ANALOGUES OF THE ERDŐS INTEGER DISTANCE PRINCIPLE IN FINITE FIELDS

BENJAMIN DEES

1. Introduction

The Erdős Integer Distance Principle shows that the distances that a set in $\mathbb{R}^n$ determines can provide information about the structure of the set itself. In particular, it shows that if $\Delta(E) \subset \mathbb{Z}$, where $\Delta(E) = \{|x-y| : x, y \in E\}$, then E is a subset of a line. It is not difficult to show, however, that there are arbitrarily large finite subsets of $\mathbb{R}^2$ not contained in a line that still determine only integral distances.

We can ask similar questions in other contexts: what if the distance set $\Delta(E)$ is a subset of another subset of $\mathbb{R}$? What if we consider other subrings of other fields (possibly with different notions of distance), or other operators similar to these notions of distance? Given some set of points A determining (say) integral distances, is there some "maximal" set E such that $A \subset E$ and E determines only integral distances? This paper investigates some of these questions.

2. Background

The motivating example for these results, as mentioned above, is the Erdős Integer Distance Principle itself. A proof of this result, which can be found in [3], for a subset of $\mathbb{R}^2$ follows, as it is in this vein that the results of this paper are proved.

Theorem 2.1. *If $E \subset \mathbb{R}^2$ is such that $\Delta(E) = \{|x-y| : x, y \in E\}$ is contained in $\mathbb{Z}$, then E is finite or E is a subset of a line.*

Proof. Suppose that E is not a subset of any line, and $\Delta(E) \subset \mathbb{Z}$. Then we can take three points x, y, z which are noncolinear, and let $|x-y| = a$, $|x-z| = b$, where $a, b \in \mathbb{Z}$. For any other $m \in E$, $|m-x|$, $|m-y|$, and $|m-z|$ are integers as well, and moreover, by the triangle inequality:

$$\big|\big(|x-m| - |y-m|\big)\big| \leq |x-y| = a \qquad \text{and} \qquad \big|\big(|x-m| - |z-m|\big)\big| \leq |x-z| = b$$

However, since $|x-m|$, $|y-m|$, and $|z-m|$ are all integers, their differences are integers as well, so that the quantities on the left hand sides of the above inequalities are nonnegative integers. Then, recognizing that $\big|\big(|x-m| - |y-m|\big)\big| = d$ describes a hyperbola, we find that every point of E lies on one of $a+1$ hyperbolae with foci at x and y and on one of $b+1$ hyperbolae with foci at x and z.

Since x, y, z are noncolinear, a hyperbola with foci at x and y will intersect a hyperbola with foci at x and z in at most four points. Then, since every point of E corresponds to one such point, and there are $(a+1)(b+1)$ ways to choose a pair of hyperbolae in the above manner:

$$|E| \leq 4(a+1)(b+1) < \infty$$

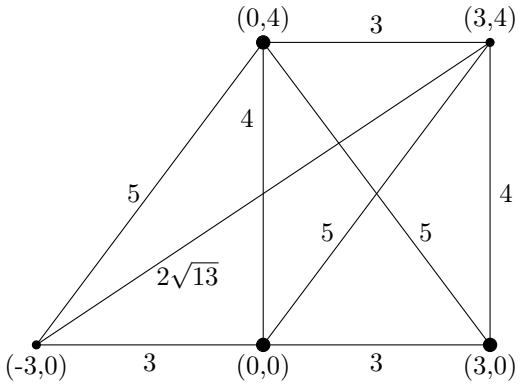
□

An immediate consequence of this theorem is that if we have a set $A \subset \mathbb{R}^2$ such that $\Delta(A) \subset \mathbb{Z}$ and A is not a subset of a line, then there is a "largest set" containing A that determines integral distances. More formally, there is a set E such that $A \subset E$, $\Delta(E) \subset \mathbb{Z}$, and the property that if $|F| > |E|$, $A \subset F$, then $\Delta(F) \nsubseteq \mathbb{Z}$.

In some settings, it is natural to expect that such a "largest" object would be unique or a maximum in some sense other than size (e.g. to be a maximum under inclusion) but this is not the case here. To see this, consider the following example.

Example 2.1: $A = \{(0,0), (3,0), (0,4)\}$ (see diagram on next page) then A has the desired properties, and we can extend A by adding either of the points $(-3,0)$ or $(3,4)$ to it and the new set will still determine integral distances, but $|(3,4) - (-3,0)| = \sqrt{6^2+4^2} = \sqrt{52} \notin \mathbb{Z}$, so $A \cup \{(-3,0), (3,4)\}$ does not determine integral distances. Thus, although a "largest set" E containing A could contain either of $(-3,0)$ or $(3,4)$, it could not contain both and thus there

would be at least one subset of the plane containing A but not contained in E, determining only integral distances.



Above is the situation described in the previous paragraph: the set A consists of the three bolded points, with two points that individually determine integral distances to each point of A shown and distances between each pair of points labeled (except between $(-3,0)$ and $(3,0)$; this distance is simply 6).

However, in the case where a set A, $|A| > 1$ determining integral distances is entirely contained in a line, there is a completely natural (if somewhat trivial) maximal set E containing A contained in this line such that $\Delta(E) \subset \mathbb{Z}$ and if $A \subset F$, $\Delta(F) \subset \mathbb{Z}$, then $F \subset E$. This set is intuitively clear: in this case, A is a subset of some possibly translated, rotated copy of $\mathbb{Z}$.

More formally, choose $x, y \in A$ and let $|x - y| = d \in \mathbb{Z}$. Then, as a line is determined by two points, $A \subset \{y + t(x - y) : t \in \mathbb{R}\}$, and $\Big|y - \big[y + t(x - y)\big]\Big| = |t(x - y)| = td$, for $a \in \{y + t(x - y) : t \in \mathbb{R}\}$, $|y - a| \in \mathbb{Z}$ if and only if $ta \in \mathbb{Z}$, which means that $A \subset \big\{y + t(x - y) : t \in \frac{1}{a}\mathbb{Z}\big\}$ and, moreover, that any set containing A determining integral distances is a subset of this set. It is also easy to see that this set determines integral distances, as if $t, s \in \frac{1}{a}\mathbb{Z}$, then $|t(x - y) - s(x, y)| = |t - s| \cdot |x - y| = \frac{k}{a}a = k$ for some $k \in \mathbb{Z}$, so that it is this set that is the "maximal" set determining integral distances containing A.

Considering other additive subgroups of $\mathbb{R}$, if $\Delta(E) \subset \frac{1}{n}\mathbb{Z}$ for any $n \in \mathbb{N}$, then consideration of $nE = \{(nx_1, nx_2) : (x_1, x_2) \in E\}$ reveals that either nE is finite or nE is a subset of a line—but then E is finite or a subset of a line, itself! However, if we instead try to consider $\Delta(E) \subset \mathbb{Q}$, then there are sets E that are neither finite nor contained in a line satisfying this criterion, as in the following example.

