Proceedings of Symposia in Applied Mathematics
Volume 00, 0000

# Codes, quadratic forms and finite geometries

WILLIAM M. KANTOR

ABSTRACT. We study nonlinear binary error–correcting codes closely related to finite geometries and quadratic forms, and having links with extremal Euclidean line–sets and with recently introduced codes over $\mathbb{Z}_4$. Our emphasis is on geometric and combinatorial properties of highly structured families of codes.

The following diagram gives an indication of the main topics and interconnections arising in this paper.

$$\begin{tikzcd}
\text{line–set in } \mathbb{R}^{2^n} & & \text{line–set in } \mathbb{C}^{2^{n-1}} & \\
\text{Kerdock set} \arrow[u] \arrow[d] \arrow[r, leftarrow, shift left, "many"] \arrow[r, shift right] & \text{orthogonal spread} \arrow[r, shift left, "many"] \arrow[r, leftarrow, shift right] & \text{symplectic spread} \arrow[u] \arrow[d] \arrow[r, leftrightarrow] & \overset{(symplectic)}{\text{affine plane}} \\
\text{Kerdock code} & & \mathbb{Z}_4\text{–Kerdock code} \arrow[ll] &
\end{tikzcd}$$

## 1. Introduction

A binary error–correcting code $C$ of length $N$ is just a subset of the vector space $\mathbb{Z}_2^N$, the most standard ones being *linear codes*: subspaces of $\mathbb{Z}_2^N$. The *Hamming distance* $d_H$ between two vectors is the number of places they differ:

$$d_H\big((x_i),(y_i)\big) = (\#i : x_i \neq y_i).$$

One of the basic problems in the theory of error–correcting codes is to construct and study codes $C$ of length $N$ and large size $|C|$ subject to the condition that the minimum of the distances between any two different "codewords" in $C$ is some given integer $d$, the *minimum distance* of the code. Of particular interest are those codes that are extremal subject to such a condition. Evidently, such questions are highly combinatorial. Our purpose is to discuss aspects that are also within finite geometry and algebra. We will only touch on one type of connection of coding theory with finite geometry, one rich in a number of additional directions (projective planes, quadratic forms, Euclidean geometry, and groups). Many other connections with finite geometry are provided in [**MS**].

Another way to view a code of length $N$ is as a set of subsets of an $N$-element set. The transition between these views is elementary, just using an ordering

1991 *Mathematics Subject Classification.* Primary 94B27, 51A40; Secondary 05B25, 51A35.

The preparation of this paper was supported in part by the NSF.

I am grateful to A. R. Calderbank and M. E. Williams for many helpful remarks concerning preliminary versions of this paper.

of the $N$–set to associate an $N$-tuple (i.e., vector) with each subset. Addition becomes symmetric difference, and the Hamming distance is just the size of the symmetric difference of the corresponding subsets. The size of a subset is the *weight* of the corresponding word: the number of nonzero coordinates.

One way to search for families of subsets of a set is to impose additional structure on the set. We will assume throughout that the underlying set is itself a vector space $\mathbb{Z}_2^n$, so that $N = 2^n$. The most obvious family of subsets consists of all affine hyperplanes: all $n - 1$–dimensional subspaces and their translates. It is easy to see that, if the empty set and the whole space $\mathbb{Z}_2^n$ are also included, the result is a subspace of $\mathbb{Z}_2^{2^n}$, called the *first order Reed–Muller code* $RM(1, n)$. Note that $RM(1, n)$ is closely tied to the binary *affine space* $AG(n, 2)$ based on $\mathbb{Z}_2^n$: the analogue, in this binary setting, of real affine space based on a real vector space. Thus, we will never be far from this binary affine space.

Forms other than linear ones can be used. An especially rich source of "good" codes is the *second order Reed–Muller code* $RM(2, n)$, consisting of all of the zero–sets of all binary polynomials of degree at most 2 in $n$ variables. That is, $RM(2, n)$ has subcodes that behave in interesting manners, and in particular, extremal subcodes; historically, this has been the reason for the time devoted to subcodes of $RM(2, n)$ by numerous authors. We will focus on some of those arising from unions of cosets of $RM(1, n)$ in $RM(2, n)$.

These Reed–Muller codes can be written as follows in terms of coordinates (where some ordering is chosen for $\mathbb{Z}_2^n$):

$$RM(1, n) = \{\, (s \cdot v + \varepsilon)_{v \in \mathbb{Z}_2^n} \mid s \in \mathbb{Z}_2^n, \varepsilon \in \mathbb{Z}_2 \,\}$$

$$RM(2, n) = \{(Q(v) + s \cdot v + \varepsilon)_{v \in \mathbb{Z}_2^n} \mid Q \text{ is a quadratic form on } \mathbb{Z}_2^n,\ s \in \mathbb{Z}_2^n, \varepsilon \in \mathbb{Z}_2\},$$

where quadratic forms will be defined in the next section. Quadratic forms will then be used to construct (nonlinear) subcodes of $RM(2, n)$ called *Kerdock codes*. In §§3,4 we will see entirely different views of these codes in terms of orthogonal geometries and projective planes, which will lead us to structural properties and nonuniqueness results for the codes and for the various geometric objects associated with them.

The codes we focus on are *nonlinear*: they are not subgroups of $\mathbb{Z}_2^n$. Historically, linear codes have been the most important codes, since they are easier to discover, describe, encode and decode. On the other hand, the nonlinear codes studied here have the advantage of being superior from a combinatorial point of view: they have at least twice as many codewords as any linear code with same length and minimum distance. A surprising breakthrough in coding theory is that some of these nonlinear codes can be viewed as linear codes over $\mathbb{Z}_4$, rather than $\mathbb{Z}_2$ [**CHKSS**], and hence have the best of both worlds: superior description and implementation, and yet combinatorial optimality.

Much of this paper can be viewed as an introduction to Kerdock and other interesting subcodes of $RM(2, n)$ described in detail in [**MS**], codes which have just been investigated anew in [**CHKSS**] and [**CCKS**] from the vantage point of $\mathbb{Z}_4$ (cf. §6). The smallest example is the Nordstrom-Robinson code, described in Calderbank's paper in these Proceedings. Subcodes of $RM(2,n)$ also arose in the study, by Cameron and Seidel [**CS**], of extremal line–sets in Euclidean spaces (cf. §§5,6).

We will describe large numbers of codes by means of projective planes and nonassociative "algebras". This will lead to a better understanding of some of the mathematical underpinnings of the newly–discovered aspects of codes over $\mathbb{Z}_4$, besides producing new connections with other areas of mathematics.

## 2. Quadratic forms and Kerdock codes

Quadratic forms are standard in algebra and geometry. Care is needed when dealing with characteristic 2, but is well worth the effort: large numbers of important geometric and combinatorial objects (as well as groups) arise from them (cf. [**MS**] for many examples).

**Quadratic and alternating forms.** A *quadratic form* on a binary vector space $V$ is a map $Q: V \to \mathbb{Z}_2$ such that

$$(u,v) := Q(u+v) - Q(u) - Q(v) \tag{2.1}$$

is a symmetric bilinear form on $V$. In terms of coordinates, if $V = \mathbb{Z}_2^n$ then $Q$ looks like

$$Q(x_1,\dots,x_n) = \sum_{i \le j} a_{ij}x_ix_j$$

for some scalars $a_{ij}$, and $((x_i),(y_i)) = \sum_{ij} b_{ij}x_ix_j$, where $b_{ij} = b_{ji} = a_{ij}$ for $i < j$ and $b_{ii} = 0$. The *rank* of $Q$, or of its associated bilinear form, is just the rank of the matrix $B = (b_{ij})$ representing the bilinear form; this is also the codimension of the subspace consisting of all of the vectors $v$ such that $(v,V) = 0$. Also, $Q$ and the bilinear form are called *nonsingular* if $B$ is (equivalently, if $((v,V) = 0 \Longrightarrow v = 0)$.

The matrix $B$ is *skew–symmetric* (symmetric with 0 diagonal); its rank necessarily is even. The associated bilinear form is called an *alternating bilinear form* $(\ ,\ )$: $(u,v) = (v,u)$ and $(v,v) = 0$ for all $u, v$.

A vector space equipped with a nonsingular alternating bilinear form $(\ ,\ )$ is called a *symplectic space*. By a linear change of variables, any such form on $\mathbb{Z}_2^{2r}$ can be transformed into the form $\sum_{i=1}^r (x_iy_{i+r} - x_{i+r}y_i)$ (i.e., any nonsingular alternating bilinear form is *equivalent* to the indicated one). An *isometry*, or a *symplectic transformation*, is a nonsingular linear transformation $g$ of the vector space that preserves the form: $(ug, vg) = (u,v)$ for all vectors $u, v$.

**Singular vectors of quadratic forms.** It is easy to count the number of zeros (*singular vectors*) of quadratic forms over finite fields. Here we will only deal with the field $\mathbb{Z}_2$.

Let $Q_{2r}$ denote the quadratic form on $\mathbb{Z}_2^r \oplus \mathbb{Z}_2^r$ defined by $Q_{2r}(x,y) = x \cdot y$ for all $x, y \in \mathbb{Z}_2^r$ (ordinary dot product). Any quadratic form on $\mathbb{Z}_2^n$ can be transformed, using an affine change of coordinates (i.e., a transformation $v \mapsto vA + c$ for some invertible $n \times n$ matrix $A$ and some $c \in \mathbb{Z}_2^n$) to one of the following:

$$Q_{2r},\ Q_{2r} + 1 \textit{ or } Q_{2r} + z,$$

where $z$ is a variable not among those used for $Q_{2r}$ ("Dickson's Theorem" [**Dic**, **p. 197**], cf. [**MS p. 438**]). This makes it very easy to determine the number of zeros (*singular vectors*) of $Q$: an easy calculation shows that

$$\text{(2.2)} \qquad \begin{array}{l} Q_{2r} \textit{ and } Q_{2r} + 1 \textit{ have, respectively, exactly} \\ 2^{n-1} + 2^{n-1-r} \textit{ and } 2^{n-1} - 2^{n-1-r} \textit{ zeros in } \mathbb{Z}_2^n. \end{array}$$

It is straightforward to deduce that any *coset* $(Q(v))_v + RM(1,n)$, where $Q$ has rank $2r$, has weight distribution as follows:

| *weight* | *# of vectors of that weight* |
|---|---|
| $2^{n-1} - 2^{n-r-1}$ | $2^{2r}$ |
| $2^{n-1}$ | $2^{n+1} - 2^{2r+1}$ |
| $2^{n-1} + 2^{n-r-1}$ | $2^{2r}$ |

(Note that the complements of the vectors of weight $2^{n-1} + 2^{n-r-1}$ are those of weight $2^{n-1} - 2^{n-r-1}$.)

We will be interested in the largest possible weights, and hence will restrict to the case in which $n = 2r$. Then any quadratic form of rank $n$ in $n$ variables can be transformed, as above, to either $Q_n$ or $Q_n + 1$, in which case our coset becomes $(Q_n(v))_v + RM(1,n)$. Moreover, the above weight distribution is then even simpler (there are no vectors of weight $2^{n-1}$).

