Unique Differences in Symmetric Subsets of $\mathbb{F}_p$

Tai Do Duc
Division of Mathematical Sciences
School of Physical & Mathematical Sciences
Nanyang Technological University
Singapore 637371
Republic of Singapore

Bernhard Schmidt
Division of Mathematical Sciences
School of Physical & Mathematical Sciences
Nanyang Technological University
Singapore 637371
Republic of Singapore

April 2, 2015

Abstract

Let p be a prime and let A be a subset of $\mathbb{F}_p$ with $A = -A$ and $|A \setminus \{0\}| \leq 2\log_3(p)$. Then there is an element of $\mathbb{F}_p$ which has a unique representation as a difference of two elements of A.

1 Introduction

Let p be a prime and let $\mathbb{F}_p$ denote the field with p elements. Let A be a nonempty subset of $\mathbb{F}_p$. We say that A has a **unique difference** if there is $x \in \mathbb{F}_p$ such that there is exactly one ordered pair (a, b), $a, b \in A$, with

$x = a - b$. Unique sums are defined similarly. A subset B of $\mathbb{F}_p$ is called **symmetric** if $B = -B$.

According to [13], the following problem was first proposed by W. Feit.

Problem 1.1. *Given a prime p, what is the largest number $f(p)$ such that every subset of $\mathbb{F}_p$ with at most $f(p)$ elements has a unique difference?*

Straus [13] showed $f(p) \geq 1 + \log_4(p-1)$. This result was improved by Browkin, Divis, and Schinzel [1] who obtained

$$f(p) \geq \log_2 p, \tag{1}$$

which is the best known lower bound for $f(p)$. It is not known whether this bound is asymptotically sharp.

In [13, Thm. 2], subsets of $\mathbb{F}_p$ were constructed which do not have unique differences and are of cardinality $(2+o(1))\log_3(p)$. These sets are symmetric. Thus [13, Thm. 2] implies

$$g(p) \leq (2 + o(1)) \log_3(p) \tag{2}$$

where $g(p)$ denotes the largest number such that every *symmetric* subset of $\mathbb{F}_p$ with at most $g(p)$ elements has a unique difference.

Note that a symmetric subset of $\mathbb{F}_p$ has a unique difference if and only if it has a unique sum. Thus we might as well formulate the results of this paper in terms of unique sums.

The above-mentioned result of Browkin, Divis, and Schinzel implies

$$g(p) \geq \log_2 p. \tag{3}$$

The following theorem is the main result of this paper. It implies $g(p) \geq 2\log_3(p)$, which is a substantial improvement upon (3).

Theorem 1.2. *Let p be a prime and let A be a symmetric subset of $\mathbb{F}_p$ with $|A \setminus \{0\}| \leq 2\log_3(p)$. Then A has a unique difference.*

In view of Straus' result (2), Theorem 1.2 is sharp in the following sense.

Corollary 1.3. *We have $g(p) \geq 2\log_3(p)$ for every prime p. Moreover, for every $\varepsilon > 0$, there exists a constant $C(\varepsilon)$ such that*

$$g(p) \leq (2+\varepsilon)\log_3(p) \tag{4}$$

for every prime $p > C(\varepsilon)$.

Results like (1) and Theorem 1.2 have applications in various areas, see [1, 5, 6, 7, 8, 9, 10], for instance. In the last section, we present a new application to cyclotomic integers X for which $|X|^2$ is an integer.

2 Preliminaries

In this section, we state some well known results which will be needed later. We include proofs for the convenience of the reader. For a ring R, let $M_{m,n}(R)$ denote the set of $m \times n$ matrices with entries from R. The Euclidean norm of $x \in \mathbb{R}^n$ is denoted by $||x||$.

Result 2.1. *Let $m \leq n$ and let $A \in M_{m,n}(\mathbb{R})$ with rows $r_1, \ldots, r_m$. Set $d(1) = ||r_1||$. For $2 \leq j \leq m$, let $d(j)$ be the distance of r_j from the subspace of $\mathbb{R}^n$ spanned by $r_1, \ldots, r_{j-1}$. We have*

$$\det(AA^T) = \prod_{j=1}^{m} d(j)^2. \tag{5}$$

Proof. If $\mathrm{rank}_{\mathbb{R}}(A) < m$, then left hand and right side of (5) are both zero. Hence we may assume $\mathrm{rank}_{\mathbb{R}}(A) = m$. By Gram-Schmidt orthogonalization, there is a nonsingular lower triangular matrix $L \in M_{m\times m}$ with diagonal entries $d(j)^{-1}$, $j = 1, \ldots, m$, such that the rows of $Q = LA$ are an orthonormal basis of the subspace of $\mathbb{R}^n$ spanned by $r_1, \ldots, r_m$. Thus $QQ^T = I_m$ where I_m denotes the $m \times m$ identity matrix. We conclude $AA^T = L^{-1}QQ^T(L^{-1})^T = L^{-1}(L^{-1})^T$. Hence

$$\det(AA^T) = (\det(L^{-1}))^2 = \prod_{j=1}^{m} d(j)^2.$$

□

Result 2.2. *Let $A \in M_{u,n}(\mathbb{R})$, $B \in M_{w,n}(\mathbb{R})$, and $C = \binom{A}{B}$. Then*