Example 2.2: Consider the set $\{(0,0),(0,1)\}$ and extend it by adding $\big(\frac{p}{q}, 0\big)$ for p, q being the lengths of the two legs of a Pythagorean triple. Then, $\frac{p^2}{q^2} + 1$ is a square in $\mathbb{Q}$ and thus the distance from $(0,1)$ to $\big(\frac{p}{q}, 0\big)$ is a rational number. Also the distance between any other two points is trivially rational as all the other points are in $\mathbb{Q} \times \{0\}$ which determines rational distances. Thus, the set $E = \{(0,0),(0,1)\} \cup \big\{\frac{p}{q} : p^2 + q^2 = b^2 \text{ for some } b \in \mathbb{Z}\big\}$ is an infinite set not contained in any line because there are infinitely many primitive Pythagorean triples and as $(0,0),(0,1)$, and $(0, \frac{3}{4})$ are noncolinear there is no line containing all of E.

3. Dot Products in $\mathbb{R}^n$

Working along similar lines, what, if anything, can we say about the structure of A if $A \cdot A \subset \mathbb{Z}$, where $A \cdot A = \{a \cdot a' : a, a' \in A\}$? It is immediately clear that A need not be a subset of any line, as $\mathbb{Z}^n \subset \mathbb{R}^n$ determines only integral dot products and is not a subset of any line, plane, or affine subspace of $\mathbb{R}^n$. However, it still has a great deal of structure, as each point of it is a $\mathbb{Z}$-linear combination of the standard orthonormal basis vectors of $\mathbb{R}^n$, and we may thus consider $\mathbb{Z}^n$ as a lattice. Indeed, if A is the lattice generated by some orthonormal basis (over $\mathbb{R}$) for $\mathbb{R}^n$, it will be the case that $A \cdot A \subset \mathbb{Z}$. To make the above more precise, we define a lattice as follows:

Definition 3.1: Given a field F, a subring $\subset F$, and a finite-dimensional vector space F^n over F, an *R-lattice* is an R-module generated by a subset of F^n that is linearly independent over F. When no field or subring are specified, we usually mean a $\mathbb{Z}$-lattice, with $\mathbb{Z}$ considered as a subring of $\mathbb{R}$

Example 3.1: A few examples of lattices are provided here, for context.
The set $\mathbb{Z}^n \subset \mathbb{R}^n$ is a $\mathbb{Z} - lattice$, generated by the standard orthonormal basis of $\mathbb{R}^n$.
The set $\{0\}$ is the only $\{0\}$-lattice, because it is the only $\{0\}$-module.

Returning to the previous topic, a few immediate remarks may be made on the structure of a set A such that $A \cdot A \subset \mathbb{Z}$—firstly, for any $a \in A$, $|a|^2 \in \mathbb{Z}$, as $a \cdot a = |a|^2$, so any point of A is on a circle of radius $\sqrt{n}$ for some $n \in \mathbb{Z}$, centered at the origin. This is rather restrictive, especially compared to the setting of distances, where as $|a-a| = 0 \in \mathbb{Z}$, there are no *a priori* restrictions on what elements can be in a set determining only integral distances. However, we can view this as analogous to the fact that, once we know one of the elements of A with $\Delta(A) \subset \mathbb{Z}$, we know that all of the other points are on one of the circles of integral radius about this element. In some sense, this is because while distance is translation-invariant, dot products are automatically "centered" at the origin and the dot product is not translation invariant.

First, we shall show that any set $A \subset \mathbb{R}^n$ determining only integral dot products is a subset of some lattice, and then we shall demonstrate, with a few examples, how A may in fact have even more necessary structure than merely being an arbitrary subset of this lattice. In fact, as it is no harder to prove this in slightly greater generality, this is how we state it:

Lemma 3.1. *If F is a field with $R \subset F$ a subring, and $\cdot$ is an inner product[1] on F^n, and $A \subset F^n$ is such that $A \cdot A \subset R$, then A is a subset of an R-lattice.*

Proof. Suppose that the points of A span a d-dimensional subspace W of F^n and take d points of A in general position, $\{a_1, a_2, \ldots, a_d\}$.

Then, suppose that $b \in A$. Then $b \cdot a_i \in R$ for each i, so that we may naturally identify b with the d-tuple $(b \cdot a_1, b \cdot a_2, \ldots, b \cdot a_d) \in R^d$. Indeed, this identification is one-to-one, for elements of W: this is simply solving d independent linear equations on a d-dimensional space, which necessarily yields a unique solution.

To generate the lattice that A is contained in, then, take d vectors $b_i \in W$ such that $b_i \cdot a_j = \delta_{ij}$, the Kronecker delta. These are sure to exist by the above remarks, and we can then obtain any $a \in A$ as an R-linear combination of these vectors, for if $(a \cdot a_1, \ldots, a \cdot a_d) = (n_1, \ldots, n_d)$, then $a = \sum_{j=1}^{d} n_j b_j$, because as previously noted, a vector of W is uniquely defined by its dot products with each of the a_i, and for each i:

$$\Big(\sum_{j=1}^{d} n_j b_j\Big) \cdot a_i = \sum_{j=1}^{d} n_j (b_j \cdot a_j) = n_i$$

Thus, A is contained in the R-lattice generated by the set $\{b_1, b_2, \ldots, b_d\}$, which are linearly independent because if some F-linear combination of the b_i is 0, then $0 = c_1 b_1 + c_2 b_2 + \cdots + c_d b_d$, so that, $0 \cdot a_i = 0$ for all a_i but $(c_1 b_1 + c_2 b_2 + \cdots + c_d b_d) \cdot a_i = c_i$ for all a_i by construction of the b_i, so that $c_i = 0$ for all i. □

Remark: It is clear that the assumption that $\cdot$ is an inner product, rather than, say, a semi-inner product, is necessary, as otherwise we have $a \in F^n$ such that $a \cdot a = 0$, so that the set $A = \{0, a\}$ satisfies $A \cdot A \subset \{0\}$, but A is not contained in any $\{0\}$-module. We use the hypothesis that $\cdot$ is an inner product when we construct the b_i, because if $\cdot$ is not an inner product, it may be the case that for one of the a_i, $a_i \cdot b = 0$ for all $b \in F^n$.

The way to visualize the above proof is similar to the proof of the Erdős Integer Distance Principle, with affine hyperplanes taking the place of hyperbolae (as an affine hyperplane is the set of all points such that, say $a \cdot b = k$ for fixed $a \in F^n, k \in F$). However, it is worth noting that the lattice generated here is extremely dependent on the choice of the a_i, and that there is no guarantee that two points of this lattice will determine integer dot product with each other—nor even a guarantee that a point of this lattice will even have an integer dot product with itself.

Example 3.1: For example, if we let $A = \{(\sqrt{2}, 1), (-1, \sqrt{2})\} \subset \mathbb{R}^2$, any point with integral dot product with the points of A is in the lattice generated by $\left\{\left(\frac{\sqrt{2}}{3}, \frac{1}{3}\right), \left(\frac{-1}{3}, \frac{\sqrt{2}}{3}\right)\right\}$; in fact, this is simply the lattice constructed in the above proof. However, $\left(\frac{\sqrt{2}}{3}, \frac{1}{3}\right) \cdot \left(\frac{\sqrt{2}}{3}, \frac{1}{3}\right) = \frac{1}{3}$, so this lattice is, in some sense, too large.

[1]The important conditions here are that $v \cdot v \neq 0$ for any nonzero $v \in F^n$ and the linearity of the inner product in each component. Actual positive-definiteness is not necessary, and this "inner product" is implicitly assumed to map into F.