*cosets.* Since our characteristic is 2, unlike the familiar situation with real quadratic forms it is not possible to recover $Q$ from the bilinear form $(u,v)$: many different quadratic forms determine the same symmetric bilinear form. For any quadratic form $Q$ on $\mathbb{Z}_2^n$, *the coset* $(Q(v))_v + RM(1,n)$ *"contains" all quadratic forms* $Q'$ *determining the same bilinear form as* $Q$. Namely, if $Q$ and $Q'$ determine the same bilinear form, then

$$(u,v) = Q(u+v) - Q(u) - Q(v) = Q'(u+v) - Q'(u) - Q'(v)$$

for all $u, v \in \mathbb{Z}_2^n$, and then $Q - Q'$ is clearly a linear functional $\mathbb{Z}_2^n \to \mathbb{Z}_2$; this argument can be reversed.

We now turn to the behavior of a *union* of cosets $(Q(v))_v + RM(1,n)$, where $Q$ is allowed to run over a family $\mathcal{F}$ of quadratic forms on $\mathbb{Z}_2^n$ *each* of which has rank $n$. However, we require even more: we want the distance between the zero sets of any two different forms $Q, Q' \in \mathcal{F}$ to be large, which means that $Q - Q'$ should be another quadratic form of rank $n$. This condition is easier to understand in terms of the corresponding bilinear forms–or, better yet, in terms of the skew–symmetric matrices $B, B'$ determined by these bilinear forms: our requirement is that $B - B'$ is nonsingular. Thus, we are led to consider a set $\mathcal{K}$ of skew–symmetric $n \times n$ matrices the difference of any two of which is nonsingular. Then any two of these matrices have different first rows, so that there can be at most $2^{n-1}$ such matrices. The extremal case is the one of special interest here.

**Kerdock sets and Kerdock codes.** Kerdock sets and their associated codes and geometries are the principal subject of this paper.

A *Kerdock set* of $n \times n$ binary matrices is a family $\mathcal{K}$ of $2^{n-1}$ skew–symmetric $n \times n$ binary matrices, containing $O$, such that the difference of any two is nonsingular. (Note that $n$ is even since we are dealing with skew–symmetric matrices.) For the reason indicated above, this number $2^{n-1}$ is extremal. In combinatorial settings, extremal configurations frequently have rich structures. This is very much the case with Kerdock sets. We will construct such sets very soon, but first we construct codes using them.

Each Kerdock set $\mathcal{K}$ determines a *Kerdock code* $C(\mathcal{K}) = \bigcup_{B \in \mathcal{K}}[(Q_B(v))_v + RM(1,n)]$, where $Q_B$ denotes any quadratic form whose associated bilinear form is $uBv^t$. Thus, if $B = (b_{ij})$ and if $U$ denotes the upper triangular matrix obtained from $B$ by replacing all entries below the diagonal by 0 (so that $U + U^t = B$), then we may assume that $Q_B(v) = vUv^t$. Explicitly, in terms of vectors we then have

$$C(\mathcal{K}) := \left\{ \left(Q_B(v) + s \cdot v + \varepsilon\right)_{v \in \mathbb{Z}_2^n} \mid B \in \mathcal{K}, s \in \mathbb{Z}_2^n, \varepsilon \in \mathbb{Z}_2 \right\}. \tag{2.3}$$

$C(\mathcal{K})$ is a code of length $N = 2^n$ (where $n$ is even), consisting of $2^{n-1}2^n 2 = 2^{2n}$ codewords (i.e., vectors). Any $c \in C(\mathcal{K})$ partitions $C(\mathcal{K})$ in terms of distances, as follows:

| *distance from c* | *# of words at that distance* | *comments* |
|---|---|---|
| $0$ | $1$ | $c$ |
| $2^{n-1} - 2^{(n-2)/2}$ | $2^n(2^{n-1} - 1)$ | |
| $2^{n-1}$ | $2^{n+1} - 2$ | |
| $2^{n-1} + 2^{(n-2)/2}$ | $2^n(2^{n-1} - 1)$ | |
| $2^n$ | $1$ | $c +$ (the all–1 vector) |

Minimum distance: $2^{n-1} - 2^{(n-2)/2}$.

The property that the distribution of distances is independent of the choice of $c$ is called *distance–invariance*. It as an approximation to the linearity of a code, this property being trivial for such codes. (That is, if $C$ is a linear code then the translations $v \mapsto v + c$ form a group of automorphisms transitive on the set of codewords, so that distance–invariance is obvious.) as we will see below (Theorem 5.1), $C(\mathcal{K})$ is nonlinear.

*Recovering $\mathcal{K}$ from $C(\mathcal{K})$.* $RM(1,n)$ is the set of codewords of weight $0$, $2^{n-1}$ or $2^n$. $C(\mathcal{K})$ is a union of cosets of $RM(1,n)$, each of which corresponds to a unique skew–symmetric matrix. These matrices comprise $\mathcal{K}$.

*Examples of Kerdock sets.*

Here is an example of a set of $2^{4-1}$ quadratic forms when $n=4$:

$\{0,\ x_1x_2+x_3x_4,\ x_1x_3+x_1x_4+x_2x_3+x_2x_4,$
$x_1x_3+x_2x_4+x_3x_4,\ x_1x_2+x_1x_3+x_1x_4+x_2x_4, x_1x_4+x_2x_3+x_2x_4+x_3x_4,$
$x_1x_2+x_1x_4+x_2x_3,\ x_1x_2+x_1x_3+x_2x_3+x_2x_4\}.$

This description is opaque. It is not motivated (nor will our other examples be well–motivated until later, when we get to projective planes), and it is tedious to verify that all differences of these quadratic forms are nonsingular (cf. [**Li**]).

All remaining examples discussed here will be obtained from the field GF$(2^m)$, where $m$ is *odd* (this will be "$n-1$" in our previous notation). Let $T:\mathrm{GF}(2^m)\to \mathrm{GF}(2)$ be the trace map: $T(x)=\sum_{i=0}^{m-1}x^{2^i}$. This determines an inner product $T(xy)$ on GF$(2^m)$. There is an orthonormal basis that lets us identify GF$(2^m)$ equipped with this inner product and $\mathbb{Z}_2^m$ equipped with its usual dot product.

We are searching for Kerdock sets, which require even–dimensional spaces whereas $m$ is odd. Hence, we boost the dimension by 1, and consider $\mathrm{GF}(2^m)\oplus\mathbb{Z}_2$, equipped with the inner product $\big((x,a),(y,b)\big):=T(xy)+ab$. In place of skew–symmetric matrices we will use linear operators $M\colon \mathrm{GF}(2^m)\oplus\mathbb{Z}_2\to \mathrm{GF}(2^m)\oplus\mathbb{Z}_2$ such that $((x,a),(x,a)M)=0$ for all $(x,a)\in \mathrm{GF}(2^m)\oplus\mathbb{Z}_2$. We will construct families of such linear operators $M$ by using suitable binary operations on GF$(2^m)$.

EXAMPLE 2.4. Consider the set $\mathcal{K}$ of $2^m$ linear operators $M_s\colon \mathrm{GF}(2^m)\oplus\mathbb{Z}_2\to \mathrm{GF}(2^m)\oplus\mathbb{Z}_2$ given by

$$(x,a)M_s=\big(xs^2+sT(sx)+as, T(sx)\big)\,.$$

This definition has been pulled out of the blue, and will be motivated later in terms of projective planes (§4). For now, we note that *these matrices $M_s$ form a Kerdock set $\mathcal{K}$.* First, $M_s$ is skew–symmetric:

$$\big((x,a),(x,a)M_s\big)=T\big(x[xs^2+sT(sx)+as]\big)+aT(sx)=0$$

since $T(xxs^2)=T(xs)^2=T(xsT(sx))$. Next, $M_r-M_s$ is nonsingular whenever $r\neq s$: if

$$(xr^2+rT(rx)+ar, T(rx))=(xs^2+sT(sx)+as, T(sx))\,,$$

then $T(rx)=T(sx)$ and $xr^2+rT(sx)+ar=xs^2+sT(sx)+as$, so that

$$x(r+s)^2+(r+s)T(sx)+a(r+s)=0\,,$$

$$rx+sx+T(sx)+a=0\,,$$
$$T(rx)+T(sx)+T(sx)+a=0$$

(this uses the fact that $m$ is odd, so that $T(1)=1$), and hence

$$rx+sx=0=T(sx)+a.$$

Thus, $M_r-M_s$ is nonsingular, as required.

By (2.3), this Kerdock set produces a Kerdock code $C(\mathcal{K})$. This is the original code discovered by Kerdock [**Ke**] in 1972 (in rather different language); cf. [**Di;MS;Ka1**]. When $m=3$ this is the Nordstrom–Robinson code of length 16. Moreover, we now see that *a Kerdock code of length* $2^n$ *exists for every even* $n$.

**Kerdock set equivalence.** Kerdock sets $\mathcal{K}_1$ and $\mathcal{K}_2$ of $n\times n$ matrices are *equivalent* if there is an invertible $n\times n$ matrix $A$ and a skew–symmetric matrix $M$ such that $A^t\mathcal{K}_1A+M=\mathcal{K}_2$. One of our goals is to describe large numbers of inequivalent Kerdock sets (and corresponding codes) by means of projective planes and nonassociative "algebras".

Example 2.4 used field multiplication (the term $xs^2$). As we will seen in §4, important types of projective planes are described using more general types of binary operations. Hence, we are led to introduce those operations important in our coding–theoretic context. These have the advantage of being the quickest way to write down what amount to *all* Kerdock sets (cf. Theorem 4.2)..

**Binary operations.** Consider a binary operation $*$ on $\mathrm{GF}(2^m)$ related to field multiplication by the following conditions (for all $x,y,z\in\mathrm{GF}(2^m)$):
**(i)** $(x+y)*z=x*z+y*z$ (left distributivity),
**(ii)** $x*y=x*z\Longrightarrow x=0$ or $y=z$,
**(iii)** $T(x(x*y))=T(xy)$, and
**(iv)** $x*y=0\iff x=0$ or $y=0$.
Also, **(iii)** implies
**(v)** $T(x(y*z))=T(y(x*z))$
(namely, in **(iii)** replace $x$ by $x+z$, $x$ and $z$, and subtract). This condition is more useful for our purposes; but if **(v)** holds and **(iii)** does not, it is easy to modify $*$ insignificantly so that **(iii)** will hold.

A fundamental aspect of the subject matter in this paper is that **(i)** and **(ii)** are familiar in the theory of projective planes (cf. §4). These amount to some distorted versions of fields; for example, if both distributive laws hold then we are dealing with a special type of (nonassociative) division algebra (except for the lack of an identity element). It is just such field–like algebras that arise in the coordinatization of projective planes. Thus, one can expect that there will be further interactions between coding theory and planes, with a great deal to be learned in each discipline from the other one.

In Example 2.4, $x*y=xy^2$ (note that $T(x(x*y)=T(x^2y^2)=T(xy)$, which explains the use of $xy^2$ instead of the more natural–looking $xy$). The argument used in that example essentially shows that

Proposition 2.5. *The maps*

$$M_s\colon (x,a) \mapsto \bigl(x*s + sT(sx) + as, T(sx)\bigr),\ \ s \in \mathrm{GF}(2^m)\,,$$

*form a Kerdock set.*

Namely, proceed exactly as before in order to obtain the equation $x*r+x*s = (r+s)[T(sx)+a]$. If $(x,a)$ also satisfies $T(sx)+a=0$ then, by **(v)**, $x=0$. It follows that the dimension of the kernel of $M_r - M_s$ is at most 1, and hence is 0 since $m+1$ is even and $M_r - M_s$ is skew–symmetric.