$$\det(CC^T) \leq \det(AA^T)\det(BB^T).$$

Proof. Let $r_1, \ldots, r_u$ be the rows of A and let $r_{u+1}, \ldots, r_{u+w}$ be the rows of B. For $1 \leq i < j \leq u+w$, let $d(i,j)$ denote the distance of r_j from the subspace of $\mathbb{R}^n$ generated by $r_i, \ldots, r_{j-1}$. Furthermore, set $d(i,i) = ||r_i||$ for all i. By Lemma 2.1, we have

$$\begin{aligned}
\det(AA^T) &= \prod_{j=1}^{u} d(1,j)^2, \\
\det(BB^T) &= \prod_{j=u+1}^{u+w} d(u+1,j)^2, \\
\det(CC^T) &= \prod_{j=1}^{u+w} d(1,j)^2.
\end{aligned}$$

Moreover, $d(1,j) \leq d(u+1,j)$ for $j \geq u+1$ by the definition of the $d(i,j)$'s. Hence

$$\begin{aligned}
\det(CC^T) &= \prod_{j=1}^{u+w} d(1,j)^2 \\
&\leq \prod_{j=1}^{u} d(1,j)^2 \prod_{j=u+1}^{u+w} d(u+1,j)^2 \\
&= \det(AA^T)\det(BB^T).
\end{aligned}$$

□

A repeated application of Result 2.2 gives the following.

Corollary 2.3. *Let $A_i \in M_{u_i,n}(\mathbb{R})$, $i = 1, \ldots, k$, and*

$$C = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_k \end{pmatrix}.$$

Then

$$\det(CC^T) \leq \prod_{i=1}^{k} \det(A_i A_i^T).$$

3 Set-up

In this section, we introduce some notation and assumptions which will be implicitly assumed in the rest of the paper. Let p be an odd prime and suppose that A is a symmetric subset of $\mathbb{F}_p$ which has no unique difference.

First of all, Theorem 1.2 trivially holds for $p = 3$. Hence we may assume $p \geq 5$. Write $A = \{\pm a_1, ..., \pm a_n\}$ such that $a_j \neq \pm a_i$ for $i \neq j$. We set $a_n = 0$ if $0 \in A$. Let $m = \lfloor |A|/2 \rfloor$. Note that $m = n$ if $0 \notin A$ and $m = n - 1$ if $0 \in A$.

Next, we show that we may assume $|A| \geq 4$. If $|A| = 1$, then A certainly has a unique difference. If $2 \leq |A| \leq 3$, then $A = \{\pm a_1\}$ or $A = \{0, \pm a_1\}$ for some $a_1 \neq 0$. Then $a_1 - (-a_1) = 2a_1$ is a unique difference in A. Hence we can indeed assume $|A| \geq 4$.

We now set up a linear system arising from A and derive some useful properties of its coefficient matrix. Let i be arbitrary with $1 \leq i \leq m$. As $2a_i = a_i - (-a_i)$ is not a unique difference in A and $2a_i \neq \pm a_i - a_i$, there exists an ordered pair $(\sigma(i), \tau(i)) \neq (i, i)$ with

$$2a_i \pm a_{\sigma(i)} \pm a_{\tau(i)} = 0. \tag{6}$$

Here "$\pm a_{\sigma(i)} \pm a_{\tau(i)}$" means that *any* combination of signs is possible including $a_{\sigma(i)} - a_{\tau(i)}$ and $-a_{\sigma(i)} + a_{\tau(i)}$. We use this convention throughout the rest of the paper.

We consider the homogeneous linear system corresponding the these equations:

$$2x_i \pm x_{\sigma(i)} \pm x_{\tau(i)} = 0, \ i = 1, \ldots, m. \tag{7}$$

Here we use the convention $x_n = 0$ if $0 \in A$ (and thus $a_n = 0$).

Note that the coefficient vectors corresponding to the system (7) all have at most 3 nonzero entries. The nonzero coefficients, however, are not necessarily $2, \pm 1$, since $i, \sigma(i), \tau(i)$ are not necessarily distinct. We now determine exactly which coefficient vectors can occur.

Case 1 $\sigma(i) = i$ or $\tau(i) = i$. By symmetry, we can assume $\tau(i) = i$. Thus $2a_i \pm a_{\sigma(i)} \pm a_i = 0$. As $(\sigma(i), \tau(i)) \neq (i, i)$, we have $\sigma(i) \neq i$. If

$2a_i \pm a_{\sigma(i)} - a_i = 0$, then $a_i \pm a_{\sigma(i)} = 0$, contradicting the assumption $a_i \neq \pm a_j$ for $i \neq j$. Hence $2a_i \pm a_{\sigma(i)} + a_i = 3a_i \pm a_{\sigma(i)} = 0$. Recall that we assume $p \geq 5$. If $a_{\sigma(i)} = 0$, then $3a_i = 0$ and thus $a_i = 0$, contradicting the assumptions. Thus $a_{\sigma(i)} \neq 0$, i.e., $\sigma(i) \leq m$.

Hence (6) can we written as

$$3x_i \pm x_{\sigma(i)} = 0 \tag{8}$$

with $\sigma(i) \neq i$ and $\sigma(i) \leq m$. We call (8) an equation of **type 1**.