In this case, however, it is easy to see what the "correct" lattice is: it is simply the lattice generated by the set A. To see this, let $b_1 = \left(\frac{\sqrt{2}}{3}, \frac{1}{3}\right)$ and $b_2 = \left(\frac{-1}{3}, \frac{\sqrt{2}}{3}\right)$ and note that $b_1 \cdot b_1 = b_2 \cdot b_2 = \frac{1}{3}$ but $b_1 \cdot b_2 = 0$. Then, for $e = n_1 b_1 + n_2 b_2$:

$$e \cdot e = n_1^2(b_1 \cdot b_1) + 2n_1 n_2(b_1 \cdot b_2) + n_2^2(b_2 \cdot b_2) = \frac{1}{3}(n_1^2 + n_2^2)$$

so that if $e \cdot e \in \mathbb{Z}$, $3 \mid (n_1^2 + n_2^2)$, which by elementary number theory implies that $3 \mid n_1$ and $3 \mid n_2$. This implies that e is simply a $\mathbb{Z}$-linear combination of elements of A, as $A = \{3b_1, 3b_2\}$. Further, for any two c, d in the lattice generated by A, $c \cdot d$ will also be an integer.

Indeed, as in the above example, the lattice generated by A is always a lattice determining only integral dot products as $A \cdot A \subset \mathbb{Z}$ and it is easy to extend this to the lattice that A generates, because for $a_i, a_j \in A$, $n_i, m_j \in \mathbb{Z}$:

$$\left(\sum_{i=1}^{d_1} n_i a_i\right) \cdot \left(\sum_{j=1}^{d_2} m_j a_j\right) = \sum_{i=1}^{d_1}\sum_{j=1}^{d_2} n_i m_j (a_i \cdot a_j)$$

so that as each $a_i \cdot a_j \in \mathbb{Z}$, the sum is as well. However, this lattice is not always the largest lattice with this property: see the following example.

Example 3.2: Let $A = \{(1,2),(2,1)\}$. Then it is clear that the lattice generated by A is a subset of $\mathbb{Z}^2$, but this lattice is a proper subset of $\mathbb{Z}^2$ (as, for example, $(0,1)$ is not in the lattice). However, as is easy to check, the maximal lattice E containing A such that $E \cdot E \subset \mathbb{Z}$ is simply $E = \mathbb{Z}^2$.

It is tempting to assume that there is some "maximal lattice" corresponding to A (the rough counterpart of the lattice generated by A, which is the "minimal lattice"), i.e., a lattice E with $E \cdot E \subset \mathbb{Z}$ such that any lattice F determining only integral dot products and containing A is a subset of E. However, this is not immediately obvious, as, in general, there are lattices determining only integral dot products such any lattice containing them fails to do so.

Example 3.3: Consider $b_1 = (1,0)$ and $b_2 = (\frac{1}{4}, \frac{\sqrt{3}}{4})$. Then notice that the lattice determined by $\{b_1, 4b_2\}$ determines only integral dot products, because $b_1 \cdot 4b_2 = 1$ so the earlier remarks apply. Similarly, as $4b_1 \cdot b_2 = 1$, $\{4b_1, b_2\}$ also determines a lattice with only integral dot products. However, the smallest lattice containing both of these lattices is $b_1\mathbb{Z} + b_2\mathbb{Z}$, and $b_1 \cdot b_2 = \frac{1}{4}$.

However, the lattice here was not generated in the manner outlined in the lemma, so it is possible that any lattice so generated would have some "maximal" lattice determining integral dot products.

Thus, although the above lemma is simple, actual characterization of what these sets A can be is slightly more involved. In the setting of finite fields, however, matters turn out to be far simpler—in particular, we will often find that we do have "maximal lattices" of the form discussed earlier.

4. Distances and Dot Products over $\mathbb{F}_q$

Before beginning the discussion here, we must define what is meant by a distance in $\mathbb{F}_q^d$, a d-dimensional vector space over $\mathbb{F}_q$, the finite field with q elements (where $q = p^n$ for p prime, $n \in \mathbb{Z}_+$).

Definition 4.1: As elsewhere (e.g. [1] and [2]), we define the "distance" between $x = (x_1, x_2, \ldots, x_d)$ and $y = (y_1, y_2, \ldots, y_d)$ with $x, y \in \mathbb{F}_q^d$ to be $d(x,y) = \sum_{i=1}^d (x_i - y_i)^2 \in \mathbb{F}_q$, and write it as $\|x - y\|$. In a similar vein, we write $\|x\| = \sum_{i=1}^d x_i^2$

Note that this corresponds roughly to Euclidean distance: if we define a dot product between $x, y \in \mathbb{F}_q^d$ in a natural way, $x \cdot y = \sum_{i=1}^d x_i y_i$, then we find that $\|x\| = x \cdot x$ so that the main difference is the omission of a square root. However, considered as an analogue of distance, this operator is, as noted in [1], invariant when acted upon by orthogonal matrices, so it does share some useful properties with the Euclidean norm.

Then we define the "distance set" of E in the familiar way: $\Delta(E) = \{\|x - y\| : x, y \in E\}$. This is roughly analogous to the notion of Euclidean distance in $\mathbb{R}^n$, but without taking a square root as there is no way to choose a canonical square root in finite fields. One other fact keeping $\|\cdot\|$ from being a norm, and thus keeping $d(x,y)$ from being a "distance" in the normal way is that

$\|x\|$ can be 0 even for $x \neq 0$. The following construction demonstrating this can be found in [2] and the work of Alex Iosevich:

Example 4.1: If -1 is a square in $\mathbb{F}_q$, with one square root of it being i, then the set $\{(t, it) : t \in \mathbb{F}_q\} \subset \mathbb{F}_q^2$ determines a single distance, because for $s, t \in \mathbb{F}_q$, $(s, is) - (t, it) = (s - t, i(s - t))$ and then $(s - t)^2 + (i(s - t))^2 = 0$.

As this occurs whenever $q \equiv 1$ modulo 4, we can already see a marked contrast to the Euclidean case, where the only sets that determine a single-point distance set are singleton sets. Indeed, moving into higher dimensions, we can always construct a non-singleton set determining only the distance 0, because for any x_1, we can write $-x_1^2$ as the sum of two squares in $\mathbb{F}_q$ and thus can write 0 as the sum of three squares.

However, if we consider any (nonzero, though depending on our definition of subring this may be unnecessary) subring R of a finite field F—in fact, R will actually be a subfield of F—and suppose that $A \subset F^d$ is such that $\Delta(A) \subset R$, we will be able to say a great deal about the structure of A. For convenience and adherence to standard notation, we shall denote R by $\mathbb{F}_q$ (as R will indeed be a finite field), and we may refer to F as $\mathbb{F}_{q^n}$, as these are the only finite fields containing $\mathbb{F}_q$ so as $\mathbb{F}_q$ is a subfield of F, F is $\mathbb{F}_{q^n}$ for some n.

In this case, we have a result that is analogous to the Erdős Integer Distance Principle, in that it gives a bound on the size of a set $A \subset \mathbb{F}_{q^n}^d$ determining $\mathbb{F}_q$ distances such that A is not a subset of an affine hyperplane.

Lemma 4.1. *Let $\mathbb{F}_q$ be a finite field of characteristic $p \neq 2$. If $A \subset \mathbb{F}_{q^n}^d$ and $\Delta(A) \subset \mathbb{F}_q$, then one of the following holds:*
(1) *A is contained in an affine hyperplane over $\mathbb{F}_{q^n}^d$*
(2) *A is a subset of an affine $\mathbb{F}_q$–lattice in $\mathbb{F}_{q^n}^d$*
Thus, combining these bounds, $|A| \leq \max\left\{\left(q^n\right)^{d-1}, q^d\right\}$.