Example 2.6. Let $T_1$ denote the trace map from $\mathrm{GF}(2^m)$ to some proper subfield $F \neq \mathbb{Z}_2$. Then

$$x*s\colon = xs^2 + T_1(x)s + T_1(xs)$$

satisfies **(i-v)**. (N.B.—If we allowed $F=\mathbb{Z}_2$ here, then the Kerdock set obtained in this manner would be the same as the one obtained in Example 2.4.)

Once again, these maps appear to have come from nowhere, but will turn out to be motivated by projective planes.

Example 2.7. Let $T_1$ again denote the trace map from $\mathrm{GF}(2^m)$ to some proper subfield $F \neq \mathbb{Z}_2$, let $\alpha \in F - \mathbb{Z}_2$, and write

$$x*s\colon = xs^2 + \alpha s T_1(xs).$$

Then this operation satisfies **(i-v)**, but this time only one distributive law is satisfied. Once again a Kerdock set is obtained using the preceding proposition.

It might appear that the above conditions **(i-v)** are so strong as to prevent the existence of many examples. This is not the case. Every Kerdock set is, in a suitable sense, equivalent to one of those in Proposition 2.5 (see the Remark following Proposition 3.6). One can, of course, ask whether $\mathcal{K}$ determines the "algebra" $(\mathrm{GF}(2^m),+,*)$ uniquely up to something like isomorphism. However, this also is not the case if $m \geq 5$: there is a strong version of non-uniqueness, which is essentially the content of Theorem 4.4. Moreover, there are large numbers of inequivalent Kerdock sets (cf. Theorem 3.5).

## 3. Finite orthogonal geometries[1]

Quadratic forms arise for us in two different ways. On the one hand, we have used them in order to construct codes (second order Reed–Muller codes and Kerdock codes). On the other hand, we will use them in order to construct a very different-looking type of configuration in larger vector spaces ("orthogonal spreads"). We will also see that the two ways to use quadratic forms are nicely linked, albeit in a somewhat indirect manner.

[1] Much of this section only uses the fact that the field is finite of characteristic 2, rather than $\mathbb{Z}_2$. The exception is the part of Theorem 3.4 that concerns codes.

**An orthogonal geometry.** In order to have a more geometric view of skew–symmetric matrices, we will double the dimension and consider one concrete type of orthogonal geometry. Let $V = \mathbb{Z}_2^{2n} = X \oplus Y$ for subspaces $X$ and $Y$ both of which are identified with $\mathbb{Z}_2^n$. Equip $V$ with the quadratic form $Q = Q_{2n}$ (so that $Q(x, y) = x \cdot y$), with associated bilinear form $(\ ,\ )$. The notion of perpendicularity is as usual. However, note that every vector is perpendicular to itself (since $(u, u) = 0$). If $W$ is any subspace of $V$ then $W^\perp := \{v \in V \mid (v, W) = 0\}$ is a subspace of dimension $\dim V - \dim W$. For example, $X^\perp = X$ and $Y^\perp = Y$.

*Totally singular subspaces.* A subspace $W$ is ***totally singular*** if $Q(W) = 0$, in which case it also is perpendicular to itself (i.e., $W \subseteq W^\perp$ since $(W, W) = 0$), and hence $\dim W \le n$ (since $2n = \dim W + \dim W^\perp \ge 2 \dim W$). Thus, $X$ and $Y$ are examples of totally singular $n$–spaces.[2]

*Orthogonal spreads.* By (2.2), $V$ has $(2^n - 1)(2^{n-1} + 1)$ nonzero singular vectors. Each totally singular $n$–space consists of singular vectors, and contains $2^n - 1$ nonzero ones. This number divides the number of singular vectors and suggests that there might be families $\Sigma$ of totally singular $n$–spaces that partition the set of all nonzero singular vectors. Such a family of $2^{n-1} + 1$ subspaces is called an *orthogonal spread*. We will see soon that such a family cannot exist unless $n$ is even, and that there is always at least one such family when $n$ is even.

*Isometries.* An *isometry* of $V$ is a nonsingular linear transformation preserving $Q = Q_{2n}$ (i.e., a nonsingular linear transformation $T$ such that $Q(vT) = Q(v)$ for all $v \in V$); these form a group, the *orthogonal group* $O^+(2n, 2)$ of $V$. (Here, the "+" refers to the fact that $V$ has totally singular $n$–spaces.) This group is transitive on the ordered pairs of totally singular $n$–spaces having only 0 in common: from the point of view of this orthogonal geometry of $V$, the pair $X, Y$ we started with is indistinguishable from any other such pair.

Fix a basis $x_1, \ldots, x_n$ of $X$ and let $y_1, \ldots, y_n$ be the dual basis of $Y$: $(x_i, y_j) = \delta_{ij}$. Write matrices with respect to the basis $x_1, \ldots, x_n, y_1, \ldots, y_n$.

Lemma 3.1.

(i) *The isometries of $V$ that fix every vector of $Y$ are just those linear transformations whose matrices are $\left(\begin{smallmatrix} I & M \\ O & I \end{smallmatrix}\right)$ for some binary skew–symmetric $n \times n$ matrix $M$.*

(ii) *These isometries form a group isomorphic to the additive group of all binary skew–symmetric $n \times n$ matrices.*

(iii) *The isometries fixing $Y$ are just those linear transformations whose matrices are $\left(\begin{smallmatrix} A^{-t} & O \\ O & A \end{smallmatrix}\right)\left(\begin{smallmatrix} I & M \\ O & I \end{smallmatrix}\right)$, where $A$ runs through the group $GL(n, 2)$ of all nonsingular $n \times n$ binary matrices and $M$ is as in* (i).

[2] Similarly, in the case of a $2m$–dimensional symplectic space, a subspace $W$ is *totally isotropic* if $W \subseteq W^\perp$, in which case $\dim W \le m$.

PROOF.

(i) Any such isometry must look like $\begin{pmatrix} I & M \\ O & B \end{pmatrix}$ for some $n \times n$ matrices $M$ and $B$. Then $Q$ is preserved if and only if $x \cdot (xM + yB) = x \cdot y$ for all $x, y$, which is the case if and only if $M$ is skew–symmetric (use $y = 0$) and $B = I$.

(ii) $\begin{pmatrix} I & M \\ O & I \end{pmatrix}\begin{pmatrix} I & N \\ O & I \end{pmatrix} = \begin{pmatrix} I & M+N \\ O & I \end{pmatrix}$.

(iii) Any such isometry $g$ induces a linear transformation on $Y$, with matrix $A$, say. The matrix $\begin{pmatrix} A & O \\ O & A^{-t} \end{pmatrix}$ arises from an isometry $h$ of $V$, and $h^{-1}g$ is as in (ii). □

LEMMA 3.2.

(i) *Every totally singular $n$–space $Z$ of $V$ such that $Y \cap Z = 0$ has the form $X\begin{pmatrix} I & M \\ O & I \end{pmatrix}$ for a unique skew–symmetric matrix $M$. Conversely, if $M$ is a skew–symmetric binary $n \times n$ matrix, then $X\begin{pmatrix} I & M \\ O & I \end{pmatrix}$ is a totally singular $n$–space having only 0 in common with $Y$.*

(ii) *The dimension of the intersection of any two such $n$–spaces $X\begin{pmatrix} I & M \\ O & I \end{pmatrix}$ and $X\begin{pmatrix} I & N \\ O & I \end{pmatrix}$ is $n - \operatorname{rank}(M - N)$*

PROOF.

(i) We can write $Z = \{(x, xM) \mid x \in X\}$ for a unique $n \times n$ matrix $M$. The requirement that $Z$ be totally singular is equivalent to having $x \cdot xM = 0$ for all $x \in X$; and this is precisely the condition of skew–symmetry.

(ii) The desired dimension is that of the set of solutions to the following system of linear equations: $(x, xM) = (x, xN)$. □

Part (i) says that $Z$ consists of all of the vectors of the form $(x, xM)$; or, in more familiar terms, $Z$ is the subspace "$y = xM$". Note that the group in Lemma 3.1(ii) is transitive on the set of subspaces $Z$ occurring in Lemma 3.2(i). By Lemma 3.2(ii), *if some such pair of subspaces meet only at* 0, *then $n$ must be even.*

In view of the definition of Kerdock sets in §2, we have the

COROLLARY 3.3.

(i) *If $\mathcal{K}$ is a Kerdock set of $n \times n$ skew–symmetric binary matrices, then*

$$\Sigma := \{Y\} \bigcup \left\{ X\begin{pmatrix} I & M \\ O & I \end{pmatrix} \,\middle|\, M \in \mathcal{K} \right\}$$

*is an orthogonal spread of $V$.*

(ii) *Conversely, every orthogonal spread of $V$ that contains both $X$ and $Y$ arises as in (i).*

Evidently, in (ii) $\Sigma$ depends on $\mathcal{K}$. Is it possible that inequivalent choices for $\mathcal{K}$ produce "equivalent" orthogonal spreads $\Sigma$? The answer is "yes": see

Theorem 3.4 and the remark following it. In any event, we now see that *an orthogonal spread exists if and only if $n$ is even.*

**Equivalence of equivalences.** We now turn to the relationships between pairs of Kerdock codes, pairs of Kerdock sets, and pairs of orthogonal spreads.

An *equivalence* between binary codes of length $N$ is a permutation of coordinates sending one to the other. Automorphisms of codes are then equivalences of a code with itself. Since our coordinates are indexed by vectors in $\mathbb{Z}_2^n$, an equivalence will look like $(a_v)_{v\in\mathbb{Z}_2^n} \mapsto (a_{v\sigma})_{v\in\mathbb{Z}_2^n}$ for a permutation $\sigma$ of $\mathbb{Z}_2^n$. For example, each translation $v \mapsto v+c$ of $\mathbb{Z}_2^n$ is an automorphism of $C(\mathcal{K})$ since it leaves invariant each coset $(Q(v))_v+RM(1,n)$. (Namely, $Q(v+c) = Q(v)+Q(c)+(v,c)$, where $(Q(c)+(v,c))_v \in RM(1,n)$.)

A *quasi-equivalence* between binary codes of length $N$ is a map of the form $(a_v)_v \mapsto (a_{v\sigma}+c_v)_v$, sending one to the other, where $\sigma$ is a permutation of coordinates and $(c_v)_v$ is some vector in $\mathbb{Z}_2^N$. Thus, two codes are quasi–equivalent if and only each is the image of the other by means of an isometry of the underlying metric space ($\mathbb{Z}_2^N$, Hamming metric). In the case of linear codes, this notion is almost the same as equivalence. For a nonlinear code $C$, even one containing 0, it is noticeably weaker: if $w \in C$, then $C$ and $C+w$ are quasi–equivalent but not equivalent, and yet clearly they are not “significantly” different.

Equivalence of Kerdock sets was defined earlier.

Orthogonal spreads $\Sigma_1$ and $\Sigma_2$ of $V$ are *equivalent* if there is an isometry of $V$ sending $\Sigma_1$ to $\Sigma_2$.

