Connections between Linear Complementary Dual Codes, Permanents and Geometry

Adel N Alahmadi, Husain S Alhazmi, Hatoon Shoaib, David G Glynn, Saeed Ur Rehman, Patrick Solé

mathematics

MDPI

Article

Connections between Linear Complementary Dual Codes, Permanents and Geometry

Adel N. Alahmadi [1,*], Husain S. Alhazmi [1], Hatoon Shoaib [1], David G. Glynn [2], Saeed Ur Rehman [2] and Patrick Solé [3]

[1] Research Group of Algebraic Structures and Applications (ASA), Mathematics Department, Faculty of Science, King Abdulaziz University, Jeddah 21589, Saudi Arabia; hsalhazmi@kau.edu.sa
[2] College of Science and Engineering, Flinders University, G.P.O. Box 2100, Tonsley, SA 5001, Australia; saeed.rehman@flinders.edu.au
[3] 12M Lab, (Centrale Marseille, CNRS, Aix-Marseille University), 13288 Marseilles, France; sole@enst.fr
* Correspondence: analahmadi@kau.edu.sa

Abstract: Linear codes with complementary duals, or LCD codes, have recently been applied to side-channel and fault injection attack-resistant cryptographic countermeasures. We explain that over characteristic two fields, they exist whenever the permanent of any generator matrix is non-zero. Alternatively, in the binary case, the matroid represented by the columns of the matrix has an odd number of bases. We explain how Grassmannian varieties as well as linear and quadratic complexes are connected with LCD codes. Accessing the classification of polarities, we relate the binary LCD codes of dimension k to the two kinds of symmetric non-singular binary matrices, to certain truncated Reed–Muller codes, and to the geometric codes of planes in finite projective space via the self-orthogonal codes of dimension k.

Keywords: error-correcting code; invariant; rectangular matrix; permanent; complementary code; geometric code; Reed–Muller; Grassmannian; polarity; self-orthogonal; matroid

MSC: 05B35; 11C20; 11T71; 14D10; 14G15; 14G17; 14L24; 14J20; 15A15; 15A72; 94A60; 94B05; 94B27

Citation: Alahmadi, A.N.; Alhazmi, H.S.; Shoaib, H.; Glynn, D.G.; Rehman, S.U.; Solé, P. Connections between Linear Complementary Dual Codes, Permanents and Geometry. *Mathematics* **2023**, *11*, 2774. https://doi.org/10.3390/math11122774

Received: 18 April 2023
Revised: 16 May 2023
Accepted: 23 May 2023
Published: 20 June 2023

1. Introduction and Some Definitions

We assume standard details to be found in textbooks [1,2] about codes over finite fields, their duals, generator matrices and so on. A subspace of dimension k of a vector space (with a fixed coordinate system) over field F is called a *code* with parameters $[n,k]$. The starting point here is a kind of linear code called a *Linear Complementary Dual* (LCD) code [3–7]. These have been used recently in applications involving protection schemes against side-channel and fault injection attacks [8–10].

Let C be a subspace of dimension k of a vector space of dimension n (C is a linear code). A *complement* D of C (or complementary subspace to C) is a subspace of the complementary dimension $n-k$ such that both subspaces together generate the whole space.

The Grassmann dimension rule for any subspaces X and Y is that

$$\dim(X)+\dim(Y)=\dim(X\cap Y)+\dim(\langle X,Y\rangle),$$

where $\langle X,Y\rangle$ denotes the smallest linear space containing both X and Y.

Thus, for C of dimension k and for any subspace D of dimension $n-k$, we have $k+n-k=n=\dim(C\cap D)+\dim(\langle C,D\rangle)$ showing that $\langle C,D\rangle=F^n$ if and only if $C\cap D=\{0\}$.

A linear code C over a field F with parameters $[n,k]$ has the LCD property if its dual code $C^\perp$ is a complement of C in the vector space F^n. Since $\dim(C^\perp)=n-k$, we have an equivalent condition that the only word of C orthogonal to all the words of C is the zero

word. Note that for the real field $F = \mathbb{R}$, every code is LCD, but for finite fields $GF(q)$ this is false.