Case 2 $\sigma(i) \neq i$ and $\tau(i) \neq i$. If $\tau(i) = \sigma(i)$, then (6) implies $a_i = 0$ or $a_i = \pm a_{\sigma(i)}$, contradicting the assumptions. Thus $\tau(i) \neq \sigma(i)$. Hence $i, \sigma(i), \tau(i)$ are pairwise distinct.

Recall that $x_n = 0$ if $0 \in A$. Hence, if $0 \in A$ and $n \in \{\sigma(i), \tau(i)\}$, then one of the variables $x_{\sigma(i)}, x_{\tau(i)}$ occurs with coefficient zero in (6). By symmetry, we can assume that $x_{\tau(i)}$ occurs with coefficient zero in this case.

Hence, in Case 2, we can write (6) as

$$\begin{aligned} 2x_i \pm x_{\sigma(i)} &= 0 \ \text{ if } 0 \in A \text{ and } \tau(i) = n, \\ 2x_i \pm x_{\sigma(i)} \pm x_{\tau(i)} &= 0 \ \text{ otherwise,} \end{aligned} \tag{9}$$

where $i, \sigma(i), \tau(i)$ are pairwise distinct. In both cases, we call (9) an equation of **type 2**.

In summary, as Case 1 and Case 2 cover all possible cases, for every $i \in \{1, \ldots, m\}$, one of the following equations is contained in the linear system (7).

$$\begin{aligned} 3x_i \pm x_{\sigma(i)} &= 0 \ \text{ (type 1)}, \\ 2x_i \pm x_{\sigma(i)} &= 0 \ \text{ (type 2)}, \\ 2x_i \pm x_{\sigma(i)} \pm x_{\tau(i)} &= 0 \ \text{ (type 2)}. \end{aligned}$$

Furthermore, the following hold.

- $\sigma(i) \leq m$ and $\tau(i) \leq m$,
- $i, \sigma(i), \tau(i)$ are pairwise distinct.

Of course, the statements involving $\tau(i)$ only apply if the equation $2x_i \pm x_{\sigma(i)} \pm x_{\tau(i)} = 0$ is contained (7).

We use a similar terminology for the coefficient vectors of the system (7): We say a coefficient vector is of **type 1** if it has exactly one entry 3, exactly one entry ± 1, and all its remaining entries are zero. A coefficient vector is of **type 2** if it has exactly one entry 2, at most two entries ± 1, and all its remaining entries are zero.

4 A Congruence for Minors of the Coefficient Matrix

Let M be the coefficient matrix M of the linear system (7). Recall that all entries of M are from $\{0, \pm 1, 2, 3\}$. Note that M can be considered as a matrix with rational entries as well as a matrix with entries from $\mathbb{F}_p$. In the following, we make use of both interpretations. Let r be the rank of the coefficient matrix M over $\mathbb{Q}$. We now prove that all $r \times r$-minors of M are divisible by p. This result is useful, as it can be combined with estimates for minors of M which we get from Result 2.1.

Theorem 4.1. *Suppose that $A = \{\pm a_1, ..., \pm a_n\} \subset \mathbb{F}_p$ does not have a unique difference and let M be the coefficient matrix of the corresponding linear system (7). Write $r = \mathrm{rank}_{\mathbb{Q}}(M)$. Every $r \times r$-minor of M is divisible by p.*

Proof. Recall that M is an $m \times m$-matrix and that M has entries from $\{0, \pm 1, 2, 3\}$ only. In this proof, we consider the entries of M as *integers* (not as elements of $\mathbb{F}_p$), and we will work with the Smith Normal Form of M over the integers.

Let $D = \mathrm{diag}(d_1, ..., d_r, 0, ..., 0)$ be the Smith Normal Form of M. Note that $d_1, \ldots, d_r$ are integers such that d_i divides d_{i+1} for $i = 1, \ldots, r-1$. Furthermore, there are $B, C \in M_{m,m}(\mathbb{Z})$ with $\det(B) \pm 1$, $\det(C) = \pm 1$, and $M = BDC$. It is important for this proof to note that B, C, D are matrices with *integer* entries and the equation $M = BDC$ holds over $\mathbb{Z}$ (not only over $\mathbb{F}_p$). Furthermore, note that all entries of C^{-1} are integers, as $\det(C) = \pm 1$.

Write $x = (a_1, \ldots, a_m)^T$. Note $x \in \mathbb{F}_p^m$. We have $x \neq 0$, as $|A| > 1$. Recall that $Mx = 0$ over $\mathbb{F}_p$ by (7).

Suppose that p does not divide d_r. Then p does not divide any of the integers $d_1, \ldots, d_r$. Write $Cx = (b_1, \ldots, b_m)^T$ where $b_1, \ldots, b_m \in \mathbb{F}_p$. Note that $DCx = 0$ over $\mathbb{F}_p$ if and only if $b_1 = \cdots = b_r = 0$, since we assume that $d_1, \ldots, d_r$ are not divisible by p. Hence $DCx = 0$ if and only if $x = C^{-1}(0, \ldots, 0, b_{r+1}, \ldots, b_m)^T$ for some $b_i \in \mathbb{F}_p$. As $BDCx = Mx = 0$ over $\mathbb{F}_p$, we indeed have $DCx = 0$ over $\mathbb{F}_p$. Thus

$$x = C^{-1}b \text{ with } b = (0, \ldots, 0, b_{r+1}, \ldots, b_m)^T \in \mathbb{F}_p^m. \tag{10}$$

For convenience, we set $b_i = 0$ for $i = 1, \ldots, r$.