THEOREM 3.4. *Let $\mathcal{K}_1$ and $\mathcal{K}_2$ be Kerdock sets of $n\times n$ binary matrices. Then the following are equivalent:*

(i) *$\mathcal{K}_1$ and $\mathcal{K}_2$ are equivalent;*

(ii) *$C(\mathcal{K}_1)$ and $C(\mathcal{K}_2)$ are quasi–equivalent;*

(iii) *The orthogonal spreads $\Sigma_1$ and $\Sigma_2$ of $V$, determined, respectively, by $\mathcal{K}_1$ and $\mathcal{K}_2$ via* Corollary 3.3(i), *are equivalent by an isometry of $V$ sending $Y$ to itself.*

PROOF. (ii)⇒(i) Suppose that $g\colon (a_v)_v \mapsto (a_{v\sigma}+c_v)_v$ is a quasi-equivalence sending $C(\mathcal{K}_1)$ to $C(\mathcal{K}_2)$. Then $(0)_v g = (c_v)_v$ is in $C(\mathcal{K}_2)$, and hence has the form $(Q(v)+s\cdot v+\varepsilon)_v$ for some quadratic form $Q$, some $s\in\mathbb{Z}_2^n$, and $\varepsilon=\pm1$. Let $h$ be the map $(a_v)_v \mapsto (a_v+c_v)_v$. Then $gh$ sends $(a_v)_v \mapsto (a_{v\sigma})_v$, and sends $C(\mathcal{K}_2)$ to $C(\mathcal{K}_2)+(c_v)_v = C(\mathcal{K}_2)+(Q(v))_v$, which still contains $RM(1,n)$.

The words in $C(\mathcal{K}_1)$ of weight $2^{n-1}$, and the words in $C(\mathcal{K}_2)+(Q(v))_v$ of weight $2^{n-1}$, are the hyperplanes of $AG(n,2)$, corresponding to $RM(1,n)$. It follows that $\sigma$ has the form $v\mapsto vA+w$ for some invertible $n\times n$ matrix $A$ and some $w\in\mathbb{Z}_2^n$; we may assume that $w=0$ since $v\mapsto v+w$ is an automorphism of $C(\mathcal{K}_1)$. Each word of $C(\mathcal{K}_1)$ or $C(\mathcal{K}_2)$ containing 0 (i.e., having 1 in the 0 position) arises from a quadratic form. Hence, each quadratic form $Q_1$ giving rise to a codeword of $C(\mathcal{K}_1)$ produces, via $A$, a quadratic form $Q_2+Q$ giving rise to a codeword of $C(\mathcal{K}_2)+(Q(v))_v$; that is, $Q_1(vA) = (Q_2+Q)(v)$ for all

$v \in \mathbb{Z}_2^n$. If $M$ and $B_i$ are the matrices of the alternating bilinear forms produced by $Q$ and $Q_i$, then this means that $(uA)B_1(vA)^t = u(B_2+M)v^t$ for all $u, v$, and hence that $AB_1A^t = B_2 + M$.

(i)⇒(ii) Reverse the above argument.

(i)⇔(iii) If $A$ is an invertible $n \times n$ matrix and $M$ is a skew–symmetric matrix such that $A^t\mathcal{K}_2A + M = \mathcal{K}_1$, then the matrix pictured in Lemma 3.1(iii) sends $\Sigma_2$ to $\Sigma_1$. For the converse, reverse the argument. □

The above theorem needs to be examined carefully. Inequivalent Kerdock sets can produce equivalent orthogonal spreads, a possibly confusing fact that has that has occasionally been overlooked **[Li;CL]**. Many examples of this phenomenon exist. In fact, this situation is the norm: it occurs whenever the group $G(\Sigma)$ of orthogonal transformations preserving an orthogonal spread $\Sigma$ is intransitive on $\Sigma$, and there appear to be few examples where $G(\Sigma)$ is actually transitive on $\Sigma$.

*Inequivalent codes.* Assume that $m$ is odd and $m \geq 5$. In **[Ka1]** it is shown that Kerdock codes arising from Examples 2.6 and 2.7 are not quasi–equivalent; and that two codes arising from Example 2.6 or 2.7 are quasi–equivalent if and only if they are equivalent under a permutation of $\mathrm{GF}(2^m)$ of the form $x \mapsto ax^\sigma + b$ for some $a, b \in \mathrm{GF}(2^m)$, $a \neq 0$, and some $\sigma \in \mathrm{AutGF}(2^m)$. By using intermediate fields in order to vary these constructions (a hint of this is in Examples 2.6 and 2.7), the following much stronger result has been proved by Williams:

THEOREM 3.5 **[Wi]**. *Let $m$ be an odd integer $> 1$. Let $m, m_1, \dots, m_r, 1$ be a sequence of $r+2 \geq 3$ divisors of $m$ such that each is a proper divisor of the preceding one. If $m \geq 7m_1$ then there are more than $2^{(r-1)m}/m$ pairwise inequivalent Kerdock sets of $(m+1) \times (m+1)$ matrices, and hence at least that many pairwise quasi–inequivalent Kerdock codes of length $2^{m+1}$.*

A similar result from **[Wi]** is found below in Theorem 6.6. Williams expects to prove similar results for analogues of Example 2.7, producing Kerdock sets admitting a cyclic automorphism group fixing one member and transitive on the remaining ones (and, more generally, producing orthogonal spreads admitting a cyclic automorphism group fixing two members and transitive on the remaining ones). Much weaker versions of this type of result are contained in **[Ka1,Ka2]**. There is also the following related result:

THEOREM 3.5′ **[KW]**. *Let $m$ be an odd integer $> 1$. Let $m, m_1, \dots, m_r, 1$ be a sequence of $r+2 \geq 3$ divisors of $m$ such that each is a proper divisor of the preceding one. Then there are at least $[\Pi_1^r(2^{m_i}+1)]/2m_1$ pairwise inequivalent orthogonal spreads $\Sigma$ in the usual binary orthogonal $2m+2$–space such that $G(\Sigma)$ has a cyclic subgroup transitive on $\Sigma$.*

These results, and the determination of the automorphism groups of the Kerdock sets or codes as well as the automorphism groups of the orthogonal spreads, rest on Theorem 3.4 together with Theorem 4.4 below. First we need to see how to construct orthogonal spreads from other geometric objects.

**To symplectic spreads.** Let $z$ denote any nonsingular 1–space of $V$, so that $Q(z) \neq 0$. If $Z$ is any totally singular $n$–space of $V$ then $Z^\perp = Z$, so that $Z \not\subseteq z^\perp$. Consequently, if $\Sigma$ is any orthogonal spread of $V$, then

$$\{Z \cap z^\perp \mid Z \in \Sigma\}$$

is a family of totally singular $n-1$–spaces of $z^\perp$ such that every nonzero singular vector is in exactly one of these subspaces.

Recall that $z$ is contained in the hyperplane $z^\perp$. The $2n-2$–space $z^\perp/z$ inherits the nonsingular alternating bilinear form from $V$ (but not the quadratic form):

$$(u+z, v+z) := (u, v)$$

is well-defined on $z^\perp$ (but "$v + z \mapsto Q(v)$" is not). This turns $z^\perp/z$ into a symplectic space. (Recall from the fourth paragraph of §2 that any two symplectic spaces of the same dimension are equivalent.)

Now we can "project" $\Sigma$ into $z^\perp/z$, obtaining a set

$$\Sigma_z := \{\langle Z \cap z^\perp, z\rangle/z \mid Z \in \Sigma\}$$

consisting of $|\Sigma| = 2^{n-1}+1$ totally isotropic $n-1$–spaces of $z^\perp/z$ such that any two meet only in 0. This is called a *symplectic spread* of the symplectic space $z^\perp/z$: each nonzero vector of $z^\perp/z$ is contained in exactly one member of $\Sigma_z$.

**From symplectic spreads.** The preceding construction can be reversed, proceeding from symplectic spreads to orthogonal ones. This can be accomplished geometrically or in terms of matrices. We will use the latter approach, since it requires no additional background.

Let $V' := X' \oplus Y'$ be the direct sum of two $m$–dimensional subspaces $X'$ and $Y'$, each of which we identify with $\mathbb{Z}_2^m$. Equip $V'$ with a nonsingular alternating bilinear form $((x'_1, y'_1), (x'_2, y'_2)) = x'_1 \cdot y'_2 - x'_2 \cdot y'_1$ for $x'_1, x'_2 \in X'$, $y'_1, y'_2 \in Y'$, so that both $X'$ and $Y'$ are totally isotropic $m$–spaces. Fix dual bases $x'_1, \ldots, x'_m$ and $y'_1, \ldots, y'_m$ of $X'$ and $Y'$; write matrices with respect to the basis $x'_1, \ldots, x'_m, y'_1, \ldots, y'_m$. As in Lemma 3.2, every totally isotropic $m$–space $Z'$ such that $X' \cap Z' = 0$ can be written as $X'\begin{pmatrix} I & P \\ O & I \end{pmatrix}$ for a unique *symmetric* matrix $P$. Two such $m$–spaces $X'\begin{pmatrix} I & P \\ O & I \end{pmatrix}$ and $X'\begin{pmatrix} I & R \\ O & I \end{pmatrix}$ have only 0 in common if and only if $P - R$ is nonsingular.

Thus, a symplectic spread in $V'$ containing $X'$ and $Y'$ arises from a set $\mathcal{S}$ of $2^m$ symmetric matrices, containing $O$, such that the difference of any two is nonsingular. We have seen above that any orthogonal spread $\Sigma$ of $V$, together with a nonsingular 1-space $z$, determines a symplectic spread $\Sigma_z$ in the symplectic space $z^\perp/z$ of dimension $2m = 2n-2$. In terms of the basis for $V$ introduced earlier, assume that $z = \langle x_n + y_n\rangle$. We identify $X'$ and $Y'$ with $\langle X \cap z^\perp, z\rangle/z$ and

$\langle Y \cap z^\perp, z\rangle/z$, respectively; let the basis chosen for $z^\perp/z$ consist of the vectors $x_i' = x_i + z$ and $y_i' = y_i + z$, $1 \le i \le m = n - 1$.

PROPOSITION 3.6 [**CCKS**].

$$\mathcal{S} = \left\{ M_1 + d^t d \,\middle|\, \begin{pmatrix} M_1 & d^t \\ d & 0 \end{pmatrix} \in \mathcal{K} \right\} \textit{ and}$$

$$\mathcal{K} = \left\{ \begin{pmatrix} P + d(P)^t d(P) & d(P)^t \\ d(P) & 0 \end{pmatrix} \,\middle|\, P \in \mathcal{S} \right\},$$

*where $d(P)$ is the vector in $\mathbb{Z}_2^m$ whose coordinates are the diagonal entries of $P$ in their natural order.*

PROOF (SKETCH). Consider a totally singular subspace

$$\{(x, xM) \mid x \in X\} = \left\{ \left( (x', a), (x', a) \begin{pmatrix} M_1 & d^t \\ d & 0 \end{pmatrix} \right) \,\middle|\, (x', a) \in X \oplus \mathbb{Z}_2 \right\}$$

as in Lemma 3.2(i). Its intersection with $z^\perp = (0,1,0,1)^\perp$ consists of those vectors $(x', a, x'M_1 + ad, x'd^t)$ such that $a = x'd^t$, and hence is the union of the cosets $(x', 0, x'M_1 + x'd^t d, 0) + z = (x', 0, x'[M_1 + d^t d], 0) + z$ with $(x', 0) \in X$. □

REMARKS. The equation $M = \begin{pmatrix} P + d(P)^t d(P) & d(P)^t \\ d(P) & 0 \end{pmatrix}$ defines a bijection $P \mapsto M$ from symmetric $m \times m$ matrices $P$ to skew–symmetric $(m+1) \times (m+1)$ matrices $M$. Indeed, given a skew–symmetric $(m+1) \times (m+1)$ matrix $M$, let its last row be $(d\ 0)$ and find $P$ from the principal minor indicated in the above equation. Conversely, given a symmetric $m \times m$ matrix $P$, observe that the matrix $M$ defined above is, indeed, skew–symmetric (the diagonal of $d(P)^t d(P)$ is that of $P$ since our field is $\mathbb{Z}_2$). This bijection $P \mapsto M$ is not linear. It is an easy exercise to show that, since $n$ is even, if $P \mapsto M$ and $P' \mapsto M'$, then $P - P'$ is nonsingular if and only if $M - M'$ is.