Proof. In the proof of the Erdős Integer Distance Principle, we began by choosing 3 noncolinear points of A. Then, unless A is contained in some affine hyperplane, we may simply choose $d + 1$ points of A to be in general position, $\{a_0, a_1, \ldots, a_d\}$. Then, for any other point $b \in A$, $\|b - a_i\| \in \mathbb{F}_q$ for all the a_i, so that in particular, for $1 \leq i \leq m$, $\|b - a_0\| - \|b - a_i\| = k_i \in \mathbb{F}_q$, and as

$$\begin{aligned}\|b - a_0\| - \|b - a_i\| &= \big((b - a_0) \cdot (b - a_0)\big) - \big((b - a_i) \cdot (b - a_i)\big) \\ &= \|b\| - 2b \cdot a_0 + \|a_0\| - \|b\| + 2b \cdot a_i - \|a_i\| \\ &= \|a_0\| - \|a_i\| + 2b \cdot (a_i - a_0)\end{aligned}$$

and then we notice that, since we have only q choices for k_i, b must be on one of q shapes defined by the equation:

$$b \cdot (a_i - a_0) = \frac{1}{2}\big(k_i + \|a_i\| - \|a_0\|\big)$$

for some $k_i \in \mathbb{F}_q$, which we recognize as the equation of a hyperplane in $\mathbb{F}_{q^n}^d$. Then, since the vectors $(a_i - a_0)$ are linearly independent (by construction of the set $\{a_0, a_1, \ldots, a_d\}$) implies that if we specify $\{k_i\}$ for $1 \leq i \leq d$, we are solving d independent linear equations, which must have a unique solution in our d-dimensional vector space $\mathbb{F}_{q^n}^d$. Then, as there are exactly q choices for each of the d variables k_i, $|A| \leq q^d$, and in fact this provides a way (as long as A is not contained in an affine hyperplane) to construct a set E containing A such that any set containing A (or, indeed, any set merely containing $\{a_0, a_1, \ldots, a_d\}$) and determining only $\mathbb{F}_q$-distances is contained in E. By the previous remarks, $|E| = q^m$.

Continuing to suppose that A is not a subset of an affine hyperplane, we wish to see that A is contained in an affine $\mathbb{F}_q$-lattice. To prove this, we construct such a set L containing $\{a_0, a_1, \ldots, a_d\}$ of size q^d, such that $\Delta(L) \subset \mathbb{F}_q$. Then by the above remarks, $L \subset E$ but then because $|L| = |E| < \infty$, $L = E$.

Consider the set $\{a_1 - a_0, a_2 - a_0, \ldots, a_d - a_0\}$, and for convenience let us set $b_i = a_i - a_0$. Notice firstly that $\|b_i\| \in \mathbb{F}_q$ for all i, by hypothesis and definition of the distance set. Then, as $\|a_i - a_j\| = \|b_i - b_j\|$ and $\|a_i - a_j\| \in \mathbb{F}_q$ by hypothesis, $\|b_i - b_j\| \in \mathbb{F}_q$ and

$$\|b_i - b_j\| = \|b_i\| - 2b_i \cdot b_j + \|b_j\|$$

so that since $\|b_i - b_j\|, \|b_i\|, \|b_j\|$, and $\frac{-1}{2}$ are elements of $\mathbb{F}_q$ (as $\mathbb{F}_q$ is not of characteristic 2), $b_i \cdot b_j$ is an element of $\mathbb{F}_q$ as well.

Now consider the $\mathbb{F}_q$ lattice L' generated by $\{b_1, \ldots, b_d\}$. We want $\Delta(L') \subset \mathbb{F}_q$, so let us take $x, y \in L'$ and write these as $x = x_1b_1 + x_2b_2 + \cdots + x_db_d$ and $y = y_1b_1 + y_2b_2 + \cdots + y_db_d$ where the $x_i, y_i \in \mathbb{F}_q$. Then let us calculate $\|x - y\|$:

$$\|x-y\| = (x-y)\cdot(x-y) = \Big(\sum_{i=1}^{d}(x_i-y_i)b_i\Big)\cdot\Big(\sum_{j=i}^{d}(x_j-y_j)b_j\Big)$$

$$= \sum_{i=1}^{d}\sum_{j=1}^{d}(x_i-y_i)(x_j-y_j)(b_i\cdot b_j)$$

but then simple term-by-term inspection reveals this to be an element of $\mathbb{F}_q$, as for all i, j, x_i, y_i, x_j, y_j, and $b_i \cdot b_j$ are elements of $\mathbb{F}_q$, and a finite sum of products of elements of $\mathbb{F}_q$ is again an element of $\mathbb{F}_q$.

Thus, L' is, by construction, a lattice containing $\{0, b_1, \ldots, b_d\}$ such that $\Delta(L') \subset \mathbb{F}_q$. If we let $L = L' + a_0$, L is an affine lattice containing $\{a_0, a_1, \ldots, a_d\}$ such that $\Delta(L) \subset \mathbb{F}_q$, which is precisely what we set out to construct. Therefore, if A is not a subset of an affine hyperplane, A is a subset of an affine lattice. □

Although this does give some bounds on the size of A, one of which can always be achieved—namely, q^d, simply by taking the natural embedding of $\mathbb{F}_q^d \subset \mathbb{F}_{q^n}^d$—the other bound is usually the larger of the two and can be improved. In particular, if either d or n is greater than 2, then the larger of the two bounds is $\left(q^n\right)^{d-1}$, and this bound is too high. However, in the case where A is a subset of a lattice, we can easily construct a "maximal lattice" containing A determining only $\mathbb{F}_q$-distances, which is more than we did in the Euclidean case of dot products!

It is tempting to, by analogy with the dot product case earlier, to think that we can do better by considering the affine subspace W that A spans, which has, say, m dimensions, taking $m + 1$ points in general position in W and then deriving m linear equations from these vectors. However, though we can do all of this, we cannot then conclude that these m equations actually give us any more information about the points of A because it is entirely possible that one or more—even possibly all!—of the equations are satisfied for every point of W. To see this, consider the following example, which should be considered an extension of Example 4.1:

Example 4.2: Let $\mathbb{F}_q$ be a field of characteristic $p \neq 2$ such that $\mathbb{F}_q$ contains a square root of -1, and call one of these square roots i. Then $\mathbb{F}_q^{2d}$ contains sets determining only the distance $\{0\}$ that are not lattices over any proper subfield of $\mathbb{F}_q$. One such set $A = \{(s_1, is_1, s_2, is_2, \ldots, s_d, is_d) : s_i \in \mathbb{F}_q\}$, which is in fact a subspace of F_q^{2d}. Seeing that any two points of A determine a distance of 0 is a simple computation analogous to that in Example 4.1.

This is not a lattice over any proper subfield F of $\mathbb{F}_q$, as the linear subspace $E = \{(s, is, 0, 0, \ldots, 0, 0) : s \in \mathbb{F}_q\}$ has dimension strictly greater than 1 over F, so any basis for A over F will need more than one vector to span E and this basis will thus be linearly dependent over $\mathbb{F}_q$ as E is one-dimensional over $\mathbb{F}_q$.

