $\log q$. We generalize methods used by Kaczorowski and Molteni and strengthen their results. In particular, we prove that if $\tau \geq \kappa(\lfloor \log_2(q) \rfloor + 4) + 5$ then $\max_{(a,q)=1} |s(a/q)| < \tau - 2(\kappa + 1)$. We further improve the constant 2 to some larger constant (about 2.37) at a cost of increasing τ slightly.

1. Introduction

Bounding exponential sums is a very active area of research. Here we consider the special case of sums over subgroups generated by 2. If the order of a subgroup is large, then there is a general result proved by Bourgain, Glibichuk and Konyagin, which gives a good upper bound.

Theorem 1 ([1]). *Let $F = \mathbb{F}_p$ be a finite field of prime order, and let H be a multiplicative subgroup of F such that $|H| \geq p^\delta$ for some $0 < \delta < 1$. Then if p is sufficiently large depending on δ, for some $\epsilon(\delta) > 0$ we have*

$$\sup_{\xi \in \mathbb{Z}_p \setminus \{0\}} \left| \sum_{x \in H} e(x\xi) \right| \leq p^{-\epsilon} |H|.$$

Throughout the rest of this paper we concentrate on small subgroups. This line of investigation was essentially started by the work of Molteni [3]. We are going to use the following notation. For some fixed odd integer q:

- $\tau := ord_q(2)$
- $\mathcal{L} := \lfloor \log_2(q) \rfloor$

- $e(x) := \exp(2\pi i x)$
- $s(a/q) := \sum_{r=1}^{\tau} e(a2^r/q)$

When subgroups are small much less cancellation is expected. In fact, Kaczorowski and Molteni provided infinitely many examples showing that in general the cancellation may be as small as some explicit constant.

Theorem 2 ([2]). *There exists a positive constant c and a sequence of integers $q \to \infty$ such that*

$$\max_{(a,q)=1} |s(a/q)| \geq \tau - c + O\left(\frac{1}{q}\right).$$

Moreover $c \leq 2\sum_{r=1}^{\infty} \sin^2\left(\frac{\pi}{2^r}\right) = 3.394\ldots$.

They also proved the following upper bound.

Theorem 3 ([2]). *If $\tau \geq \kappa(\mathcal{L}+1)+2$ for a nonnegative integer κ and $q > 3$, then*

$$\max_{(a,q)=1} |s(a/q)| < \tau - \kappa - 1.$$

We improve the above bound. Here is the main result of our paper.

Theorem 4. *If $\tau \geq \kappa(\mathcal{L}+4)+5$ for some positive integer κ, then*

$$\max_{(a,q)=1} |s(a/q)| < \tau - 2(\kappa+1). \tag{1}$$

2. Proof of Theorem 4

The following fact plays a key role in the proof of Theorem 3.

Lemma 1 ([2]). *Suppose $\zeta = e(\theta)$ for some real number θ with $\Re(\zeta) \leq 0$ and $\zeta \neq -1$. Then*

$$|\zeta^2 - 1| < |\zeta - 1| \qquad \text{or} \qquad |\zeta^4 - 1| < |\zeta^2 - 1|.$$

Similarly, our proof relies on the following lemma.

Lemma 2. *Suppose $\zeta = e(\theta)$ for some real number θ with $\Re(\zeta) \leq 0$ and $\zeta \neq -1$. Then*

$$|\zeta + \zeta^2 + \zeta^4 + \zeta^8 + \zeta^{16}| < 3.$$

Proof. Let $f(\theta) = |\zeta + \zeta^2 + \zeta^4 + \zeta^8 + \zeta^{16}|$. Based on the well known Euler identity $e^z = \cos z + i \sin z$, we have

$$\begin{aligned} f(\theta) &:= ((\sin 2\pi\theta + \sin 4\pi\theta + \sin 8\pi\theta + \sin 16\pi\theta + \sin 32\pi\theta)^2 \\ &\quad + (\cos 2\pi\theta + \cos 4\pi\theta + \cos 8\pi\theta + \cos 16\pi\theta + \cos 32\pi\theta)^2)^{\frac{1}{2}}. \end{aligned} \tag{2}$$