Proposition 3.6 is very closely related to Proposition 2.5: see Theorem 4.2.

EXAMPLE 3.7. Suppose that $\Sigma'$ is a *desarguesian spread*: the set of 1-spaces of $\mathrm{GF}(2^m)^2$. There is an obvious alternating bilinear form on $\mathrm{GF}(2^m)^2$, given by $\det\binom{u}{v}$. When followed by the trace map $\mathrm{GF}(2^m) \to \mathbb{Z}_2$, this produces a nonsingular alternating bilinear form on $\mathbb{Z}_2^{2m}$ such that $\Sigma'$ is still a symplectic spread. Now identify $\mathbb{Z}_2^{2m}$ with $z^\perp/z$. Then Corollary 3.3 produces an orthogonal spread $\Sigma$ of $V$, which is in fact the orthogonal spread arising from the Kerdock set in Example 2.4. This reflects the prominence of field multiplication there (the term $xs^2$). The group $SL(2, 2^m)$ that acts on $\mathrm{GF}(2^m)^2$, preserving its set $\Sigma'$ of 1-spaces, also preserves the alternating bilinear form on $\mathbb{Z}_2^{2m} \equiv z^\perp/z$, and lifts to a subgroup of $O^+(2m+2, 2)$ that acts on $\Sigma$ as it does on the set of 1-spaces of $\mathrm{GF}(2^m)^2$ (in particular, this subgroup is 3-transitive on $\Sigma$).

*Up and down.* The Kerdock sets arising from Examples 2.6 and 2.7 are obtained from a slight variation on the example, using different choices of $z$ and taking into account an intermediate field between $\mathrm{GF}(2^m)$ and $\mathbb{Z}_2$ in order to get a different orthogonal spread $\Sigma$.

Starting with a symplectic spread $\Sigma'$ in a $2m$–dimensional binary symplectic space, we now can produce an orthogonal spread in a $2m+2$–dimensional orthogonal space, in such a way that there is a nonsingular 1-space $z$ for which $\Sigma_z$ is $\Sigma'$. Once we have $\Sigma$, we can then form a *different* symplectic spread $\Sigma_{z'}$ using a *different* nonsingular 1-space $z'$. When combined with passage to subfields [**Ka1**], this type of up and down process leads to the proof of Theorem 3.5.

## 4. Projective planes

We now wander even further from the traditional coding theory questions we started with: an entirely different type of geometric view of symplectic spreads is provided by projective planes. For this purpose we first need to ignore, temporarily, the word "symplectic".

**From spreads to projective planes.** Let $V'$ be a $2m$–dimensional vector space over $\mathrm{GF}(q)$ (no restriction is placed even on the parity of $q$ or $m$).

A *spread* of $V'$ is a family $\Sigma'$ of $q^m+1$ subspaces of dimension $m$ whose union is all of $V'$. This means that every nonzero vector is in a unique member of $\Sigma'$. *Any family of $q^m+1$ $m$–spaces in a $2m$–space, any two of which have only* 0 *in common, is a spread.* (N.B.—An orthogonal spread is not a spread in this sense, but a symplectic spread is.)

*Affine planes.* The importance of spreads is that they produce affine planes: Let $\mathbf{A}(\Sigma')$ denote the point–line geometry whose points are vectors and whose lines are the cosets $W+v$ with $W \in \Sigma', v \in V'$. Then $\mathbf{A}(\Sigma')$ is an *affine plane of order $q^m$*:

- Any two different points $u, v$ are on a unique line (namely, the line $W+v$ where $u-v \in W \in \Sigma'$);
- Given a line $L$ and a point $v$ not on it, there is a unique line through $v$ disjoint from $L$ (namely, $W+v$ if $L$ is a coset of $W \in \Sigma'$);
  and
- Each line has exactly $q^m$ points.

There is an obvious notion of parallelism, and by adjoining a new "line at infinity" that "contains" all parallel classes we obtain a projective plane (of order $q^m$). For each $c \in V$ the translation $v \mapsto v+c$ is an automorphism fixing every parallel class. These affine planes (and their associated projective planes) are called *translation planes*. Any isomorphism between two such planes is induced by a semilinear transformation of the underlying vector spaces.

EXAMPLE 4.1. If $V'$ is a 2-dimensional vector space over $\mathrm{GF}(q)$, its set $\Sigma'$ of 1-spaces is a desarguesian spread (cf. Example 3.17), and $\mathbf{A}(\Sigma')$ is called a *desarguesian plane*. This plane is pictured in the following figure. The figure also suggests why it makes no difference whether we view lines as the sets "$y = xs$" or the sets "$y = xs^2$" (i.e., $s \mapsto s^2$ is bijective).

Needless to say, the study of translation planes focusses on nondesarguesian planes. Nevertheless, within the context of this paper desarguesian planes play very special roles: firstly, they produce the original Kerdock codes [**Ke**] (cf. Example 2.4); and secondly, the translation planes constructed via Examples 2.6 and 2.7, as well as those in Theorem 3.5, all are very closely related to desarguesian planes.

*Spreads* ↔ *spread sets.* As in Lemma 3.2, write $V' = X' \oplus Y'$ with $X', Y' \in \Sigma'$, and write matrices with respect to a basis of $V'$ consisting of a basis of $X'$ together with one of $Y'$. Then every member of $\Sigma' - \{Y\}$ can be written uniquely in the form $X'\begin{pmatrix} I & P \\ O & I \end{pmatrix}$ = "$y = xP$" for an $m \times m$ matrix $P$. The set $\mathcal{S}$ of such matrices is essentially what is called a *spread set*, and lets the plane be described in a very familiar manner, using the lines

$$\text{“}x = c\text{'' and “}y = xP + b\text{''},\ b, c \in V',\ \ P \in \mathcal{S}.$$

$\Sigma'$ and $\mathcal{S}$ determine one another in an obvious manner.

*Binary operations* ↔ *spreads.* A binary operation $*$ satisfying conditions **(i)** and **(ii)** in §2 also determines a spread, consisting of $Y'$ and the subsets "$y = x * s$" of $\mathrm{GF}(2^m) \oplus \mathrm{GF}(2^m)$. The lines of the associated affine plane have a familiar appearance:

$$\text{“}x = c\text{'' and “}y = x * s + b\text{''},\ b, c, s \in \mathrm{GF}(2^m).$$

Here, $\mathcal{S}$ consists of matrices of the maps $x \mapsto x * s$. Conversely, starting with a spread $\Sigma'$ in a vector space of characteristic 2, and distinct $X', Y' \in \Sigma'$, choose any bases for $X'$ and $Y'$, and obtain obtain a spread set $\mathcal{S}$ of $m \times m$ matrices as above. Fix an arbitrary bijection $s \mapsto P_s$ of $\mathrm{GF}(2^m) \to \mathcal{S}$ with $P_0 = 0$, and define $x * s := xP_s$ for all $x, y \in \mathrm{GF}(2^m)$. Then conditions **(i)** and **(ii)** are straightforward to check. See [**De**,**§5.1**] for this and additional information concerning these ways of representing translation planes, as well as for a survey of projective planes of this type.

**Symplectic translation planes.** Now that we have seen how to go back and forth between spreads, spread sets, translation planes and binary operations, it is time to see how this works for symplectic spreads $\Sigma'$ and the corresponding planes $\mathbf{A}(\Sigma')$, called *symplectic translation planes.*

EXAMPLE 4.1 *continued.* Starting with a desarguesian spread $\Sigma'$ in $\mathbb{Z}_2^{2m}$, where $m$ is odd, by Corollary 3.3 and Proposition 3.6 we obtain an orthogonal spread $\Sigma$ in $\mathbb{Z}_2^{2m+2}$, and hence a Kerdock set and Kerdock code. This latter Kerdock set just the one in Example 2.4: the one giving rise to the original Kerdock code of length $m+1$.

*Symplectic spreads ↔ symmetric spread sets.* Suppose we start with a symplectic spread $\Sigma'$. Fix distinct $X', Y' \in \Sigma'$, and a basis of $X'$, but this time choose the basis of $Y'$ to be the dual basis (as was also done just before Lemma 3.1). Then the resulting spread set $\mathcal{S}$ consists of symmetric matrices. Conversely, each spread set consisting of symmetric matrices produces a symplectic spread.

*Binary operations ↔ Kerdock sets.* Now suppose that we start with a binary operation $*$ satisfying condition **(v)** in §2. Then the transformations $x \mapsto x * s$ are self–adjoint with respect to the inner product $T(xy)$ on $\mathrm{GF}(2^m)$. In terms of an orthonormal basis of $\mathrm{GF}(2^m)$, this means that $x \mapsto x * s$ is represented by a symmetric matrix $P_s$. Consequently, we obtain a spread set consisting of symmetric matrices, and we have seen that this produces a symplectic spread. Lift $\mathcal{S}$ to a Kerdock set $\mathcal{K}$ using Proposition 3.6. A calculation shows that *this Kerdock set is precisely the one appearing in* Proposition 2.5.

Conversely, starting with a Kerdock set $\mathcal{K}$, pass to the set $\mathcal{S}_{\mathcal{K}}$ given in Proposition 3.6, fix a bijection $s \mapsto P_s$ of $\mathrm{GF}(2^m) \to \mathcal{S}_{\mathcal{K}}$ with $P_0 = 0$, and again define $x * s := xP_s$ for all $x, y \in \mathrm{GF}(2^m)$. Then conditions **(i,ii,iv,v)** are straightforward to check, and **(iii)** can be made to hold by suitably modifying the bijection $s \mapsto P_s$. In other words,

THEOREM 4.2. *Every Kerdock set of* $(m+1) \times (m+1)$ *binary matrices is equivalent to one obtained in Proposition* 2.5 *using some binary operation.*

We have now seen how to go back and forth between various objects:

| | | |
|---|---|---|
| orthogonal spread | ↔ | symplectic spread |
| orthogonal spread | ↔ | Kerdock set |
| Kerdock set | ↔ | symmetric spread set |
| Kerdock set | ↔ | binary operation |

With each object on the left are associated many on the right; with each on the right is associated essentially just one on the left. Here "many" and "one" mean "up to whatever notion of equivalence is appropriate". Choices are made in each case, though this is is most evident in the first two listed instances, where a choice of nonsingular point $z$ or of $Y \in \Sigma$ was explicitly made. In Proposition 3.6 we chose a basis, and even distinguished the last basis vector. In Proposition 2.5 and Theorem 4.2 a specific identification was chosen between an $m+1$–dimensional vector space and $\mathrm{GF}(2^m) \oplus \mathbb{Z}_2$.

EXAMPLE 4.3. Start with a desarguesian symplectic spread, go up and then down (at the end of Section 3). This produces another symplectic spread. The binary operations in Examples 2.6 and 2.7 were obtained in this manner [**Ka1**].