Now, we can quickly verify that $A \cdot A = 0$ as well: this is easily computed directly. But this means that if we tried to construct linear equations as outlined above, inspection reveals that the affine hyperplanes defined by these equations either contain A or are disjoint from A—they contain A when they are actually hyperplanes (i.e. subspaces) and do not contain A when they are translated nontrivially.

It is worth pausing here to note why this was not a problem in the earlier Euclidean dot product case—this is because in the Euclidean case, if $0 \neq x_1 \in A$, it is impossible for A to be contained in the hyperplane orthogonal to x_1 because $x_1 \cdot x_1 \neq 0$, because $\cdot$ is assumed to be an inner product. This, in turn, implies that this orthogonal hyperplane intersects A in a subspace of A of codimension 1, and we can iterate this process to work through the all details of Lemma 3.1. Moreover, we can use the above ideas to find an upper bound on the size of subspaces that are similar to the one constructed in Example 4.2:

Lemma 4.2. *If W is a subspace of $\mathbb{F}_q^d$ such that $W \cdot W = \{0\}$, then the dimension of W is at most $\frac{d}{2}$ over $\mathbb{F}_q$.*

Proof. Suppose that W, a subspace of $\mathbb{F}_q^d$ has dimension m over $\mathbb{F}_q$, and $W \cdot W = \{0\}$.

Defining $W^\perp = \{x \in \mathbb{F}_q^d : \forall w \in W : x \cdot w = 0\}$, we can easily see that $W^\perp$ has dimension $d - m$ by taking a basis $\{w_1, \ldots, w_m\}$ for W and considering the linear transformation $T : \mathbb{F}_q^d \to \mathbb{F}_q^m$

where $x \mapsto (x \cdot w_1, x \cdot w_2, \ldots, x \cdot w_m)$. Because the w_i are linearly independent, T has rank m so that $\dim(\mathrm{Im}(T)) = m$.

Then $T(\mathbb{F}_q^d)$ has dimension m, $\mathbb{F}_q^d$ has dimension d, so by the rank-nullity theorem, $\mathrm{Ker}(T) = W^\perp$ has dimension $d-m$. But then, by assumption, $W \subset W^\perp$, so $m \leq d-m$ and therefore $m \leq \frac{d}{2}$. □

This lets us immediately bound the size of a set determining the distance $\{0\}$, again exploiting the close relationship between dot products and distances over finite fields:

Corollary 4.2.1. *If $A \subset \mathbb{F}_q^d$ is such that $\Delta(A) = \{0\}$, then A is a subset of an affine subspace of $\mathbb{F}_q^d$ of dimension at most $\frac{d}{2}$ over $\mathbb{F}_q$.*

Proof. Let $a_0 \in A$, and let W be the linear span of $A - a_0 = \{a - a_0 : a \in A\}$. By translation invariance, $\Delta(A - a_0) = \{0\}$, and moreover for each $x \in A - a_0$, $\|x\| = 0$. Then, as $\|x - y\| = \|x\| - 2x \cdot y + \|y\|$, for any $b, \prime \in A - a_0$, $b \cdot b' = \{0\}$, so that, finally, if $v, w \in W$, since both v and w are linear combinations of elements of $A - a_0$:

$$v \cdot w = \Big(\sum_{a_i \in A - a_0} \alpha_i a_i \Big) \cdot \Big(\sum_{a_j \in A - a_0} \beta_j a_j \Big) = \sum_{a_i, a_j \in A - a_0} \alpha_i \beta_j (a_i \cdot a_j) = 0$$

Thus, $W \cdot W = \{0\}$ so that Lemma 4.2 applies. Therefore $A - a_0$ is a subset of a subspace of $\mathbb{F}_q^d$ of dimension at most $\frac{d}{2}$ and thus A is a subset of an affine subspace of $\mathbb{F}_q^d$ of dimension at most $\frac{d}{2}$. □

Further, Example 4.2 shows that this bound is sharp. The next theorem characterizes the sets handled in Lemma 4.1 more precisely, thereby allowing a better bound on the size of such sets. Informally, it shows that we can consider the "lattice" structure and the "nonlattice" structure separately, with the nonlattice component corresponding to subspaces determining a distance of $\{0\}$

Theorem 4.3. *Let $\mathbb{F}_q$ be a finite field of characteristic $p \neq 2$. If $A \subset \mathbb{F}_{q^n}^d$ and $\Delta(A) \subset \mathbb{F}_q$, then A is contained in a translate of a direct sum of an $\mathbb{F}_q$-lattice Λ and an $\mathbb{F}_{q^n}$ subspace N of $\mathbb{F}_{q^n}^d$ such that $\Delta(N) = \{0\}$, and where the dimension of Λ over $\mathbb{F}_q$ plus the dimension of N over $\mathbb{F}_{q^n}$ is less than d.*

Proof. Begin by translating A to include 0, and work with the resultant set B, noting that $0 \in B$ and that $\Delta(B) \subset \mathbb{F}_q$. Let the linear span of B over $\mathbb{F}_{q^n}$ be W, and let $N = \{w \in W : \forall v \in W : w \cdot v = 0\} = W \cap W^\perp$. N is a subspace of W and $\Delta(N) = \{0\}$ as for $v, w \in N$, $\|v - w\| = \|v\| - 2v \cdot w + \|w\| = 0$ because $v, w \in W$ as well as being in N.

N has certain useful properties due to its definition. Firstly, for any $x, y \in W$, $(x + N) \cdot (y + N)$ is well-defined: for all $v, w \in N$:

$$(x + v) \cdot (y + w) = x \cdot y + x \cdot w + v \cdot y + v \cdot w = x \cdot y$$

and due to this, $\|(x + v)\| = \|x\|$ for all $x \in W$. Thus, we may consider W/N to inherit these operators from W, and speak of e.g. $\|(x+N)\|$ in a sensible manner. Also, if $(x+N)$ is orthogonal to W/N in the sense that $(x + N) \cdot (y + N) = 0$ for all y, then $x + N = N$.

Then, if we pick $\{x_1 + N, \ldots, x_k + N\}$ to be a basis for W/N with $x_i \in B$, because x_i is not orthogonal to W, the subspace $\mathrm{Ann}(x_i) = \{w \in W : w \cdot x_1 = 0\}$ has codimension 1 and contains N, so $\mathrm{Ann}(x_i)/N$ is a subspace of W/N that also has codimension 1. Also, the intersection of all of these subspaces is easily seen to be trivial because if $(w + N) \cdot (x_i + N) = 0$ for all i, then $(w+N) \cdot (W/N) = 0$. This, in turn, means that if we have a system of linear equations of the form $(b + N) \cdot (x_i + N) = c_i$ with the c_i being fixed constants, this has a unique solution in W/N.

Now, let b be an arbitrary point of B. Then $\|b+N\| \in \mathbb{F}_q$ and moreover $\|(b+N)-(x_i+N)\| \in \mathbb{F}_q$ for all i. Then, in precisely the same manner as in Lemma 4.1, we note that $\|b + N\| - \|(b + N) - (x_i + N)\| \in \mathbb{F}_q$ and:

$$\begin{aligned} \|b + N\| - \|(b + N) - (x_i + N)\| &= \|b + N\| - \|b + N\| + 2(b + N) \cdot (x_i + N) - \|x_i + N\| \\ &= 2(b + N) \cdot (x_i + N) - \|x_i + N\| \end{aligned}$$

thus, if $b + N$ is a point such that this difference of distances is s_i, then $b + N$ is a solution to the equation:

$$(b + N) \cdot (x_i + N) = \frac{1}{2}\big(s_i + \|x_i + N\|\big)$$

Then, as we have q choices for each s_i (as $\|(b + N) - (x_i + N)\| \in \mathbb{F}_q$), and k indices i, and each choice of a vector $(s_1, s_2, \ldots, s_k)$ corresponds to a unique point in W/N, we have a set E of size

at most q^k in W/N such that each point of E determines integral distances to each of the vectors $\{0, x_1, x_2, \ldots, x_k\}$.