In order to prove the lemma, it suffices to show that values of the function f on the interval $[\frac{1}{4}, \frac{3}{4}]$ are less than 3 (except of the point $\theta = \frac{1}{2}$). Repeatedly using the formulae $\sin(2x) = 2\sin(x)\cos(x)$ and $\cos(2x) = 2\cos(x)^2 - 1$, and then using substitution $x = \cos 2\pi\theta$ we get the following polynomial

$$\begin{aligned} w(x) &= 32768x^{15} + 16384x^{14} - 122880x^{13} - 53248x^{12} \\ &\quad + 184320x^{11} + 66560x^{10} - 140800x^9 - 39680x^8 + 57728x^7 \\ &\quad + 11200x^6 - 12320x^5 - 1216x^4 + 1240x^3 + 12x^2 - 48x + 5. \end{aligned} \tag{3}$$

We need to show that it is bounded by 9 on the interval $(-1, 0]$. By standard tools (we used wxMaxima 16) one can verify that the 14 roots of $w'(x)$ are:
$-1.057176\ldots$, $-0.948631\ldots$, $-0.855344\ldots$, $-0.720103\ldots$, $-0.531527\ldots$, $-0.344771\ldots$, $-0.123226\ldots$, $0.148074\ldots$, $0.266689\ldots$, $0.405528\ldots$, $0.631112\ldots$, $0.794703\ldots$, $0.907195\ldots$, $0.960809\ldots$.

Only the points $-0.948631\ldots$, $-0.855344\ldots$, $-0.720103\ldots$, $-0.531527\ldots$, $-0.344771\ldots$, $-0.123226\ldots$ belong to the considered interval. The polynomial $w(x)$ takes the values $0.8492539\ldots$, $5.0979332\ldots$, $0.0739295\ldots$, $7.3947072\ldots$, $2.1874524\ldots$ and $8.8596675\ldots$ at those points; furthermore $w(-1) = 9$ and $w(0) = 5$. Hence $w(x) < 9$ for any $x \in (-1, 0]$. Since $f(\theta) = \sqrt{w(\cos 2\pi\theta)}$ the assertion follows. □

The graphs of $f(\theta)$ and $w(x)$ in the relevant ranges are shown in Figures 1 and 2, respectively.

Lemma 3. *Let $(a, q) = 1$ and $q > 5$. Then for any integer $m \geq 0$ there exists an integer l such that $m \leq l < \mathcal{L} + m$ and*

$$s_5(a2^l/q) := \left| e\left(\frac{2^l a}{q}\right) + e\left(\frac{2^{l+1} a}{q}\right) + e\left(\frac{2^{l+2} a}{q}\right) + e\left(\frac{2^{l+3} a}{q}\right) + e\left(\frac{2^{l+4} a}{q}\right) \right| < 3.$$

Proof. Without loss of generality we may assume that $m = 0$ (otherwise $2^m a$ should be considered instead of a). If $\Re e(2^L a/q) \leq 0$ for some $0 \leq L < \mathcal{L}$, then the claim follows from Lemma 2. Further we assume that $\Re e(2^L a/q) > 0$ for any $0 \leq L < \mathcal{L}$. Denote by θ the real number satisfying $|\theta| < \frac{1}{4}$ and $e(2^{\mathcal{L}-1} a/q) = e(\theta)$. Then the numbers $e(\frac{2^l a}{q})$ for $0 \leq l \leq \mathcal{L} - 1$ are equal to $e(\frac{\theta}{2^k})$ for $\mathcal{L} - 1 \geq k \geq 0$, correspondingly. In particular $e(a/q) = e(\theta/2^{\mathcal{L}-1})$ and so

$$\frac{1}{q} \leq \left| \frac{\theta}{2^{\mathcal{L}-1}} \right| < \frac{1}{2^{\mathcal{L}+1}} < \frac{1}{q},$$

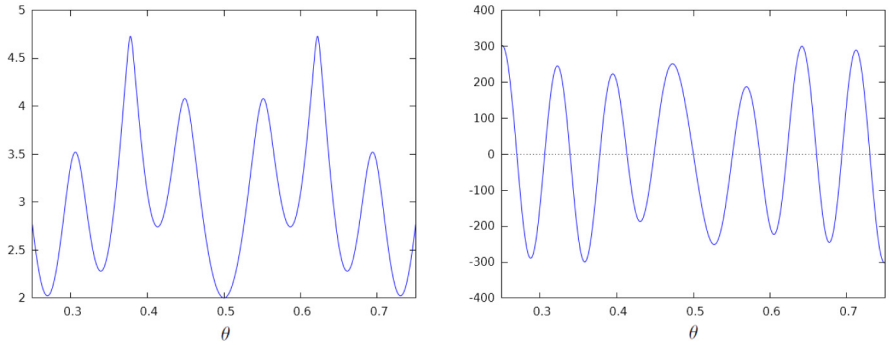


Figure 1: Values of trigonometric polynomial and its derivative.