Before discussing the various algebraic and geometric ideas related to LCD codes, we shall clarify the definitions; for example, the term "invariant" has various meanings depending on the situation.

A square matrix with a single non-zero entry in every row and every column with all the other entries as 0 is a *generalised permutation matrix GPM*; it is, more precisely, a *permutation matrix PM* if the non-zero entries are 1.

Suppose $A = (a_{ij})$ is a $k \times n$ matrix over F. Let $d \in \mathbb{Z}^+$. As usual, $|B|$ denotes the determinant of a square matrix B over F. Consider the following conditions on a polynomial p in the entries a_{ij} of A:

1. $p(BA) = |B|^d p(A)$ for any $k \times k$ matrix B;
2. $p(AC) = \pm p(A).|C|$ for any *GPM* C of size $n \times n$;
3. $p(AC) = \pm p(A).|C|$ for any *PM* C of size $n \times n$.

A polynomial function $p(A)$ homogeneous of degree kd in the kn variables a_{ij} may be known as the following:

- *invariant* if properties 1 and 2 hold;
- *quasi-invariant* if properties 1 and 3 hold.

It is clear that if $k = n$, there is only one kind of invariant that is $|A|$. For non-square matrices, there are quasi-invariants that are constructed, as in [11], from formulae involving subdeterminants of the matrix A. One of the authors in [12], for the cases $n = k(q-1)$, $d = q-1$, q being any prime power, constructed invariants X_q for prime characteristic fields. We shall see that there is a quasi-invariant that determines the binary LCD codes.

Another kind of (non-matrix) invariant is that of Tutte–Grothendieck in the theory of matroids. Here, the Tutte (bivariate) polynomial is connected with the binary LCD codes.

The well-known Reed–Muller codes are certain linear codes defined from polynomial functions [1,2]. In Section 2, we relate the $RM(k-3,k)$ binary codes to the self-orthogonal codes that satisfy $C \subseteq C^{\perp}$. There are also finite geometry codes. Given a fixed dimension d of the subspace of a finite projective space $PG(k-1,q)$ over the field $GF(q)$, there is a linear code called the *subspace* code. This has a length of the number of points $(q^k-1)/(q-1)$ of $PG(k-1,q)$ and words that are generated by the characteristic functions of the subspaces of dimension d. The dual code to this is called the *geometric* code for dimension d of subspace. It is the space of functions from the point-set to $GF(q)$ such that the sum of the values on any subspace of dimension d is zero.

The material is arranged as follows. Section 2 relates the theory of binary LCD codes to the permanent function. Section 3 considers Grassmannian varieties, linear and quadratic complexes and how they are related to LCD codes, especially looking at the binary case. Section 4 discusses constructions of LCD codes and the number of them. It uses the counts of polarities in binary space and the numbers of self-orthogonal or certain Reed–Muller codes. At the end, we discuss the Tutte invariant of matroids and how it is used to count bases in the matroid, which, in the binary case, evaluates the permanent of the corresponding binary matrix. Section 5 summarises this note.

2. Quasi-Invariants of Matrices

We recall a result characterising LCD codes in a simple algebraic way:

Theorem 1 ([4,8]). *C is LCD if and only if any generator matrix of C satisfies $|AA^t| \neq 0$.*

Note that this was restated and proved [8] in a different way using a generator matrix of standard form $A = (IM)$. The point is that the condition of Theorem 1 makes the polynomial $|AA^t|$ a quasi-invariant (with the degree $d = 2$ in the quasi-invariant definition

of Section 1). Thus, we expect that there should be an alternative formula for it using subdeterminants. We are led to the following:

Theorem 2. *If A is a $k \times n$ matrix over a field (with $k \leq n$), then*

$$|AA^t| = \Sigma_B \, |B|^2,$$

where the sum is over all $k \times k$ submatrices B of A.

Proof. This is a special case of the Cauchy–Binet formula (published in Paris, 1813), sometimes called the Lagrange identity, for the determinant of the product AB^t of two (rectangular) matrices of size $k \times n$. Let us provide a brief proof.