Then, if we once again consider the natural $\mathbb{F}_q$ lattice L generated by $\{x_1, x_2, \ldots, x_k\}$, in precisely the same manner as in Lemma 4.1 we will find that $\Delta(L) \subset \mathbb{F}_q$ because $x_i \cdot x_j \in \mathbb{F}_q$ for all i, j. Then it is clear that since $\{0, x_1, x_2, \ldots, x_k\} \subset L$ and $\Delta(L) \subset \mathbb{F}_q$, $L \subset E$, so that $L = E$ because they have the same size.

Now, for each point of B, $b + N \in L$ so that if φ is the canonical mapping $\varphi : W \to W/N$, it must be the case that $B \subset \varphi^{-1}(L)$, which in turn can be considered as the direct sum of the $\mathbb{F}_q$-lattice Λ generated by $\{x_1, \ldots, x_k\}$ and the subspace N.

If we wish to check that this is, in fact, a direct sum, it suffices to notice that for $a, b \in \Lambda$, $v, w \in N$, if $a + v = b + w$, then $a - b = w - v$ so that $a - b \in N$. But then since $x_i + N \neq N$ for each x_i, this implies that $a - b = 0$ as a, b are $\mathbb{F}_q$-linear combinations of the x_i. Thus, there are no relations between the different "coordinates" here and thus we may consider these as ordered pairs.

Further, it is clear that in fact the entire set $\Lambda \oplus N$ satisfies $\Delta(\Lambda \oplus N) \subset \mathbb{F}_q$, as if $(a, v), (b, w) \in \Lambda \oplus N$, then:

$$\|(a, v) - (b, w)\| = \|a + v - b - w\| = \|a + N - b + N\|$$

which is in $\mathbb{F}_q$ because the $\mathbb{F}_q$-lattice $L \subset W/N$ determines only $\mathbb{F}_q$ distances.

Finally, the bound on the dimension of Λ over $\mathbb{F}_q$ plus the dimension of N over $\mathbb{F}_{q^n}$ is simply from the fact that $W \leq \mathbb{F}_{q^n}^d$ and as the x_i generating Λ span W/N over $\mathbb{F}_{q^n}$, the statement in the theorem is simply a rephrasing of the fact that $\dim(N) + \dim(W/N) = \dim(W) \leq d$ □

This provides, for any set A such that $\Delta(A) \subset \mathbb{F}_q$, a "maximal structure" M containing A such that $\Delta(M) \subset \mathbb{F}_q$ as well. Thus, in some sense, we have achieved the goal that we set out to accomplish for sets of this type—given a set A with distance set contained in $\mathbb{F}_q$, we have shown that there is a natural maximal set containing A having a fairly nice algebraic structure that maintains the property that its distance set is a subset of $\mathbb{F}_q$.

Corollary 4.3.1. *If $\mathbb{F}_q$ is of characteristic $p \neq 2$ and $A \subset \mathbb{F}_{q^n}^d$ such that $\Delta(A) \subset \mathbb{F}_q$, where $n > 1$, then $|A| \leq \left(q^n\right)^{\frac{d}{2}}$*

Proof. First, we notice that in the case where $n = 1$ we obviously cannot say anything nontrivial about the size of A: $\Delta(\mathbb{F}_q^d) = \mathbb{F}_q$, so we understand the case excluded by the assumption that $n > 1$.

Then, combining the above methods with the ideas used in Lemma 4.2, suppose that the linear span of A over $\mathbb{F}_{q^n}$ is W, where $\dim(W) = m$. Then, by Lemma 4.1, A is a subset of $\Lambda \oplus N$ where Λ is an $\mathbb{F}_q$-lattice and $N = W \cap W^{\perp}$. Thus, $\dim(N) \leq \min\{m, d - m\}$ and the dimension of Λ over $\mathbb{F}_q$ is $m - \dim(N)$.

We can then see that $|\Lambda \oplus N| = (q^n)^{\dim(N)} q^{m - \dim(N)}$ For a fixed m, this size is maximized when $\dim(N)$ is as large as it can be, because $q^n > q$.

For $m \leq \frac{d}{2}$, the upper bound on $\dim(N)$ is m, so $|\Lambda \oplus N| \leq (q^n)^m$ in this case. For $m \geq \frac{d}{2}$, the upper bound on $\dim(N)$ is $d - m$, so $|\Lambda \oplus N| \leq (q^n)^{d-m} q^{2m-d}$ in this case, and simple inspection of these two bounds shows that the first increases in m and the second is nonincreasing in m because $n \geq 2$.

Therefore, the bound on the size of $|\Lambda \oplus N|$ from Lemma 4.1 reaches a maximum when $m = \frac{d}{2}$, and this maximum size is $(q^n)^{\frac{d}{2}}$. □

Further, we can see that this upper bound cannot be improved in general due to Example 4.2, which exhibits a set of this size determining the distance $\{0\}$.

For the duration of this section, we have been neglecting the case where $\mathbb{F}_q$ is of characteristic 2. The following discussion will handle this case.

Suppose that $\mathbb{F}_q$ is a finite field of characteristic 2 and $A \subset \mathbb{F}_{q^n}^d$ has $\Delta(A) \subset \mathbb{F}_q$. Then say $x = (x_1, \ldots, x_2) \in A$, so that for any other $y = (y_1, \ldots, y_d) \in A$, $\|x - y\| \in \mathbb{F}_q$. Now, exploiting the fact that $\mathbb{F}_q$ is of characteristic 2 and using Freshman's Dream:

$$\|x - y\| = \|x\| - 2x \cdot y + \|y\| = \|x\| + \|y\| = \sum_{i=1}^{d} x_i^2 + \sum_{j=1}^{d} y_j^2$$

$$= \Big(\sum_{i=1}^{d} x_i + \sum_{j=1}^{d} y_j \Big)^2$$

and then, since for $m \in \mathbb{F}_{q^n}$, $m^2 \in \mathbb{F}_q$ if and only if $m \in \mathbb{F}_q$, so that $\|x-y\| \in \mathbb{F}_q$ if and only if $\big(\sum_{i=1}^{d} x_i + \sum_{j=1}^{d} y_j \big) \in \mathbb{F}_q$ as well.