We now switch from equations over $\mathbb{F}_p$ to equations over $\mathbb{Z}$. We first define a vector c with *integer entries* which represent the residue classes $b_1, \ldots, b_m$ mod p. Formally, let $c_1, \ldots, c_m$ be the unique integers with $0 \leq c_i \leq p-1$ and

$$b_i = c_i + p\mathbb{Z}/\mathbb{Z} \tag{11}$$

(here we use the standard notation $\mathbb{F}_p = \{k + p\mathbb{Z}/\mathbb{Z} : k = 0, \ldots, p-1\}$). Note that $c_1 = \cdots = c_r = 0$, as $b_1 = \cdots = b_r = 0$ (again, note that $c_1 = \cdots = c_r = 0$ are equations over $\mathbb{Z}$, not only over $\mathbb{F}_p$).

Define $y = (y_1, \ldots, y_m) \in \mathbb{Z}^m$ by

$$y = C^{-1}(0, \ldots, 0, c_{r+1}, \ldots, c_m)^T \tag{12}$$

(recall that the c_i's are considered as integers, not as elements of $\mathbb{F}_p$). Then we have

$$My = BDCy = B\,\mathrm{diag}(d_1, ..., d_r, 0, ..., 0)(0, \ldots, 0, c_{r+1}, \ldots, c_m)^T = 0 \tag{13}$$

over $\mathbb{Z}$ (note that all entries of matrices and vectors occurring in (13) are considered as integers and to derive (13), we need the fact that the equation $M = BDC$ holds over $\mathbb{Z}$, not only over $\mathbb{F}_p$).

Let Γ_j by the jth row of C^{-1}, $j = 1, \ldots, m$. By (10) and (12), we have

$$a_j = \Gamma_j b \text{ and } y_j = \Gamma_j c \tag{14}$$

where $b = (0, \ldots, 0, b_{r+1}, \ldots, b_m)^T$ and $c = (0, \ldots, 0, c_{r+1}, \ldots, c_m)^T$. Note that (11) and (14) imply

$$a_j = y_j + p\mathbb{Z}/\mathbb{Z},\ i = 1, \ldots, m. \tag{15}$$

Recall that M is the coefficient matrix of the linear system (7) and that $My = 0$ over $\mathbb{Z}$ by (13). This implies that $\{\pm y_1, \ldots, \pm y_m\}$ does not have a unique difference (where the differences are taken in $\mathbb{Z}$). Recall that $a_i \neq \pm a_j$ for $i \neq j$ by assumption. Thus (15) implies $y_i \neq \pm y_j$ for $i \neq j$. Hence there is $k \in \{1, \ldots, m\}$ such that $|y_k| > |y_i|$ for all $i \neq k$. This implies that $2y_k = y_k - (-y_k)$ is a unique difference of $\{\pm y_1, \ldots, \pm y_m\}$, a contradiction. We conclude that p divides d_r.

From the theory of the Smith Normal Form (see [11, p. 41], for instance), it is well known that d_r divides the greatest common divisor of all $r \times r$-minors of M. □

5 Equations of Type 1

Equations of type 1 play a critical role in the proof of Theorem 1.2, as the Euclidean norm of their coefficient vectors is the largest among the equations occurring in the linear system (7). In this section, we study the structure of the set of equations of type 1 contained in (7). Recall that equations of type 1 have the form

$$3x_i \pm x_{\sigma(i)} = 0$$

where $1 \leq i \leq m$, $\sigma(i) \neq i$, and $\sigma(i) \leq m$.

Lemma 5.1. *Suppose $3^{\lfloor |A|/2 \rfloor} \leq p$. Let I be a subset of $\{1, \ldots, m\}$ such that $3x_i \pm x_{\sigma(i)} = 0$, $i \in I$, are equations of type 1 contained in (7). Let G be the directed graph with vertex set $I \cup \{\sigma(i) : i \in I\}$ and edge set $E = \{(i, \sigma(i)) : i \in I\}$. Then E can be decomposed into directed paths which are pairwise vertex disjoint.*

Proof. We show

(i) every vertex of G has outdegree at most 1,

(ii) every vertex of G has indegree at most 1,
(iii) G does not contain a directed cycle.

Note that (i-iii) imply that G indeed can be decomposed into directed paths which are pairwise vertex disjoint.

First of all, there is at most one edge $(i, \sigma(i))$ for every vertex i. This means that the outdegree of every vertex in G is at most one.

Suppose that the indegree of a vertex i is at least 2. Then there are distinct vertices j, k with $\sigma(j) = \sigma(k) = i$. By definition, this implies $3a_j \pm a_i = 0$ and $3a_k \pm a_i = 0$ and thus $3(a_j - a_k) = 0$ or $3(a_j + a_k) = 0$. As $p > 3$, we conclude $a_j = \pm a_k$ which contradicts the assumption $a_j \neq \pm a_k$ for $j \neq k$. This shows that all vertices of G have indegree at most 1.