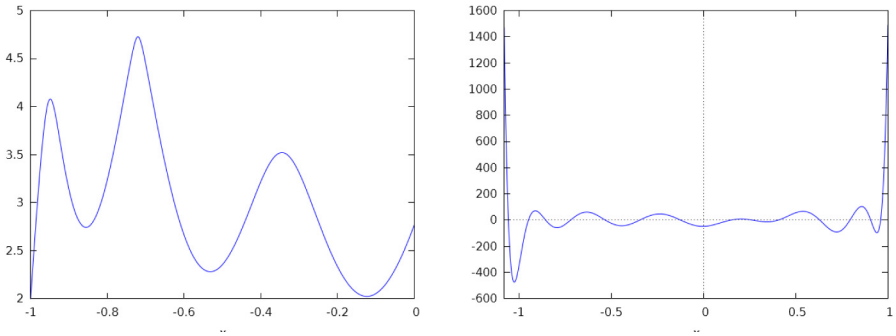


Figure 2: Values of corresponding algebraic polynomial and its derivative.

which leads to a contradiction. □

Now we are ready to prove Theorem 4.

Proof of Theorem 4. By Lemma 3 and the assumption of the theorem there exists a number l_0 such that $s_5(a2_0^l/q) < 3$. By the periodicity of $e(2^l a)$ it follows that $s(a/q) = \sum_{l=l_0}^{l_0+\tau-1} e\left(\frac{2^l a}{q}\right)$. We divide the set of summand indices into intervals: $\{l_0, l_0+1, l_0+2, l_0+3, l_0+4\}$ and at least κ intervals of length $\mathcal{L}+4$. By the previous lemma, each interval contains some number l such that $s_5(2^l a/q) < 3$; furthermore, it can be chosen from the first $\mathcal{L}$ elements of the interval. Hence, using the triangular inequality we get $|s(a/q)| < \tau - 5(\kappa+1) + 3(\kappa+1) = \tau - 2(\kappa+1)$. □

The above proof differs from the proof of Theorem 2 by considering the sum of five consecutive summands instead of only two. Apart from that, the argument is analogous.

3. Further Improvement

If we consider taking more than 5 summands, we can improve the result; however, the argument becomes more technical. The next theorem is an example of such an improvement.

Theorem 5. *If $\tau \geq \kappa(\mathcal{L}+5)+6$ for a nonnegative integer κ, then*

$$\max_{(a,q)=1} |s(a/q)| < \tau - 2.37(\kappa+1). \tag{4}$$

Proof. Let $\zeta = e(\theta)$ for some real number θ such that $-0.999118 \leq \Re(\zeta) \leq 0.021$. First we show that

$$|\zeta + \zeta^2 + \zeta^4 + \zeta^8 + \zeta^{16} + \zeta^{32}| < 3.63. \tag{5}$$

By almost the same arguments as in the proof of Lemma 2, we come to the conclusion that it is enough to bound the polynomial

$$\begin{aligned} w(x) &= 2147483648\,x^{31} + 1073741824\,x^{30} - 16642998272\,x^{29} \\ &- 7784628224\,x^{28} + 58250493952\,x^{27} + 25300041728\,x^{26} \\ &- 121701924864\,x^{25} - 48637149184\,x^{24} + 169030451200\,x^{23} \\ &+ 61446553600\,x^{22} - 164479631360\,x^{21} - 53589573632\,x^{20} \\ &+ 115135741952\,x^{19} + 32967491584\,x^{18} - 58595868672\,x^{17} \\ &- 14351925248\,x^{16} + 21655027712\,x^{15} + 4363173888\,x^{14} \\ &- 5741977600\,x^{13} - 895791104\,x^{12} + 1066528768\,x^{11} + 115973120\,x^{10} \\ &- 133433856\,x^9 - 8054272\,x^8 + 10580864\,x^7 + 131264\,x^6 \\ &- 484512\,x^5 + 15376\,x^4 + 11160\,x^3 - 704\,x^2 - 110\,x + 10 \end{aligned} \tag{6}$$

on $[-0.999118, 0.021]$.

Its extrema are approximately at the points: -1.074387, -0.989143, -0.971382, -0.939692, -0.890416, -0.829615, -0.776161, -0.717199, -0.637236, -0.564463, -0.466427, -0.359011, -0.252928, -0.159027, -0.043114, 0.173648, 0.309891, 0.406477, 0.508774, 0.579395, 0.672828, 0.766044, 0.812919, 0.849519, 0.910000, 0.950689, 0.978700, 0.990701.