Thus, once $x \in A$ is fixed, we may choose $y_1, \ldots, y_{d-1}$ with complete freedom and then simply choose $y_d \in \big(\mathbb{F}_q - \sum_{i=1}^{d} x_i - \sum_{j=1}^{d-1} y_j \big)$, so that we have q^n choices for each of $y_1, \ldots, y_{d-1}$ and q choices for y_d. Then, if we have a y, z such that $\|x-y\| \in \mathbb{F}_q$ and $\|x-z\| \in \mathbb{F}_q$, we will also have that $\|y-z\| \in \mathbb{F}_q$

This is simply because $\|x-y\|$ and $\|x-z\|$ are in $\mathbb{F}_q$ if and only if both of $\big(\sum_{i=1}^{d} x_i + \sum_{j=1}^{d} y_j \big) \in \mathbb{F}_q$ and $\big(\sum_{i=1}^{d} x_i + \sum_{j=1}^{d} z_j \big) \in \mathbb{F}_q$, so their sum is in $\mathbb{F}_q$ as well. But this sum is simply $\big(\sum_{i=1}^{d} y_i + \sum_{j=1}^{d} z_j \big)$, and by our earlier observations, this is in $\mathbb{F}_q$ if and only if $\|y-z\| \in \mathbb{F}_q$

This construction thus produces a subset A of $\mathbb{F}_{q^n}^d$ of size $(q^n)^{d-1}q$ which can, in general, be much larger than $(q^n)^{\frac{d}{2}}$. We summarize these findings in the following lemma:

Lemma 4.4. *If $\mathbb{F}_q$ is of characteristic 2 and $A \subset \mathbb{F}_{q^n}^d$ has $\Delta(A) \subset \mathbb{F}_q$, then there exists $E \subset \mathbb{F}_{q^n}^d$ such that $A \subset E$, $\Delta(E) \subset \mathbb{F}_q$ and $|E| = (q^n)^{d-1}q$.*

Proof. Selecting $x \in A$ arbitrarily and then taking $E = \{y \in \mathbb{F}_{q^n}^d : \|x-y\| \in \mathbb{F}_q\}$ will produce a set E with the desired properties by the above work. □

Thus, the case where $\mathbb{F}_q$ is of characteristic 2 is quite different from the odd characteristic case, and we can have subsets of $\mathbb{F}_{q^n}^d$ in characteristic 2 that determine $\mathbb{F}_q$ distances and have very little structure, whereas in odd characteristic knowing that $\Delta(A)$ is contained in a proper subfield of $\mathbb{F}_{q^n}$ lets us learn a great deal about the structure of A.

Finally, it is worth noting that, instead of using $\|x\| = \sum_{i=1}^{d} x_i^2$, we can work with a nondegenerate quadratic form $\|x\|_a = \sum_{i=1}^{d} a_i x_i^2$ with $a_i \neq 0$, with the analogous dot product being $x \cdot_a y = \sum_{i=1}^{d} a_i x_i y_i$. In this terminology, it is easy to show that $\|x-y\|_a = \|x\|_a - 2(x \cdot_a y) + \|y\|_a$. Further, we may analogously define $\Delta_a(E) = \{\|x-y\|_a : x, y \in E\}$

Then, it is easily seen that Lemma 4.1 still applies if $\Delta_a(E) \subset \mathbb{F}_q$ instead of the standard distance set, although instead of solving independent equations of the form $v \cdot (x_j - x_0) = k_j$, we will be solving equations of the form $\sum_{i=1}^{d} a_i v_i (x_j - x_0)_i = k_j$. Still, if we view the first case as solving the vector equation $Xv = k$ where X has $(x_j - x_0)$ as its j^{th} row, we are instead solving $AXv = k$, where A is a diagonal matrix with entries a_j along this diagonal, which is invertible since $a_j \neq 0$ for each j. Then since X is invertible (as the $x_j - x_0$ are assumed to be linearly independent), we can simply solve this system as usual and the bound in Lemma 4.1 holds.

It is not much harder to work through the details of Theorem 4.2 with a general nondegenerate quadratic form. Given $W \leq \mathbb{F}_{q^n}^d$, a general subspace, we can define $W^{\perp_a} = \{x \in \mathbb{F}_{q^n}^d : \forall w \in W : x \cdot_a w = 0\}$ and this will still be a subspace (in fact, it will still be a subspace of dimension $d-m$ if W has dimension m by the same considerations), so we may again take $N_a = W \cap W^{\perp_a}$ and work, as before, in W/N_a, where all of the previous methods will still work.

Then, an interesting question is what structure a set A has if $\Delta_a(E) \subset \mathbb{F}_q$ for all $a \in \big(\mathbb{F}_q \setminus \{0\} \big)^d$. Such sets clearly exist—$E = \mathbb{F}_q^d$ is one such set—but they will necessarily have much more structure than sets that merely have $\Delta(E) \subset \mathbb{F}_q$. For example, if E is such a set and $x, y \in E$, and $\mathbb{F}_q$ is not of characteristic 2, then $(x_i - y_i)^2 \in \mathbb{F}_q$ for all i, because if this is not the case and say $(x_{i_0} - y_{i_0})^2 \notin \mathbb{F}_q$, then it must also be the case that $\|x-y\| + (x_{i_0} - y_{i_0})^2 \notin \mathbb{F}_q$. But then if we consider the nondegenerate quadratic form corresponding to $a = (1, 1, \ldots, 2, \ldots, 1)$ with a 2 in the i_0 index, $\|x-y\|_a \notin \mathbb{F}_q$ by the earlier remarks.

Further, for any vector $v \neq 0 \in \mathbb{F}_{q^n}^d$, $\mathbb{F}_q$ not of characteristic 2, there is some nondegenerate quadratic form $\|\cdot\|_a$ such that $\|v\|_a \neq 0$. To see this, suppose that $\|v\| = 0$, with a nonzero entry at the i_0 index (which much exist since $v \neq 0$). Then let $a = (1, 1, \ldots, 2, \ldots, 1)$ with a 2 in the i_0 index, and notice that then $\|v\|_a = \|v\| + v_{i_0}^2 \neq 0$. On the other hand, if $\|v\| \neq 0$, then we are already done. This lets us prove a stronger version of Theorem 4.3 under stronger hypotheses.

Theorem 4.5. *If $\mathbb{F}_q$ is a finite field of characteristic $p \neq 2$, $E \subset \mathbb{F}_{q^n}^d$ is such that $\Delta_a(E) \subset \mathbb{F}_q$ for every $a \in \big(\mathbb{F}_q \setminus \{0\} \big)^d$, then E is a subset of an affine $\mathbb{F}_q$-lattice.*

Proof. We proceed by induction on the dimension of the affine subspace that E spans. If this dimension is 0, E is a singleton set and trivially a subset of an affine $\mathbb{F}_q$-lattice.

Suppose that for all affine subspaces $W + v$ of dimension strictly less than m, if we have a set $E' \subset W+v$ and $\Delta(E) \subset \mathbb{F}_q$, then E' is a subset of an affine $\mathbb{F}_q$-lattice. Notice that this immediately implies that $|E| \leq q^{m-1}$

Now, let E span an affine subspace of dimension m and let $\{e_0, e_1, \ldots, e_m\}$ be in general position in E. For convenience, let W be the span of $\{e_1 - e_0, \ldots, e_m - e_0\}$. Choose a nondegenerate quadratic form $\|\cdot\|_a$ such that $\|e_m - e_0\|_a \neq 0$. Then let $x \in E$, and notice that $\|x-e_0\|_a - \|x-e_m\|_a$ can take only values in $\mathbb{F}_q$, so that any point x of E is a solution of the following equation for some $k \in \mathbb{F}_q$:

$$x \cdot_a (e_m - e_0) = \frac{1}{2}\big(k + \|e_m\|_a - \|e_0\|_a\big)$$

Then, since $\|e_m - e_0\|_a \neq 0$, the hyperplane that is a-orthogonal to $e_m - e_0$ does not contain all of W and thus intersects W in a subspace of dimension $m-1$. Solving q different equations of the above form yields q possible affine subspaces of dimension $m-1$ that can contain points of E, but by the inductive hypothesis, each of these subspaces contains at most q^{m-1} points.