Now suppose that G contains a directed cycle. Then there are vertices $v_1, v_2, \ldots, v_k$ with $v_k = v_1$ such that $(v_i, v_{i+1}) \in E$ for $i = 1, \ldots, k-1$, i.e., $v_{i+1} = \sigma(v_i)$ for $i = 1, \ldots, k-1$. By definition, this implies $3a_{v_i} = \pm a_{v_{i+1}}$ for $i = 1, \ldots, k-1$. Hence

$$a_{v_1} = a_{v_k} = \pm 3a_{v_{k-1}} = \pm 9a_{v_{k-2}} = \cdots = \pm 3^{k-1} a_{v_1}.$$

Thus $(\pm 3^{k-1} - 1)a_{v_1} = 0$. Note that $k \leq m + 1$, as the cycle visits every vertex $\neq v_1$ at most once. Moreover, $p \geq 3^m = 3^{\lfloor |A|/2 \rfloor}$ by assumption and thus $p \geq 3^m + 2$, as $p > 3$ is a prime. We conclude $3^{k-1} + 1 \leq 3^m + 1 < p$. Hence $(\pm 3^{k-1} - 1)a_{v_1} = 0$ implies $a_{v_1} = 0$, contradicting the assumption $a_i \neq 0$ for $i = 1, \ldots, m$. This shows that G does not contain a directed cycle.

In summary, we have shown that (i)-(iii) hold, and this completes the proof. □

Remark 5.2. Note that Lemma 5.1 implies that G does not contain *any* cycle, directed or undirected.

Next, we compute determinants arising from matrices whose rows are coefficient vectors of equations of type 1.

Lemma 5.3. *Let G be the graph defined in Lemma* 5.1. *Suppose J is a subset of $\{1, \ldots, m\}$ such that $\{(i, \sigma(i)) : i \in J\}$ is a directed path in G and let B be*

the coefficient matrix of the corresponding equations $3x_i \pm x_{\sigma(i)} = 0$, $i \in J$. *Then*

$$\det(BB^T) = \frac{1}{8}\left(-1 + 9^{|J|+1}\right). \tag{16}$$

Proof. Write $v = |J|$. By relabeling vertices of G, if necessary, we may assume that the directed path consists of the edges $(i, i+1)$, $i = 1, \ldots, v$. The corresponding equations are $3x_i \pm x_{i+1} = 0$, $i = 1, \ldots, v$. Hence B is a $v \times m$-matrix of the form

$$\begin{pmatrix} 3 & \pm 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 3 & \pm 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 3 & \pm 1 & 0 & \cdots & 0 \end{pmatrix}.$$

Thus

$$BB^T = \begin{pmatrix} 10 & 3\delta_1 & 0 & \cdots & 0 & 0 & 0 \\ 3\delta_1 & 10 & 3\delta_2 & \cdots & 0 & 0 & 0 \\ 0 & 3\delta_2 & 10 & 3\delta_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \\ 0 & \cdots & 0 & 3\delta_{v-3} & 10 & 3\delta_{v-2} & 0 \\ 0 & \cdots & 0 & 0 & 3\delta_{v-2} & 10 & 3\delta_{v-1} \\ 0 & \cdots & 0 & 0 & 0 & 3\delta_{v-1} & 10 \end{pmatrix}$$

with $\delta_i = \pm 1$, $i = 1, \ldots, v-1$. We now prove (16) by induction on v. It is straightforward to check that (16) holds for $v = 1$. Suppose $v \geq 2$. Using Laplace expansion with respect to the last column of BB^T and the inductive hypothesis, we find

$$\det(BB^T) = 10\left(\frac{1}{8}\left(-1 + 9^v\right)\right) - (3\delta_{v-1})^2\left(\frac{1}{8}\left(-1 + 9^{v-1}\right)\right).$$

Note $(3\delta_{v-1})^2 = 9$. Hence

$$\det(BB^T) = \frac{1}{8}\left(-10 + 9 + (10-1)9^v\right) = \frac{1}{8}\left(-1 + 9^{v+1}\right).$$

□

Corollary 5.4. *Suppose* $3^{\lfloor |A|/2 \rfloor} \leq p$. *Let* I *be a subset of* $\{1, \ldots, m\}$ *such that* $3x_i \pm x_{\sigma(i)} = 0$, $i \in I$, *are equations of type* 1. *Let* G *be the directed*

graph defined in Lemma 8 and let $P_1, \ldots, P_t$ be vertex disjoint directed paths in G with $E = \bigcup_{j=1}^{t} P_j$. Let B be the matrix whose rows are the coefficient vectors of the equations $3x_i \pm x_{\sigma(i)} = 0$, $i \in I$. Then

$$\det(BB^T) < 11^t 9^{|I|-t}.$$

Proof. For $j = 1, \ldots, t$, let B_j be the matrix whose rows are the coefficient vectors of the equations $3x_i \pm x_{\sigma(i)} = 0$, $(i, \sigma(i)) \in P_j$. By Lemma 5.3, we have

$$\det(B_j B_j^T) = \frac{1}{8}\left(-1 + 9^{|P_j|+1}\right).$$

Note that

$$\frac{1}{8}\left(-1 + 9^{x+1}\right) < 11 \cdot 9^{x-1}$$

for all $x \geq 1$. Using Corollary 2.3 and $\sum_{j=1}^{t} |P_j| = |I|$, we conclude

$$\begin{aligned}
\det(BB^T) &\leq \prod_{j=1}^{t} \det(B_j B_j^T) \\
&= \prod_{j=1}^{t} \left(\frac{1}{8}\left(-1 + 9^{|P_j|+1}\right)\right) \\
&< \prod_{i=1}^{t} \left(11 \cdot 9^{|P_j|-1}\right) \\
&= 11^t 9^{|I|-t}.
\end{aligned}$$

□

6 Proof of Theorem 1.2

Let p be prime and let A be a symmetric subset of $\mathbb{F}_p$ with

$$|A \setminus \{0\}| \leq 2\log_3(p). \tag{17}$$

We have to prove that A has a unique difference. As shown in Section 3, we may assume $p \geq 5$.