For a clearer view, let us first calculate the values of the function $h(x) = 6 - \sqrt{w(x)}$. At the first 16 points we obtain: $-94.6222693\ldots$, $4.5876861\ldots$, $3.6328312\ldots$, $5.9968304\ldots$, $3.0354921\ldots$, $4.8466566\ldots$, $3.9512462\ldots$, $4.7487580\ldots$, $3.4408042\ldots$, $4.3572044\ldots$, $2.6267897\ldots$, $5.4964278\ldots$, $2.9328713\ldots$, $4.0850766\ldots$, $2.4415242\ldots$, 6.0. At the point -0.999118 it takes the value $2.3703688\ldots$. We see that all the values are greater than 2.37 so $w(x) < 3.63$ for $x \in [-0.999118, 0.021]$ unless there exists some another minimum of h in this interval.

To exclude this possibility, we consider the second and the third derivative of $w(x)$. The second derivative has a root $0.0211231\ldots$, while the third derivative has roots $0.0683720\ldots$ and $0.1498680\ldots$. If f has an additional minimum in the interval $[-0.999118, 0.021]$, then w' has two additional roots in this interval. As a derivative always has some zero between two zeros of a function, that would imply that w'' has 16 roots smaller than 0.02, a root $0.0211231\ldots$ and 12 roots greater than $0.1736481\ldots$. That in turn would imply that w''' has 29 roots: 16 roots smaller than $0.0211231\ldots$, points $0.0683720\ldots$ and $0.1498680\ldots$, and 11 roots greater than $0.1736481\ldots$. But this is a polynomial of degree 28, so we come to the contradiction. We conclude that $h(x) > 2.37$ and thereby $w(x) < 3.63$ for $x \in [-0.999118, 0.021]$.

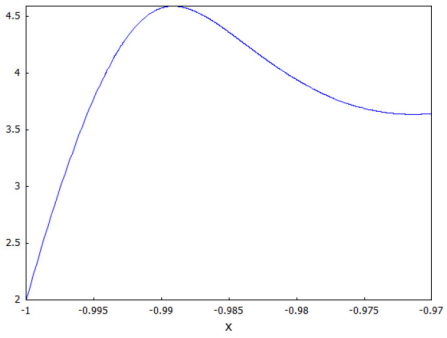
Now we show that there exists an integer l such that $m \le l < \mathcal{L} + m$ and

$$s_6(a2^l/q) := \left| \sum_{i=0}^{5} e\left(\frac{2^{l+i}a}{q}\right) \right| < 3.63.$$

For this purpose we repeat the argument from the proof of Lemma 3. If $\Re e(2^L a/q) > 0$ for any $0 \le L < \mathcal{L}$, then the argument is the same. If $\Re e(2^L a/q) \le 0$ for some $0 \le L < \mathcal{L}$, then the claim follows from (5) by taking $l = L$ or $l = L - 1$, as $\cos(2\arccos(0.021)) = -0.999118$.

The proof of Theorem 5 proceeds in the same way as the proof of Theorem 4. □

The graphs for $h(x)$ and the derivative of $w(x)$ in the ranges critical to the twist in the argument are shown in Figure 3.



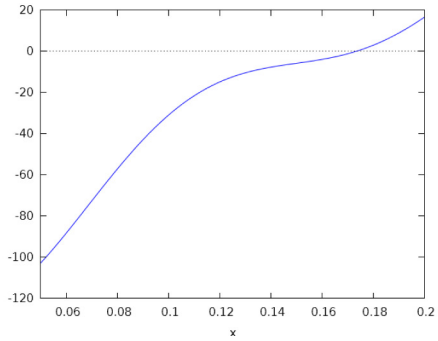


Figure 3: $h(x)$ for points close -1. Derivative of $w(x)$ for points near 0.

It seems reasonable to conjecture that with this method the constant 2.37 in the bound (4) may be replaced with any number less than the constant $c = 3.394\ldots$ from Theorem 2.

4. Concluding Remarks

We conclude the paper by providing another infinite family of small subgroups generated by 2 for which the cancellation may be bounded by some constant. The constant is worse than that in [2], but subgroups are a bit larger.

Proposition 1. *For* $q = \frac{2^{3n}+1}{2^n+1}$ *we have*

$$\max_{(a,q)=1} |s(a/q)| \geq \tau - c' + O\left(\frac{1}{\sqrt{q}}\right)$$

with $c' = 4c = 4 \cdot 2\sum_{r=1}^{\infty} \sin^2\left(\frac{\pi}{2^r}\right) = 13.57\ldots$.