*Isomorphisms between planes.* Each orthogonal spread appears to produce large numbers of symplectic spreads $\Sigma_z$. This leads us to the isomorphism question: when are two planes $\mathbf{A}(\Sigma_z)$ obtained in this manner isomorphic? If there is a symplectic transformation sending one spread to the other, the planes are certainly isomorphic. It seems surprising that the converse is both true and easy to prove:

THEOREM 4.4 [**Ka1**]. *For $i = 1, 2$, let $\Sigma_i$ be a symplectic spread in a $2m$–dimensional symplectic space $V_i$ over $\mathbb{Z}_2$. Let $g\colon \mathbf{A}(\Sigma_1) \to \mathbf{A}(\Sigma_2)$ be an isomorphism that sends the point 0 to the point 0. Then there is an invertible linear transformation $s\colon V_1 \to V_2$ such that the following hold:*

(i) $(\Sigma_1)s = \Sigma_2$,
(ii) *$s$ is an isometry (i.e., $(us, vs) = (u, v)$ for all $u, v \in V_1$), and*
(iii) *$g^{-1}s$ fixes every member of $\Sigma_2$.*

The set of all nonsingular linear transformations fixing every member of $\Sigma_2$ (as in (iii)), together with $O$, is a field. It is the largest field over which the plane can most readily be viewed.

COROLLARY 4.5. *Two translation planes $\mathbf{A}(\Sigma_{z_1})$ and $\mathbf{A}(\Sigma_{z_2})$ arising from the same orthogonal spread $\Sigma$ are isomorphic if and only if $z_1$ and $z_2$ are in the same orbit of the group $G(\Sigma)$ of all orthogonal transformations preserving $\Sigma$.*

Theorem 4.4 also permits the determination of the full automorphism groups of many of these planes with little or no effort. Further information concerning some of these planes is given in [**Ka1**]. For now we merely note that the construction techniques for planes, using Kerdock sets and orthogonal and symplectic spreads, are very flexible. They have produced planes with relatively large automorphism groups [**Ka1**] as well as planes with unexpectedly small automorphism groups [**Ka5;Wi**].

## 5. Further aspects of Kerdock codes

**Nonlinearity.** *Each code $C(\mathcal{K})$ is nonlinear.* This is not at all obvious. What is easy to see is that linearity would be the same as $\mathcal{K}$ being closed under addition, which is not the case when $n > 2$. In fact, a much stronger result is true, in view of the following elegant result of Cameron (cf. [**Ka5**]):

THEOREM 5.1 (Cameron). *Let $W$ be a subspace of the space of all $2r \times 2r$ skew–symmetric matrices over a finite field. If every nonzero member of $W$ is nonsingular, then* $\dim W \leq r$.

PROOF. If $(a_{ij}) \in W$ then $\det(a_{ij}) = \mathrm{Pf}(a_{ij})^2$, where $\mathrm{Pf}(a_{ij})$ is the Pfaffian of $(a_{ij})$ and is a polynomial of degree $r$ in the $a_{ij}$ [**La, p. 373**]. Let $A_1, \dots, A_d$ be a basis of $W$. If $A = \sum_i x_i A_i$ for scalars $x_i$, then $\mathrm{Pf}(A) = f(x_1, \dots, x_d)$ for a polynomial $f$ of degree $r$. By the Chevalley-Warning Theorem [**La, p. 140**], $f$ has more than one zero if $d > r$. Thus, we must have $d \leq r$. □

Extremal subspaces (i.e., of dimension $r$) have yet to be investigated. In particular, it is not known whether there might be interesting examples.

**Strongly regular graphs.** Once again consider a $2n$–dimensional binary vector space $V$ equipped with the quadratic form $Q_{2n}$, where $n \geq 4$ and $n$ is even. There is a natural graph $(S, \not\perp)$ defined on the set $S$ of nonzero singular vectors. This is a strongly regular graph: it is regular of degree $2(2^{n-1}-1)(2^{n-2}+1)$; any two adjacent vertices are adjacent to $1 + 4(2^{n-2}-1)(2^{n-3}+1)$ others; and any two nonadjacent vertices are adjacent to $(2^{n-1}-1)(2^{n-2}+1)$ others.

Any orthogonal spread $\Sigma$ in $V$ also leads to a strongly regular graph having the exact same parameters (i.e., the same constants, associated with adjacent and nonadjacent pairs of vertices, as in the preceding paragraph). Namely, the vertices of this graph are the hyperplanes of the members of $\Sigma$; two vertices $V_1, V_2$ are adjacent if and only if $V_1 \cap V_2^{\perp} = 0$ [**DDT;Ka3**]. This graph is isomorphic to the previous one if $2n = 8$, and probably not if $2n > 8$, but this is open.

There are analogues of these graphs obtained from symplectic spreads, and similar graphs obtained over other fields [**Ka3**].

**Bounds for line–sets in $\mathbb{R}^N$ with prescribed angles.** There is a simple way to embed $\mathbb{Z}_2^N$ into $\mathbb{R}^N$, induced by the obvious isomorphism $\mathbb{Z}_2^N \to \{\pm 1\}^N$. In this manner, a code $C$ of length $N$ produces an example of a set of unit vectors; and extremal properties of sets of unit vectors imply ones for codes. This point of view is somewhat related to that of Sloane in these Proceedings.

*Line–sets from Kerdock codes.* For example, start with any Kerdock set $\mathcal{K}$ of matrices, let $C(\mathcal{K})$ be as in (2.3), and write $N = 2^n$. Then we can form the following unit vectors in $\mathbb{R}^N$ (where coordinates are again indexed by vectors in $\mathbb{Z}_2^n$ and the exponents again are just the Kerdock codewords):

$\frac{1}{2^{n/2}}\left((-1)^{Q_B(v)+s\cdot v+\varepsilon}\right)_{v\in\mathbb{Z}_2^n}$ *where* $B \in \mathcal{K}$, $s \in \mathbb{Z}_2^n$, $\varepsilon \in \mathbb{Z}_2$;
and
*the $N = 2^n$ standard basis vectors and their negatives.*
Total number of vectors: $N^2 + 2N$ in $\mathbb{R}^N$.

Another way to view this is as a set of *lines* in $\mathbb{R}^N$, namely the 1-spaces spanned by the vectors in the above list:

$$\left\langle \left( (-1)^{Q_B(v)+s\cdot v+\varepsilon} \right)_{v\in\mathbb{Z}_2^n} \right\rangle \quad \textit{where } B \in \mathcal{K},\ s\in\mathbb{Z}_2^n,\ \varepsilon\in\mathbb{Z}_2;$$

and

*the 1–spaces spanned by the $N$ standard basis vectors.*

Total number of lines: $(N^2+2N)/2$ in $\mathbb{R}^N$.

The distances between codewords in the Kerdock code imply that *any two of these lines are either perpendicular or are at an angle of* $\cos^{-1} 1/\sqrt{N}$. In fact, these lines fall into $(N+2)/2$ orthonormal frames such that the angle between members of different frames is always $\cos^{-1} 1/\sqrt{N}$. This construction is due to König [**Kö**], based on ideas in [**CS**]. Applications of these line–sets to approximation theory and to isometric embeddings of Euclidean spaces into $\ell_p$-spaces are given in [**Kö**].

One of the starting points of the paper [**CCKS**] was the observation that there is a tantalizing similarity between the construction of this set of lines from $\mathcal{K}$ and the construction orthogonal spreads from $\mathcal{K}$ in Corollary 3.3. Namely, in both situations there is an apparent asymmetry to the description, in which one member of the spread, or one frame (the standard one), appears to be somehow distinguished. In both situations, this asymmetry is merely apparent, caused by an initial choice of basis. If we had chosen one of the other orthonormal frames and written all the others in terms of it, we would have obtained a similar description. This is studied in great detail in [**CCKS**], where it is shown that these $|\mathcal{K}|+1$ orthonormal frames arise in a *natural* way from the $|\mathcal{K}|+1$ members of the orthogonal spread determined by $\mathcal{K}$. Moreover, it is shown how to go back from the line-set to $\mathcal{K}$ using a group (the stabilizer of the line–set in the real orthogonal group).

*More general line–sets.* In general, consider a set $\Omega$ of unit vectors spanning $\mathbb{R}^N$, use the usual dot product in $\mathbb{R}^N$, and assume that $|a\cdot b|\in\{0,\alpha\}$ for all $a\neq b$ in $\Omega$, where $0<\alpha<1$ is a constant: the angles between the lines of $\mathbb{R}^N$ determined by the pairs of distinct members of $\Omega$ take on only two values, one of which is 90° (so $\Omega\cap(-\Omega)=\emptyset$). Delsarte, Goethals and Seidel [**DGS**] proved that $|\Omega|\le\binom{N+2}{3}$ for any $\alpha$. They also showed that $|\Omega|\le\frac{N(N+2)(1-\alpha^2)}{3-(N+2)\alpha^2}$, provided that the denominator is positive.

It is the case of equality here that especially concerns us, where we have an *extremal line–set*. In that case, define a graph on $\Omega$, joining two vectors if they are perpendicular. Then this is a strongly regular graph (cf. [**CCKS**]). The special case $\alpha=1/\sqrt{N}$, $|\Omega|=\frac{N(N+2)(1-\alpha^2)}{3-(N+2)\alpha^2}=N^2(N+2)/2$, arises when the original code is a Kerdock code $C(\mathcal{K})$ of length $N=2^{m+1}$. In that case, $\Omega$

is a union of $(N+2)/2 = 2^m + 1 = |\mathcal{K}| + 1$ orthonormal bases, with vectors in different bases not perpendicular. (N. B.—It is not known whether the extremal case $\alpha = 1/\sqrt{N}$ can *only* arise from a Kerdock set $\mathcal{K}$ as above.)

There are many other extremal results concerning Euclidean line–sets (or sets of vectors) due to Delsarte, Goethals and Seidel [**DGS**] and Levenštein [**Le**], among others. For example, the results in [**Le**] merely make assumptions about the largest value of $|a \cdot b|$ for distinct $a, b \in \Omega$, rather than the exact nature of the set of dot products. The arguments in these papers use classical Jacobi polynomials. Besides being beautiful, the methods have the added advantage of being highly flexible, permitting natural extensions to a variety of different contexts. For example, in §6 we will be concerned with line–sets in complex vector spaces.

## 6. $\mathbb{Z}_4$–codes

A *$\mathbb{Z}_4$–code* of length $N$ is just a subset $C_4$ of $\mathbb{Z}_4^N$; it is *linear* if it is an additive subgroup. While one could use the Hamming metric here, an important discovery in [**CHKSS**] was that the *Lee metric* $d_L(u,v)$ leads to lovely results. This metric is defined as follows. The *Lee weights* of $0,1,2,3 \in \mathbb{Z}_4$ are $0,1,2,1$ respectively, the Lee weight $wt_L(v)$ of $v \in \mathbb{Z}_4^N$ is the integral sum of the Lee weights of its coordinates, and $d_L(u,v) := wt_L(u-v)$.

**Gray map.** Next we need to recall the definition of the Gray map used in [**CHKSS**]. The following figure shows the Gray encoding of the elements of $\mathbb{Z}_4$ (or of the points $1, i, -1, -i$ in the complex plane) as pairs of *binary* digits.