Suppose A does not have a unique difference. Write $A = \{\pm a_1, \ldots, \pm a_n\}$ as in Section 3 where $a_n = 0$ if $0 \in A$. Recall $m = \lfloor A/2 \rfloor$. Note $|A \setminus \{0\}| = 2\lfloor A/2 \rfloor$. Hence $3^{\lfloor A/2 \rfloor} \leq p$ by (17) and thus

$$3^{\lfloor A/2 \rfloor} < p, \tag{18}$$

as $p \neq 3$.

Let M be the coefficient matrix of the linear system (7) and write $r = \operatorname{rank}_{\mathbb{Q}}(M)$. Note $r \leq m$. Let N be a nonsingular $r \times r$ submatrix of M.

We claim that

$$|\det(N)| \leq 3^r. \tag{19}$$

We say that a row vector is of **type 3** if it has exactly one entry 3 and all its remaining entries are zero. Note that N may contain rows of type 3, as some rows of M of type 1 may turn into type 3 when columns of M are deleted. If N contains a row of type 3, then, by Laplace expansion, $|\det(N)| = 3|\det(N_1)|$ where N_1 is a $(r-1) \times (r-1)$ submatrix of M. Repeating this process, if necessary, we either get $|\det(N)| = 3^r$ or obtain a $d \times d$ submatrix N_2 of M with

$$|\det(N)| = 3^{r-d}|\det(N_2)| \tag{20}$$

such that N_2 does not contain any row of type 3. If $|\det(N)| = 3^r$, then (19) holds. Thus we may assume that (20) holds.

As N_2 is a submatrix of M, all rows of N_2 which are not of type 1 have at most at most three nonzero entries ± 1, 2, at most one of which is 2. Hence every row of N_2 which is not of type 1 has Euclidean norm at most $\sqrt{6}$.

Swapping rows, if necessary, we can write

$$N_2 = \begin{pmatrix} B \\ C \end{pmatrix}$$

where B consists of the rows of N_2 of type 1 and C of the remaining rows of N_2. Let I be the subset of $\{1, \ldots, m\}$ such that $3x_i \pm x_{\sigma(i)} = 0$, $i \in I$, are the equations corresponding to the rows of B. Note that $|I|$ is the number of rows of B and $d - |I|$ is the number of rows of C.

Let G be the directed graph with vertex set $V = I \cup \{\sigma(i) : i \in I\}$ and edge set $E = \{(i, \sigma(i)) : i \in I\}$. Note that $|V| \leq d$, as every element of V is an index of a column of N_2.

By Lemma 5.1, there is a decomposition of E into pairwise vertex disjoint directed paths. Note that these directed paths are connected components of G. Let t be the number of paths in the decomposition. As G has at most d vertices and contains no cycles, it has at most $|V| - |E|$ connected components. Hence

$$t \leq |V| - |E| \leq d - |I|. \tag{21}$$

By Corollary 5.4, we have

$$\det(BB^T) < 11^t 9^{|I|-t}. \tag{22}$$

From (21) and (22), we get

$$\det(BB^T) < 11^{d-|I|} 9^{|I|-(d-|I|)} = 11^{d-|I|} 9^{2|I|-d}. \tag{23}$$

As the rows of C all have Euclidean norm at most $\sqrt{6}$, we have

$$\det(CC^T) \leq 6^{d-|I|} \tag{24}$$

by Result 2.1. From Result 2.2, we get

$$\det(N_2 N_2^T) \leq \det(BB^T) \det(CC^T). \tag{25}$$

Putting (23-25) together, we find

$$\begin{aligned} \det(N_2 N_2^T) &< 11^{d-|I|} 9^{2|I|-d} 6^{d-|I|} \\ &= 66^{d-|I|} 9^{2|I|-d} \\ &\leq 81^{d-|I|} 9^{2|I|-d} \\ &= 9^d. \end{aligned}$$

Hence $|\det(N_2)| \leq 3^d$ and thus $|\det(N)| \leq 3^r$ by (20). This proves (19).

Finally, recall that $\det(N)$ is a nonzero $r \times r$-minor of M. Hence $|\det(N)| \geq p$ by Theorem 4.1. But $r \leq \lfloor |A|/2 \rfloor$ and thus $|\det(N)| \leq 3^r < p$ by (18), a contradiction. This completes the proof of Theorem 1.2. $\square$

7 Application to Weil Numbers

Let m, n be positive integers. Write $\zeta_m = \exp(2\pi i/m)$. An ***n*-Weil number** in $\mathbb{Z}[\zeta_m]$ is an element Y of $\mathbb{Z}[\zeta_m]$ with $|Y|^2 = n$. In this section, we show that Theorem 1.2 implies that under certain conditions Weil numbers in $\mathbb{Z}[\zeta_m]$ necessarily are contained in proper subfields of $\mathbb{Q}(\zeta_m)$. This result is a partial improvement of the "field descent" introduced in [12] and is relevant for the study of difference sets and related objects. We will assume basic algebraic number theory in this sections, as treated in [4], for instance.