Therefore $|E| \leq q^m$ and there is a set M such that $|M| = q^m$ with the property that if S is such that $\{e_0, e_1, \ldots, e_m\} \subset S$ and $\|s - e_i\|_a \in \mathbb{F}_q$ for all $s \in S, e_i \in \{e_0, \ldots, e_m\}$ and for all $a \in \big(\mathbb{F}_q \setminus \{0\}\big)^d$, then $S \subset M$—this set is implicitly constructed above in the induction argument, but notice that *a priori* there is no guarantee that these conditions hold for arbitrary subsets of M, as very few of these hypotheses are actually used in the construction of M. However, if a set satisfies these rather strong hypotheses, then we know that it must be a subset of M.

However, as in the previous examples, if we set $v_i = e_i - e_0$, then $\mathbb{F}_q$-lattice L generated by $\{v_1, \ldots, v_m\}$ actually satisfies $\Delta_a(L) \subset \mathbb{F}_q$ for all suitable a and obviously has size q^m so that, as in the other cases, $M = L + e_0$. To see that this condition holds, notice that $\|v_i\|_a \in \mathbb{F}_q$ and $v_i \cdot_a v_j \in \mathbb{F}_q$ for all i, j, and thus letting $x = \sum_{i=1}^m \alpha_i v_i$, $y = \sum_{i=1}^m \beta_i v_i$, with $\alpha_i, \beta_i \in \mathbb{F}_q$, then:

$$\|x - y\|_a = (x - y) \cdot_a (x - y) = \Big(\sum_{i=1}^m (\alpha_i - \beta_i)v_i\Big) \cdot_a \Big(\sum_{j=1}^m (\alpha_j - \beta_j)v_j\Big)$$

$$= \sum_{i=1}^m \sum_{j=1}^m (\alpha_i - \beta_i)(\alpha_j - \beta_j)(v_i \cdot_a v_j)$$

which is in $\mathbb{F}_q$ by inspection since each term being summed is an element of $\mathbb{F}_q$. Then by translating L by v_0 so that this translate of L contains $\{e_0, e_1, \ldots, e_m\}$, we establish the desired result. □

Remark: In the above proof, the only point that we need the whole family of nondegenerate quadratic forms for is finding a quadratic form such that $\|v\|_a \neq 0$, but as we have already seen we can find such a quadratic form using only $a \in \{1, 2\}^d$, and, in fact, we can use only the a such that $a_i = 1$ for all but possibly 1 index and $a_i = 2$ in one or zero indices. Thus, if we change the hypothesis "for every $a \in \big(\mathbb{F}_q \setminus \{0\}\big)^d$" to "for every $a \in \{1, 2\}^d$ with $a_i = 2$ for at most one i," the conclusion of the theorem does not change but we need only consider these $m+1$ quadratic forms. In fact, because any v such that $\|v\| = 0$ will necessarily have at least 2 nonzero components, we can exclude any one of the a of the form $(1, 1, \ldots, 2, \ldots, 1)$. However, excluding any more of the a in this collection allows some analogue of Example 4.2 to still work as a counterexample, simply by taking the vectors of the form $sw = (0, 0, \ldots, s, \ldots, is, \ldots, 0)$ where the s and is correspond to the two indices that are always 1 for the a in our chosen collection.

Then, for any quadratic form corresponding to an a in our collection:

$$\|sw\|_a = \sum_{i=1}^d a_i (sw)_i^2 = 0 + 0 + \cdots + s^2 + \cdots + (is)^2 + \cdots + 0 = s^2 - s^2 = 0$$

so that sw is self-orthogonal under any of these quadratic forms.

The above theorem uses stronger hypotheses than Theorem 4.3 to find a stronger result, and heuristically tells us that if A behaves enough like $\mathbb{F}_q^d$ with respect to (sufficiently large sets of) quadratic forms, then A must in fact look a great deal like $\mathbb{F}_q^d$ in terms of its actual algebraic structure. It is natural to then wonder whether there are any sets that are not $\mathbb{F}_q^d$ that have this property.

One last example will demonstrate the existence of a set E that is not a translate of $\mathbb{F}_q^d$ and which still satisfies $\Delta_a(E) \subset \mathbb{F}_q$ for all $a \in \left(\mathbb{F}_q \setminus \{0\}\right)^d$:

Example 4.3: Let E be the $\mathbb{F}_3$-lattice in $\mathbb{F}_9$ generated by $(i,0),(0,1)$ where i is a square root of 2 in $\mathbb{F}_9$ (there is such an element by the theory of field extensions). Then E is not simply a translate of $\mathbb{F}_3^2$ because then $(i,0) \in E$ would imply $(i+1,0) \in E$ but we can simply write down the 9 elements of E as follows: $E = \{(0,0),(i,0),(2i,0),(0,1),(i,1),(2i,1),(0,2),(i,2),(2i,2)\}$ and then as $i+1$ is not equal to any of $0, i, 2i$, we can see that $(i+1,0) \notin E$ so that E is not a translate of $\mathbb{F}_3^2$.

However, $\Delta_a(E) \subset \mathbb{F}_3$ for any $a \in \left(\mathbb{F}_3 \setminus \{0\}\right)^2$ because for $x, y \in E$, $x - y = (\alpha i, \beta)$ for $\alpha, \beta \in \mathbb{F}_3$ due to the construction of E, so that:

$$\|x-y\|_a = a_1(\alpha i)^2 + a_2(\beta)^2 = -\alpha^2 + \beta^2 \in \mathbb{F}_3$$

Now, recall our earlier remarks, which show that if $\|x-y\|_a \in \mathbb{F}_q$ for all $a \in \left(\mathbb{F}_q \setminus \{0\}\right)^d$, then $(x_i - y_i)^2 \in \mathbb{F}_q$ for all i. This means that, in some sense, any set E with $\Delta_a(E) \subset \mathbb{F}_q$ for all a will bear some resemblance to the above one due to the rather strenuous constraints on what the individual entries of the vectors in E can be.

5. Further Questions

There are other problems of these types that can be investigated: for example, rather than assuming that we have our distances in a subfield or subring of $\mathbb{R}$ or $\mathbb{F}_q$, an additive subgroup, or an arithmetic progression, or some other set with enough additive structure that we can use the method of constructing "hyperbolae" (i.e. sets having a fixed difference of distances to two fixed points, as previously in this paper) and bounding the size of the set by counting intersections of these hyperbolae. This does not work for arbitrary subrings of arbitrary fields—even in $\mathbb{R}$, trying to use $\mathbb{Q}$ instead of $\mathbb{Z}$ introduces problems because there are infinitely many rationals between any two distinct real numbers, so rather than looking at the intersections two finite sets of hyperbolae, we find ourselves looking at the intersections of two infinite sets, which is much less convenient.

References

[1] Bennett, M., Hart, D., Iosevich, A., et al. (2016). *Group actions and geometric combinatorics in* $\mathbb{F}_q^d$. Forum Math., 29(1), pp. 91-110.

[2] Bourgain, J., Katz, N., & Tao, T. *A sum-product estimate in finite fields, and applications.* Geom. Funct. Anal. (2004) 14:27.

[3] Erdős, P. *Integral distances.* Bull. Amer. Math. Soc. **51** (1945), no. 12, 996.