The 2–adic expansion $c = c_1 + 2c_2$ of $c \in \mathbb{Z}_4$ defines maps $c \mapsto c_1$ and $c \mapsto c_2$ from $\mathbb{Z}_4$ to $\mathbb{Z}_2$. Extend these in an obvious way to maps $v \mapsto v_i$ from $\mathbb{Z}_4^N$ to $\mathbb{Z}_2^N$. Then binary codes are obtained from $\mathbb{Z}_4$–codes as their images under the *Gray map* $\varphi : \mathbb{Z}_4^N \to \mathbb{Z}_2^{2N}$ given by

$$v\varphi = (v_2, v_1 + v_2),\ \ v \in \mathbb{Z}_4^N$$

[**CHKSS**]. The key property of this map is the following observation:

Theorem 6.1 [**CHKSS**]. *The Gray map is an isometry* ($\mathbb{Z}_4^N$, Lee metric) $\to$ ($\mathbb{Z}_2^{2N}$, Hamming metric).

There is really more than one Gray map: one can be followed by a permutation of the coordinates of $\mathbb{Z}_2^{2N}$ in order to obtain another one. This is significant in [**CHKSS**] and [**CCKS**] since rearrangements of coordinates are implicitly allowed or needed when Gray maps are used.

*Duality.* Starting with a *linear* code $C_4 \subseteq \mathbb{Z}_4^N$, there is a natural definition of the *dual* linear code $C_4^\perp$, using the usual dot product on $\mathbb{Z}_4^N$. The standard MacWilliams identity [**MS, p. 127**], and the fact that the Gray map is an isometry, produce detailed information concerning the metric properties of $C_4^\perp$ and of its binary image under any Gray map $\varphi$. Namely, define the *Hamming weight enumerator* of a distance–invariant *binary* code $C$ to be the following polynomial in two variables $W$ and $X$:

$$\mathrm{Ham}_C(W,X) = \sum_{c' \in C} W^{N-d_H(c,c')} X^{d_H(c,c')},$$

which is independent of the choice of $c \in C$. The standard MacWilliams identity asserts that, for a *linear binary code* $C$,

$$\mathrm{Ham}_{C^\perp}(W,X) = \frac{1}{|C|}\mathrm{Ham}_C(W+X, W-X).$$

The following variant of this was proved using Theorem 6.1:

THEOREM 6.2 [**CHKSS**]. *If $C$ is a binary code such that $C\varphi^{-1}$ is linear, then the binary codes $C$ and $((C\varphi^{-1})^\perp)\varphi$ are distance–invariant and satisfy*

$$\mathrm{Ham}_{((C\varphi^{-1})^\perp)\varphi}(W,X) = \frac{1}{|C|}\mathrm{Ham}_C(W+X, W-X).$$

$$\begin{tikzcd}
\mathbb{Z}_4\text{–}\mathit{code} \arrow[d, "\text{Gray}"] & C\varphi^{-1} \arrow[r, "\text{dual}"] & (C\varphi^{-1})^\perp \arrow[d, "\varphi"] \\
\mathit{binary\ code} & C \arrow[u, "\varphi^{-1}"] & ((C\varphi^{-1})^\perp)\varphi
\end{tikzcd}$$

We will apply this below to some of the Kerdock codes $C(\mathcal{K})$.

$\mathbb{Z}_4$**–valued quadratic forms.** *Each symmetric $m \times m$ binary matrix $P$ determines a map $F_P\colon \mathbb{Z}_2^m \to \mathbb{Z}_4$.*

The definition of $F_P$ is based on the observation that $\mathbb{Z}_4^m / 2\mathbb{Z}_4^m \cong \mathbb{Z}_2^m$. For each $v \in \mathbb{Z}_2^m$ let $\acute{v} \in \mathbb{Z}_4^m$ project onto $v$ mod 2. View the entries of $P$ as elements 0, 1 of $\mathbb{Z}_4$, and write $F_P(v) := \acute{v} P \acute{v}^t$. In detail, if $P = (p_{ij})$ with $p_{ij} \in \{0,1\}$, and if $\acute{v} = (x_i)$, then

$$F_P(v) = \sum_i p_{ij} x_i^2 + 2\sum_{i<j} p_{ij} x_i x_j.$$

It is easy to see that this is independent of the choice of "lift" $\acute{v} = (x_i)$ of $v$. Of course, we could just choose $\acute{v}$ to have coordinates equal to those of $v$, but with 0 and 1 viewed as elements of $\mathbb{Z}_4$ (as was done for $P$). However, we need to be able to state the following basic property of the $\mathbb{Z}_4$*–valued quadratic form* $F_P$ associated with $P$:

$$F_P(u+v) = F_P(u) + F_P(v) + 2\acute{u}P\acute{v}^t$$

for all $u, v \in \mathbb{Z}_2^n$. This equation should be compared with the similar one (2.1) relating binary quadratic forms and alternating bilinear forms.

$\mathbb{Z}_4$**–Kerdock codes.** By Proposition 3.6, each Kerdock set $\mathcal{K}$ is related to a set $\mathcal{S}_\mathcal{K}$ of symmetric matrices. (N.B.—Note, however that relationship presupposes that a row and column have been specified.) Equation (2.3) suggests that we consider the following subset of $\mathbb{Z}_4^{2^m}$:

$$(6.3) \qquad C_4(\mathcal{S}_\mathcal{K}) := \left\{ \left(F_P(v) + 2\acute{s}\cdot\acute{v} + \varepsilon\right)_{v\in\mathbb{Z}_2^m} \mid P \in \mathcal{S}_\mathcal{K},\, s \in \mathbb{Z}_2^m, \varepsilon \in \mathbb{Z}_4 \right\}.$$

This is called the $\mathbb{Z}_4$*–Kerdock code* associated with $\mathcal{K}$. One of the main results in [**CCKS**] is the following

THEOREM 6.4. *$C(\mathcal{K})$ is the image of $C_4(\mathcal{S}_\mathcal{K})$ under a suitable Gray map.*

In particular, these two codes are isometric when equipped, respectively, with the Hamming and Lee metrics. Note that the Gray map in the theorem depends on the manner in which $\mathcal{S}_\mathcal{K}$ was obtained from $\mathcal{K}$ in Proposition 3.6. Namely, we arbitrarily chose to single out the $n$th row and column; but we could just as well used the $j$th row and column for any $j$. Thus, some care is needed so as not to be mislead by notation. In view of the preceding theorem, $\mathbb{Z}_4$–valued quadratic forms are natural objects. They were first introduced within topology [**Br**].

The preceding theorem may leave the impression that the definition of $C_4(\mathcal{S}_\mathcal{K})$ might first have been calculated by applying $\varphi^{-1}$ to $\mathbb{C}(\mathcal{K})$. This was not the case: $\mathbb{Z}_4$–valued quadratic forms, and the definition of $C_4(\mathcal{S}_\mathcal{K})$, arose by viewing the real and complex representation theory of extraspecial 2–groups from unusual perspectives, guided by real and complex line–sets. However, we will not delve into the group–theoretic aspects of these codes, or into the structure of these line–sets, which were essential ingredients in the discovery of $\mathbb{Z}_4$–Kerdock codes. (cf. [**CCKS**]).

$\mathbb{Z}_4$**–linear Kerdock and Preparata codes.** We have seen that $C(\mathcal{K})$ is never linear (Theorem 5.1). One of the most striking discoveries in [**CHKSS**] was that the $\mathbb{Z}_4$–version $C_4(\mathcal{S}_\mathcal{K})$ of $C(\mathcal{K})$ can be linear. That paper studied this in the case of the original Kerdock code, where $\mathcal{S}_\mathcal{K}$ arises, as in Example 2.4, using the field multiplications $\mathrm{GF}(2^m) \to \mathrm{GF}(2^m)$ (although this code was written in an entirely different manner in that paper). In that case, $\mathcal{S}_\mathcal{K}$ is clearly closed under addition, so that $C_4(\mathcal{S}_\mathcal{K})$ is a linear $\mathbb{Z}_4$–code. This was generalized in [**CCKS**]:

THEOREM 6.5. *If $\mathcal{K}$ is a Kerdock set arising as in Proposition* 2.5 *by means of a binary operation* $*$, *then* $C_4(\mathcal{S}_\mathcal{K})$ *is a linear* $\mathbb{Z}_4$*–code if and only if* $*$ *is* **2–sided** *distributive.*

The next observation in [**CHKSS**] was that, if $C_4(\mathcal{S}_\mathcal{K})$ is linear, then $C_4(\mathcal{S}_\mathcal{K})^\perp$ also is linear, and Theorem 6.2 gives the exact weight distribution of the *nonlinear* binary code $P_m(\mathcal{S}_\mathcal{K})$ of length $2^{m+1}$ that is the Gray image of $C_4(\mathcal{S}_\mathcal{K})^\perp$. The codes $P_m(\mathcal{S}_\mathcal{K})$ are examples of *Preparata codes*: their weight distributions are the same as that of code $P_m$ of length $2^{m+1}$ discovered by Preparata in 1968 [**Pr**] (cf. [**MS**]). The fact that the weight distribution of the latter codes is related to that of (the original) Kerdock codes has been a perplexing fact for many years. The introduction of $\mathbb{Z}_4$–linear codes and the Gray map have narrowed this gap, providing codes in some precise sense dual to Kerdock codes. If $m = 3$ then $P_m(\mathcal{S}_\mathcal{K})$ and $P_m$ are equivalent to the Nordstrom–Robinson code. However, if $m \geq 5$ then $P_m(\mathcal{S}_\mathcal{K})$ has the property that the $\mathbb{Z}_2$–subspace of $\mathbb{Z}_2^{2^{m+1}}$ it spans has vectors of weight 2, which is not the case for the original Preparata codes: the codes $P_m$ and $P_m(\mathcal{S}_\mathcal{K})$ *are never quasi–equivalent if* $m \geq 5$ (proved in [**CHKSS**] for the case studied there, and for any Kerdock set $\mathcal{K}$ in [**CCKS**]). The fact that $P_m(\mathcal{S}_\mathcal{K})$ is a sort of dual of $C_4(\mathcal{K})$ prompted the authors of [**CHKSS**] to "propose that this is the 'correct' way to define these codes" (i.e., codes with Preparata's weight distribution).

*Extremal property.* The importance of Preparata codes (either the original versions $P_m$ or the new ones $P_m(\mathcal{S}_\mathcal{K})$) takes us back to the start of this paper. These codes are extremal in the following sense. They have length $N = 2^{m+1}$, minimum distance 6, and as many codewords as possible subject to these conditions, namely, $2^{N-2(m+1)}$. Moreover, no linear code can be extremal in this sense [**GS**]. Since the size of any linear code is a power of 2, it follows that *any linear code with minimum distance at least* 6 *has at most half as many codewords as Preparata codes of the same length.*

In [**CHKSS**] it is also shown that other nonlinear subcodes of $RM(2, n)$ are Gray images of $\mathbb{Z}_4$–linear codes. This groundbreaking paper has produced an outpouring of further research on $\mathbb{Z}_4$–codes. This has led to the consideration of codes over $\mathbb{Z}_{2^l}$ for all $l$, and even more recently to codes over the ring of 2-adic integers by using classical Hensel lifts.

As already observed in §1, Kerdock and other $\mathbb{Z}_4$–linear codes have the striking properties of being optimal from a combinatorial point of view and yet having linear descriptions that simplify both their study and implementation.