We use induction on (k, n) starting with the case $k = 1$ and any $n \geq 1$. In that case, $A = (a_1, \ldots, a_n)$ and $|AA^t| = \Sigma a_i^2$, and the formula holds. It is also clear for $n \leq k$ because the rank of A must be k requiring $n \geq k$ if $|AA^t| \neq 0$, and, if $n = k$, then $|AA^t| = |A|^2$.

Assume now that the formula is true for values at most k and n. Then, consider any $k \times (n + 1)$ matrix A. If the final column of A is all zeros, then the formula will hold by looking at the smaller matrix with that column deleted. If the final column is non-zero, by multiplying A on the left with a non-singular matrix B, we can modify A so that the final column is the unit vector e_1. We can check that the quasi-invariant is multiplied by $|B|^2$. It can also be calculated (by induction) from the matrix obtained by deleting the last column and also by deleting both the last column and the first row (the "minor"). Hence, the induction can be shown by reducing it to smaller cases. □

Corollary 1. *If A is a $k \times n$ matrix over a field of characteristic two (e.g., the binary field $\mathbb{Z}_2$), then*

$$|AA^t| = (\Sigma_B \, |B|)^2,$$

where the sum is over all $k \times k$ submatrices B of A.

Recall that the *permanent* function can be generalised to any non-square matrix A. Assuming that the number of rows is at most the number of columns, i.e., $k \leq n$, it is sum over all products of elements of A, one in each row and one in k of the columns. In general, the permanent is difficult to calculate, and, in the worst-case scenario, the time blows out exponentially with the size of the matrix. There are formulae which improve the efficiency of calculation such as those in [13], but these are not polynomial time.

Theorem 3. *The permanent function $per(A)$ of any matrix over a field of characteristic two (non-square or not) is a quasi-invariant that has the formula*

$$per(A) = \sqrt{|AA^t|}.$$

Proof. This follows directly from Corollary 1. Even if the Frobenius mapping ($x \mapsto x^2$) were not an automorphism (as the square root function is not well defined everywhere in the case of an imperfect field of characteristic two), from the formula, the square root of $|AA^t|$ will always be unique and in the same field generated by the matrix elements. □

The square root is very easy to calculate for finite fields $GF(2^h)$ because it is the inverse of the Frobenius automorphism x^2, and so $\sqrt{x} = x^{2^{h-1}}$, which is a polynomial function. For the binary field (matrices of zeros and ones modulo 2), the square root can be omitted from the formula.

3. Grassmannian Varieties and Geometrical Considerations

Here, we recall some classical geometrical ideas and use them to understand the LCD codes better. In particular, we can enumerate LCD codes in a certain way. The main

reference we use here is ([14], §7), but one should be aware of some differences with fields of prime characteristic, especially when quadrics (degree 2 hypersurfaces) are considered. However, the main results for Grassmannian and other concepts such as null polarities and linear complexes are usually identical.

As established originally by Grassmann (we refer to the elder H.G. Grassman, not his son H.E. Grassman who was also a high-level mathematician) in the 19th century, one can coordinatise subspaces of a particular dimension $k-1$ in projective space $PG(n-1,F)$, the projective space of dimension $n-1$ over the (commutative) field F, using Grassmann coordinates. We are reminded that Grassmann was one of the first to investigate higher-dimensional geometry in a systematic way. The Grassmann coordinates generalise the Plücker coordinates for lines in 3-dimensional space that were discovered earlier in that century.

We consider a set of k independent points (a basis) in $PG(n-1,F)$ for the particular subspace of dimension $k-1$. Then, we form a $k \times n$ matrix A with the coordinates of these points as the k rows, and then the Grassmann coordinates for the subspace is the list (in a particular fixed order) of $k \times k$ subdeterminants of A. Thus, the number of coordinates is $\binom{n}{k}$. Note that the choice of basis for the subspace does not affect the Grassmann coordinates since multiplying on the left of A by a non-singular matrix B will only multiply the Grassmann coordinates by $|B|$, and Grassmann coordinates are homogeneous. The same is true for non-zero multiples.