Observe that Theorem 4 (with $\kappa = 2$) gives in this case $\max_{(a,q)=1} |s(a/q)| < \tau - 7.11$. Thus the true value of the maximum for such q is in the range $[\tau - 13.57 - \epsilon, \tau - 7.11]$ if only n is large enough.

Proof. Obviously, we have $\mathcal{L} = 2n$ and $\tau = 6n = 3\mathcal{L}$. Next, observe that $q = 2^{2n} - 2^n + 1$ and so

$$2^{2n} \equiv 2^n - 1 \,(\mathrm{mod}\, q),\ \ 2^{3n} \equiv -1 \,(\mathrm{mod}\, q),\ \ 2^{5n} \equiv -2^n + 1 \,(\mathrm{mod}\, q),\ \ 2^{6n} \equiv 1 \,(\mathrm{mod}\, q).$$

We are going to bound the difference between τ and the real part of the sum. We split the range of the summation into four intervals: $[0, 2n-1], [2n, 3n-1], [3n, 5n-1], [5n, 6n-1]$. We only consider the first two sums as the calculations for the other two are analogous. Using Taylor expansion of a cosine and changing the order of summation (just as in [2]), we arrive at

$$\sum_{r=0}^{2n-1} \left(1 - \cos\left(\frac{2\pi 2^r}{q}\right)\right) = -\sum_{m=1}^{\infty} \frac{(-1)^m}{2m!} \frac{(2\pi)^{2m}}{4^m - 1} \left(\frac{q + 2^n - 1}{q}\right)^{2m} + O\left(\frac{1}{q^2}\right) \quad (7)$$

and

$$\sum_{r=2n}^{3n-1} \left(1 - \cos\left(\frac{2\pi 2^r}{q}\right)\right) = \sum_{r=0}^{n-1} \left(1 - \cos\left(\frac{2\pi 2^r (2^n - 1)}{q}\right)\right) =$$
$$= -\sum_{m=1}^{\infty} \frac{(-1)^m}{2m!} \frac{(2\pi)^{2m}}{4^m - 1} \left(\frac{q+1}{q}\right)^{2m} + O\left(\left(\frac{2^n}{q}\right)^2\right). \quad (8)$$

Now we write the series as a sum of three parts as in [2]. The first part is the same for (7) and (8) and equals

$$\Sigma_1 := -\sum_{m=1}^{\infty} \frac{(-1)^m}{(2m)!} \frac{(2\pi)^{2m}}{4^m - 1} = 2\sum_{r=1}^{\infty} \sin^2\left(\frac{\pi}{2^r}\right).$$

The second part for (7) is equal to

$$\Sigma_2 := -\sum_{m<\sqrt{q}} \frac{(-1)^m}{(2m)!}\frac{(2\pi)^{2m}}{4^m-1}\left(\left(1+\frac{2^n-1}{q}\right)^{2m}-1\right).$$

Using $e^x - 1 \ll x$ we see that $|\Sigma_2| \ll \frac{1}{\sqrt{q}}$.

The second part for (8) is the same as in [2] and also smaller than $\frac{1}{\sqrt{q}}$. The third part for (7) and (8) is negligible (see [2] for details). We infer that

$$\sum_{r=2n}^{3n-1}\left(1-\cos\left(\frac{2\pi 2^r}{q}\right)\right) = c + O\left(\frac{1}{q}\right)$$

and

$$\sum_{r=0}^{2n-1}\left(1-\cos\left(\frac{2\pi 2^r}{q}\right)\right) = c + O\left(\frac{1}{\sqrt{q}}\right).$$

We conclude that $\tau - |\max_{(a,q)=1}|s(a/q)|| \geq 4c + O(\frac{1}{\sqrt{q}})$. □

Acknowledgments. I would like to thank Tomasz Schoen for introducing me to the subject of [2] and for encouraging me to work on this problem. I would also like to thank the anonymous referee for many corrections.

References

[1] J. Bourgain, A.A. Glibichuk and S.V. Konyagin, Estimate for the number of sums and products and for exponential sums in fields of prime order, *J. London Math. Soc.* **73**(2006), 380-398.

[2] J.Kaczorowski and G. Molteni, Extremal values for the sum $\sum_{r=1}^{\tau} e(a2^r/q)$, *J. Number Theory* **132** (2012), 2595-2603.

[3] G. Molteni, Cancellation in a short exponential sum, *J. Number Theory* **130** (2010), 2011-2027.