For a finite group G and a ring R, let $R[G]$ denote the group ring of G over R. Every element B of $R[G]$ can be written as $B = \sum_{g\in G} r_g g$ with $r_g \in R$. The r_g's are called the **coefficients** of B. We write $B^{(-1)} = \sum_{g\in G} \overline{r_g} g^{-1}$ where $\overline{r_g}$ is the complex conjugate of r_g.

For $Y \in \mathbb{Z}[\zeta_m]$, let

$$\mathcal{M}(Y) = \frac{1}{\varphi(m)} \sum_{\sigma\in \mathrm{Gal}(\mathbb{Q}(\zeta_m)/\mathbb{Q})} (Y\overline{Y})^\sigma,$$

where φ denote the Euler totient function. Note

$$\mathcal{M}(Y) \geq 1 \tag{26}$$

for $Y \neq 0$ by the inequality of geometric and arithmetic means, since $\prod (Y\overline{Y})^\sigma$ is the norm of an algebraic integer and thus $\prod (Y\overline{Y})^\sigma \geq 1$. The following is due to Cassels [3].

Result 7.1. *Let $X \in \mathbb{Z}[\zeta_m]$ where $m = pm'$ and p is a prime with $(p, m') = 1$. Write $X = \sum_{i=0}^{p-1} X_i \zeta_p^i$ with $X_i \in \mathbb{Z}[\zeta_{m'}]$. We have*

$$(p-1)\mathcal{M}(X) = \sum_{i<j}^{p-1} \mathcal{M}(X_i - X_j).$$

We denote the cyclic group of order k by C_k.

Theorem 7.2. *Let p, q be distinct primes and r be a positive integer with $\gcd(r, pq) = 1$. Let $n = q^b$ where b is a positive integer. Suppose that $Y\bar{Y} = n$ for some $Y \in \mathbb{Z}[\zeta_{pr}]$. If $\mathrm{ord}_p(q)$ is even and $p > \max\{3^{n/2}, n^2 + n + 1\}$, then $Y\eta \in \mathbb{Z}[\zeta_r]$ for some root of unity η.*

Proof. Write $\mathrm{ord}_p(q) = 2f$, and define $\sigma \in \mathrm{Gal}(\mathbb{Q}(\zeta_{pr})/\mathbb{Q})$ by $\zeta_{pr}^{\sigma} = \zeta_{pr}^{q^f}$. Note that σ fixes all prime ideals above n in $\mathbb{Z}(\zeta_{pr})$ (see [12, Thm. 2.1], for instance). Hence $Y^{\sigma} = Y\alpha$ for some unit α. Note $|\alpha|^2 = Y^{\sigma}\overline{Y^{\sigma}}/(Y\overline{Y}) = n^{\sigma}/n = 1$. Thus α is a root of unity, i.e., $\alpha = \pm\zeta_p^c\zeta_r^d$ for some integers c, d. Let e be an integer with $2e \equiv c \pmod{p}$. Note that $\zeta_p^{\sigma} = \zeta_p^{-1}$ and thus

$$(Y\zeta_p^e)^{\sigma} = Y(\pm\zeta_p^c\zeta_r^d)\zeta_p^{-e} = Y(\pm\zeta_p^{c-e}\zeta_r^d) = Y(\pm\zeta_p^e\zeta_r^d).$$

Let $Y_1 = Y\zeta_p^e$. Then

$$Y_1^{\sigma} = Y_1(\pm\zeta_r^d). \tag{27}$$

Write $Y_1 = \sum_{i=0}^{p-1} a_i\zeta_p^i$ with $a_i \in \mathbb{Z}[\zeta_r]$. Note $\mathcal{M}(Y_1) = \mathcal{M}(Y) = n$. By Result 7.1, we have

$$(p-1)n = (p-1)\mathcal{M}(Y_1) = \sum_{i<j}^{p-1} \mathcal{M}(a_i - a_j). \tag{28}$$

Let t be the maximum number such that there are distinct indices $i_1, \ldots, i_t$ with $a_{i_1} = \cdots = a_{i_t}$. If $t \le p/2$, then, by (26), the right hand side of (28) is at least $p^2/4$ and thus $4n > p$, contradicting the assumption $p > n^2 + n + 1$. Hence $t \ge p/2$. Note that, by (26), the right hand side of (28) is at least $(p-t)t$. Hence $(p-t)t \le n(p-1)$. Note that $(p-t)t$ is decreasing for $t \in [p/2, p]$ and that $(p-(p-n-1))(p-n-1) = n(p-1)+p-n-1-n^2 > n(p-1)$, as $p > n^2+n+1$. Hence $t \ge p-n$. Recall that $a_{i_1} = \cdots = a_{i_t}$. Note $Y_1 = \sum_{i=0}^{p-1}(a_i - a_{i_1})\zeta_p^i$. Thus, writing $b_i = a_i - a_{i_1}$, we have $Y_1 = \sum_{i=0}^{p-1} b_i\zeta_p^i$ and $|\{i : b_i \ne 0\}| \le n$. If $|\{i : b_i \ne 0\}| = 1$, then the assertion of Theorem 7.2 holds. Hence, to complete the proof, it suffices to show that