There are many properties of these coordinates for subspaces, including formulae for dual subspaces, for how subspaces intersect, and the quadratic relations between the coordinates. Suffice it to say that the set of Grassmann coordinates, a "Grassmannian variety", is an algebraic variety that is the precise intersection of a number of quadrics in the space $PG(\binom{n}{k}-1,F)$. The smallest non-trivial example is the Klein quadric $ab-cd+ef=0$ in $PG((5,F)$ with point coordinates (a,b,c,d,e,f), which is the Grassmann coordinate for the lines in 3D space. In particular, if p_{ij} denotes the 2×2 subdeterminant with columns $i \neq j$, then $a=p_{01}, b=p_{23}, c=p_{02}, d=p_{13}, e=p_{03}, f=p_{12}$. Note that the dimension of the space of lines in $PG(3,F)$ is four, and this is the same as the dimension of the Klein quadric (or any hypersurface) in $PG(5,F)$.

Theorem 4. *An $[n,k]$-code over a field F is LCD if and only if as a subspace its Grassmann coordinates (g_i) satisfy $\Sigma_i g_i^2 \neq 0$; that is, the norm of the Grassmann coordinates' vector is non-zero.*

Proof. This follows directly from Theorem 2 and the definition of Grassmann coordinates. □

Given that, in general, the Grassmann coordinates satisfy many quadratic conditions we have the following:

Corollary 2. *There are many possible equivalent quadratic conditions on the Grassmann coordinates to determine an LCD code $[n,k]$ (when $4 \leq k+2 \leq n$).*

For example, in the case $[n,k]=[4,2]$ above, $a^2+b^2+c^2+d^2+e^2+f^2+\lambda(ab-cd+ef) \neq 0$ gives the space of all quadratic conditions. In the characteristic two case, in particular when F is the binary field $GF(2)$, the condition for LCD is that the sum of the Grassmann coordinates is non-zero, a linear condition. In that case, we can characterise the LCD codes as the points in the complement of the Grassmannian with a certain hyperplane.

In the case $[n,k]=[n,2]$, the intersection of the Grassmannian of lines of $PG(n-1,F)$ with any hyperplane corresponds, in general, to a "linear complex" of $PG(n-1,F)$. This is because any linear complex is just a set of lines satisfying a single non-trivial (homogeneous) linear condition on the Grassmann coordinates of the lines. Then, it is known [14] that a linear complex determines and is determined by a null polarity: two points are conjugate under the polarity if and only if the line joining them is in the complex. *Conjugate* means that one point is on the image (a hyperplane) of the other point under the polarity. Note that *null* means that every point is self-conjugate. The image of any point under the null

polarity will be the hyperplane that is the union of all the lines of the linear complex passing through that point.

Often, we substitute the word *symplectic* instead of "null" since symplectic groups can be defined as the non-singular linear transformations that commute with a null polarity.

Theorem 5. *In the characteristic two field case, non-LCD codes* $[n, 2]$ *correspond to lines of a linear complex of* $PG(n-1, F)$ *and a null polarity. If n is even, then the linear complex and null polarity are non-degenerate, but, if n is odd, they are degenerate.*

Proof. The condition of Corollary 1 that the sum of the 2×2 subdeterminants in the $2 \times n$ generator matrix for C is zero is a linear condition in the Grassmann coordinates for the lines of $PG(n-1, F)$. This translates ([14], §XI, p. 380) to the null polarity with matrix $J_n - I_n$, where J_n is the all ones matrix and I_n is the identity matrix of size n. By using row reductions (subtracting the first row from all the others) then adding the first column to all the other columns, we obtain an upper triangular matrix with a main diagonal of $n-1, -1, -1, \ldots$. Thus, $|J_n - I_n| = (n-1)(-1)^{n-1} \equiv 0 \pmod 2$ if n is odd, and 1 if n is even. Another way to execute this is to calculate

$$(J_n - I_n)^2 = (n-2)J_n + I_n,$$

so that modulo two $J_n - I_n$ is self-inverse $(J_n - I_n)^{-1} = J_n - I_n$, invertible, for even n. In addition, nilpotent, $(J_n - I_n)^2 = J_n - I_n$, not invertible, for odd n. Note that in the odd n case, the rank of $J_n - I_n$ is $n-1$ because it contains a non-singular $J_{n-1} - I_{n-1}$ principal submatrix so that the corresponding linear complex and null polarity are minimally degenerate. □