*Equivalence.* Two $\mathbb{Z}_4$–codes are *equivalent* if one can be gotten from the other by a permutation of coordinates followed by multiplication by a single diagonal matrix of ±1's. Two $\mathbb{Z}_4$–codes of length $N$ are *quasi–equivalent* if one is equivalent to a $\mathbb{Z}_4^N$–translate of the other.

Equivalences among the codes $C_4(\mathcal{S}_\mathcal{K})$ and among the codes $P_m(\mathcal{S}_\mathcal{K})$ are discussed at length in [**CCKS**]. The results are similar to Theorem 3.4. It is also shown that, when $m$ is composite, there is a $\mathbb{Z}_4$–linear Kerdock code $C_4(\mathcal{S}_K)$ and a Preparata code $P_m(\mathcal{S}_K)$ not quasi–equivalent to the ones studied in [**CHKSS**]. More recently, Williams proved the following far stronger result:

THEOREM 6.6 [**Wi**]. *If $m$ is odd and has $r \geq 2$ prime factors, at least one of which is $\geq 7$, then there are at least $2^{(r-1)m}/m$ pairwise quasi–inequivalent $\mathbb{Z}_4$–linear Kerdock and Preparata codes of length $2^{m+1}$.*

These codes are constructed as follows. There is a sequence $\mathrm{GF}(2^m) \supset F_1 \supset \cdots \supset F_{r-1} \supset \mathrm{GF}(2)$ of fields. For each $i \geq 1$ let $T_i\colon \mathrm{GF}(2^m) \to F_i$ be the trace map and let $\zeta_i \in F_i$. Williams greatly generalized Example 2.6 by repeated use of the up and down process at the end of Section 3. He showed that the following 2–sided distributive binary operation on $\mathrm{GF}(2^m)$ satisfies **(i-v)**:

$$x * s = xs^2 + \sum_1^{r-1} \Big( T_i(\zeta_i x)s + \zeta_i T_i(xs) \Big).$$

He then handled the quasi–equivalence of all of the resulting $\mathbb{Z}_4$–linear Kerdock codes (obtained using Propositions 2.5 and 3.6 together with (6.3)) in order to prove Theorem 6.6. Planes were a crucial tool in this, using Theorem 4.4. Thus, planes enter not only into the construction of the codes $\mathbb{C}(\mathcal{K})$, $\mathbb{C}_4(\mathcal{S}_K)$ and $P_m(\mathcal{S}_K)$, but also into the study of their structure.

**Bounds for line–sets in $\mathbb{C}^{N'}$ with prescribed angles.** As in the case of the usual Kerdock codes, the $\mathbb{Z}_4$–Kerdock codes produce line–sets in $\mathbb{C}^{N'}$ via the isomorphism $\mathbb{Z}_4^{N'} \cong \langle i \rangle^{N'}$, where $N' = 2^m$. Thus, the $\mathbb{Z}_4$–Kerdock code (6.3) produces the following set of lines of $\mathbb{C}^{N'}$ (where the exponents are just the $\mathbb{Z}_4$–Kerdock codewords):

$\Big\langle \big( i^{F_P(v)+2\hat{s}\cdot\hat{v}+\varepsilon} \big)_{v \in \mathbb{Z}_2^m} \Big\rangle$ *where* $P \in \mathcal{S}_K, s \in \mathbb{Z}_2^m, \varepsilon \in \mathbb{Z}_4$;
and
*the* 1*–spaces spanned by the* $N' = 2^m$ *standard basis vectors.*
Total number of lines: $2N'^2 + 2N'$ in $\mathbb{C}^{N'}$.

This time the distances between codewords in the $\mathbb{Z}_4$–Kerdock code imply that *any two of these lines are either perpendicular or are at an angle of* $\cos^{-1} 1/\sqrt{N'}$. These lines fall into $N' + 1$ orthonormal frames such that the angle between members of different frames is always $\cos^{-1} 1/\sqrt{N'}$. One can pass back and forth between these line-sets and those in $\mathbb{R}^{2^{m+1}}$ discussed earlier, each complex line–set producing an essentially unique real one but each real one producing many inequivalent complex ones [**CCKS**].

Once again there are general bounds on complex line–sets in [**DGS**] and [**Le**]; strongly regular graphs arise in suitable cases of equality; and the line–sets obtained from $\mathbb{Z}_4$–Kerdock codes are examples of *extremal line–sets* (cf.

[**CCKS**]). Once again there are also applications to approximation theory and to isometric embeddings [**Ko**].

*Quaternionic codes.* These ideas are pursued slightly further in [**Ka7**], where *quaternionic* line–sets are examined. Extremal ones are constructed in quaternionic space $\mathbb{H}^{N''}$, where $N'' = 2^{m-1}$. These arise, in turn, from quaternionic Kerdock codes of length $N''$, which are suitable subsets of $Q_8^{N''}$ obtained using Kerdock sets of $m+1 \times m+1$ matrices.

In general, a *quaternionic code* of length $N$ is simply a subset of $Q_8^N$. The study of such codes suffers from various apparent disadvantages. There is no way to convert to additive notation, and in particular no reasonable way to introduce a ring structure on the coordinates, as in the cases of codes over $\mathbb{Z}_2$ or $\mathbb{Z}_4$. Thus, they also lack one of the most basic aspects of codes over $\mathbb{Z}_2$ or $\mathbb{Z}_4$, the notion of a dual code: there is no natural inner product on subsets of $Q_8^N$ that generalizes the dot product on $\mathbb{Z}_2^N$ or $\mathbb{Z}_4^N$, since only one binary operation is available. Therefore, it is not surprising that, as yet, there are almost no results concerning quaternionic codes.

## 7. Further directions

In §5 we discussed subsets of $\mathbb{R}^N$ arising from binary codes $C$, obtained by replacing $\mathbb{Z}_2$ by $\{\pm 1\}$. In a similar manner, subsets of $\mathbb{C}^N$ arise from $\mathbb{Z}_4$–codes by replacing $\mathbb{Z}_4$ by $\{\pm 1, \pm i\} = \langle i \rangle$. The obvious metric in $\mathbb{C}^N$ is the one induced by the usual hermitian inner product. The natural metrics on $\mathbb{R}^N$ and $\mathbb{C}^N$ are also natural within coding theory:

- the Hamming metric on $\mathbb{Z}_2^N \equiv \{\pm 1\}^N$ is half of the square of the Euclidean metric restricted to the subset $\{\pm 1\}^N$ of $\mathbb{R}^N$; and
- the Lee metric on $\mathbb{Z}_4^N \equiv \langle i \rangle^N$ is half of the square of the hermitian metric restricted to the subset $\langle i \rangle^N$ of $\mathbb{C}^N$.

These statements are entirely elementary to check. Nevertheless, they suggested in [**Ka7**] that

- "the natural metric" on $Q_8^N$ is half of the square of the hermitian metric restricted to the subset $Q_8^N$ of $\mathbb{H}^N$.

(Note that this "hamiltonian metric" restricts to the Lee metric on $\langle i \rangle^N$.) A tentative discussion of this can be found in that reference.

We have merely hinted at group–theoretic aspects of the subject of this paper; see [**Ka1;CCKS**]. A connection with simple complex Lie algebras is surveyed in [**Ka6**].

We also have not spent much time discussing the affine planes $\mathbf{A}(\Sigma_z)$.

The methods in [**CHKSS**] involve the use of a ring of size $4^m$ in place of $\mathrm{GF}(2^m)$. This leads into the realm of cyclic codes over $\mathbb{Z}_4$. This fundamental new direction in coding theory has not been dealt with at all in the present survey.

## References

[Br] E. H. Brown, *Generalizations of Kervaire's invariant,* Annals of Math. **95** (1972), 368–383.

[CCKS] A. R. Calderbank, P. J. Cameron, W. M. Kantor and J. J. Seidel, *$\mathbb{Z}_4$-Kerdock codes, orthogonal spreads, and extremal Euclidean line-sets* (submitted).

[CHKSS] A. R. Calderbank, A. R. Hammons, Jr., P. V. Kumar, N. J. A. Sloane, and P. Solé, *The $\mathbb{Z}_4$-linearity of Kerdock, Preparata, Goethals and related codes,* IEEE Trans. Inform. Theory **40** (1994), 301–319.

[CL] P. J. Cameron and J. H. van Lint, *Designs, graphs, codes and their links,* Cambridge Univ. Press, Cambridge, 1991.

[DDT] F. DeClerck, R. H. Dye and J. A. Thas, *An infinite class of partial geometries associated with the hyperbolic quadric in $PG(4n-1,2)$,* Europ. J. Combinatorics **1** (1980), 323–326.

[DGS] P. Delsarte, J. M. Goethals and J. J. Seidel, *Bounds for systems of lines and Jacobi polynomials,* Philips Res. Repts. **30** (1975), 91–105.

[De] P. Dembowski, *Finite geometries,* Springer, Berlin-Heidelberg-New York, 1968.

[Dic] L. E. Dickson, *Linear groups, with an exposition of Galois theory,* Teubner, Leipzig 1901; reprint Dover, New York, 1958.

[Dil] J. F. Dillon, *Elementary Hadamard difference sets.* Ph.D. thesis, U. of Maryland, 1974.

[GS] J. M. Goethals and S. L. Snover, *Nearly perfect binary codes,* Discrete Math. **3** (1972), 65–88.

[Ka1] W. M. Kantor, *Spreads, translation planes and Kerdock sets.* I,II, SIAM J. Alg. Discr. Meth. **3** (1982), 151–165 and 308–318.

[Ka2] ———, *An exponential number of generalized Kerdock codes,* Inform. Control **53** (1982), 74–80.

[Ka3] ———, *Strongly regular graphs defined by spreads,* Israel J. Math **41** (1982), 298–312.

[Ka4] ———, *On the inequivalence of generalized Preparata codes,* IEEE Trans. Inform. Theory **29** (1983), 345–348.

[Ka5] ———, *Projective planes of order $q$ whose collineation groups have order $q^2$,* J. Algebraic Combinatorics **3** (1994), 405-425.

[Ka6] ———, *Some Lie algebras, finite groups and finite geometries* (to appear in: Proc. Ohio State U. Groups and Geometries Conference).

[Ka7] ———, *Quaternionic line-sets and quaternionic Kerdock codes* (submitted) .

[KW] W. M. Kantor and M. E. Williams, *New flag-transitive affine planes of even order* (to appear in JCT(A)).

[Ke] A. M. Kerdock, *A class of low-rate nonlinear binary codes,* Inform. Control **20** (1972), 182–187.

[Kö] H. König, *Isometric embeddings of euclidean spaces into finite-dimensional $\ell_p$-spaces* (manuscript).

[Le] V. I. Levenštein, *Bounds on the maximal cardinality of a code with bounded modulus of the inner product,* Soviet Math. Dokl. **25** (1982), 526–531.

[Li] J. H. van Lint, *Kerdock and Preparata codes,* Cong. Numerantium **39** (1983), 25–41.

[MS] F. J. MacWilliams and N. J. A. Sloane, *The theory of error-correcting codes,* North–Holland, Amsterdam, 1977.

[Pr] F. P. Preparata, *A class of optimum nonlinear double-error correcting codes,* Inform. Control. **13** (1968), 378–400.

[Wi] M. E. Williams, *$\mathbb{Z}_4$-linear Kerdock codes, orthogonal geometries, and non-associative division algebras,* Ph.D. thesis, University of Oregon 1995.

Department of Mathematics, University of Oregon, Eugene, OR 97403

*E-mail address*: kantor@math.uoregon.edu