$$2 \le |\{i : b_i \ne 0\}| \le n \tag{29}$$

leads to a contradiction. Define $X = \sum b_i g^i$ where g is a generator of C_p, the cyclic group of order p. Let K be the kernel of the ring homomorphism $\rho : \mathbb{Z}[\zeta_r][C_p] \to \mathbb{Z}[\zeta_{pr}]$ determined by $g \mapsto \zeta_p$. It is well known and straightforward to verify that $K = \{AC_p : A \in \mathbb{Z}[\zeta_r][C_p]\}$. Note that $\rho(X) = Y_1$, $\rho(X^{(-1)}) = \overline{Y_1}$ and thus $\rho(XX^{(-1)}) = Y_1\overline{Y_1} = n$. Hence

$$XX^{(-1)} = n + AC_p \tag{30}$$

for some $A \in \mathbb{Z}[\zeta_r][C_p]$. Suppose $A \neq 0$. Then there are at least $p-1$ nonzero coefficients on the right hand side of (30). On the other hand, by (29), there are at most $n(n-1)$ nonzero coefficients on the left hand side of (30). Hence $p \leq n(n-1)+1$ which contradicts the assumptions. Thus $A = 0$ and hence

$$XX^{(-1)} = n. \tag{31}$$

Recall that $X = \sum b_i g^i$. By (29), we can write $X = \sum_{j=1}^{z} b_{i_j} g^{i_j}$ with $2 \leq z \leq n$ and $b_{i_j} \neq 0$ for all j. Note

$$XX^{(-1)} = \sum_{k=0}^{p-1} \left(\sum_{i_r - i_s \equiv k \bmod p} b_{i_r} \overline{b_{i_s}} \right) g^k. \tag{32}$$

Write $S = \{i_1, \ldots, i_z\}$ and view S as a subset of $\mathbb{F}_p$. Note $2 \leq |S| \leq n$. Suppose S has a unique difference, say, $k = i_r - i_s \in \mathbb{F}_p$ is a unique difference in S. Note that $k \neq 0$, as $|S| \geq 2$ and thus 0 is not a unique difference in S. In view of (32), the coefficient of g^k in $XX^{(-1)}$ is nonzero. But this contradicts $XX^{(-1)} = n$. Thus S has no unique difference.

Using (27), we get

$$Y_1^\sigma = \sum_{i=0}^{p-1} b_i^\sigma \zeta_p^{-i} = \pm \zeta_r^d Y_1 = \pm \zeta_r^d \sum_{i=0}^{p-1} b_i \zeta_p^i$$

Hence, for $i \neq 0$, we have $b_i \neq 0$ if and only if $b_{p-i} \neq 0$. This implies that S is symmetric. As S has no unique difference, we have $p \leq 3^{|S|/2}$ by Theorem 1.2. As $|S| \leq n$, we conclude $p \leq 3^{n/2}$, contradicting the assumptions. □

Example Let $p = 107$ and $n = 8$ in Theorem 7.2. Note that $\mathrm{ord}_p(2)$ is even and that $p > \max\{3^{n/2}, n^2 + n + 1\}$. Thus, for any prime odd prime r and $Y \in \mathbb{Z}[\zeta_{107r}]$ with $|Y|^2 = 8$, we have $Y\eta \in \mathbb{Z}[\zeta_r]$ for some root of unity η.

References

[1] J. Browkin, B. Divis, and A. Schinzel: Addition of sequences in general fields, *Monatsh. Math.* **82** (1976), 261–268.

[2] E. Croot, T. Schoen: On sumsets and spectral gaps. *Acta Arith.* **136** (2009), 47–55.

[3] J. W. S. Cassels: On a conjecture of R. M. Robinson about sums of roots of unity. *J. Reine Angew. Math.* **238** (1969), 112–131.

[4] K. Ireland, M. Rosen: *A Classical Introduction to Modern Number Theory*. Graduate Texts in Mathematics **84**, Springer 1990.

[5] K. H. Leung, B. Schmidt: Unique Sums and Differences in Finite Abelian Groups. Submitted.

[6] V. F. Lev: The rectifiability threshold in abelian groups. *Combinatorica* **28** (2008), 491–497.

[7] J. H. Loxton: On two problems of E. M. Robinson about sums of roots of unity. *Acta Arith.* **26** (1974), 159–174.

[8] Z. Nedev: An algorithm for finding a nearly minimal balanced set in $\mathbb{F}_p$. *Math. Comp.* **268** (2009), 2259–2267.

[9] Z. Nedev: Lower bound for balanced sets. *Theoret. Comput. Sci.* **460** (2012), 89–93.

[10] Z. Nedev, A. Quas: Balanced sets and the vector game. *Int. J. Number Theory* **4** (2008), 339–347.

[11] C. Norman: *Finitely Generated Abelian Groups and Similarity of Matrices over a Field*. Springer 2012.

[12] B. Schmidt: Cyclotomic integers and finite geometry. *J. Am. Math. Soc.* **12** (1999), 929–952.

[13] E. G. Straus: Differences of residues (mod p). *J. Number Th.* **8** (1976), 40–42.