Since a line of $PG(n-1, F)$ will be either in the linear complex or not, we can count the number of LCD codes $[n, 2]$. Let us demonstrate this for the binary case, $F = GF(2)$. Suppose n is even and we have a non-degenerate null polarity α of $PG(n-1, 2)$. Given a point P, its image P^α will be a hyperplane of $PG(n-1, 2)$. Then, any point $Q \in P^\alpha \setminus \{P\}$ is conjugate to P, so every line through P in P^α is in the linear complex corresponding to α. There are $(2^{n-2} - 1)$ such lines. By counting point/line incidences in two ways, we obtain the number of lines of the linear complex as $(2^n - 1)(2^{n-2} - 1)/3$. Hence:

Theorem 6. *In the case n is even, the number of LCD codes (lines not in the linear complex) is*

$$(2^n - 1)(2^n - 2)/6 - (2^n - 1)(2^{n-2} - 1)/3 = (2^n - 1).2^{n-2}/3.$$

These counts are consistent with the formulae in [6].

Corollary 3. *For* $n = 2$, *there is a single binary LCD code* $[2, 2]$, *which is generated by* $(1, 0)$ *and* $(0, 1)$. *For* $n = 4$, *there are* $(2^4 - 1)2^2/3 = 20$ *LCD codes with parameters* $[4, 2]$.

For instance, the [4, 2] codes with generators $(1, 0, 1, 1)$ and $(0, 1, 1, 1)$ are LCD.

A *projective* code is a linear code which has a dual distance of at least three. Equivalently, any generator matrix of such a code has no zero columns or columns that are multiples of one another. For example, the longest binary projective LCD code $[14, 4]$ will be obtained by puncturing the simplex code $[15, 4]$ at any column so that one could consider a 4×14 binary matrix as a generator matrix with all the possible non-zero columns except for $(1, \ldots, 1)^t$. We shall learn more about this in the next section.

4. Constructions and Numbers of Projective Binary LCD Codes

Since the main applications have, so far, been in the binary case, we will continue to consider binary LCD codes here. Suppose C is an LCD code over a characteristic two field with generator matrix A. We can assume that A has no zero or repeated columns because of the following:

Lemma 1. *If a zero column is deleted from or appended to the matrix A, or if a pair of identical binary columns are deleted from or appended to it, then the product AA^t is unchanged.*

Proof. The first part is elementary, and the second part uses the fact that $2 = 0$ in the field. □

Once we assume the simplification of non-zero or repeated columns in a generator matrix, and also the equivalence of codes under permutations of the element positions, we can make a geometrical leap and consider sets of points in projective space $PG(k-1,2)$ of dimension $k-1$ over the finite field GF(2). Every point of $PG(k-1,2)$ has homogeneous coordinates, a non-zero vector in $GF(2)^k$. A subset of n points of $PG(k-1,2)$ then corresponds to a $k \times n$ matrix A. There are no zero or repeated columns in this matrix. Thus, it is the generator matrix for some binary code of length n that has a dimension of at most k. The property of this code being LCD, self-orthogonal or similar is an inherent property of the set of points, as is the code's distance. Equivalent codes will be related by non-singular (row) transformations; that is, collineations of the space $PG(k-1,2)$.

When considering LCD codes for which AA^t is non-singular, the product AA^t is important. It can be any symmetric binary $k \times k$ matrix. Any set S of n points of $PG(k-1,2)$ gives a $k \times n$ matrix $A(S)$ with the columns as point coordinates of the set. This, in turn, gives the symmetric binary matrix $A(S)A(S)^t$. If its determinant is 1, then the code generated by the rows of A is LCD.

Suppose we have two sets of points S and T such that their corresponding symmetric matrices $A(S)A(S)^t = A(T)A(T)^t$. Then, the symmetric difference $S\Delta T$ will be another set of points with $A(S\Delta T).A(S\Delta T)^t = A(S).A(S)^t + A(T)A(T)^t = 0$. Note that by Lemma 1, columns which are the same in both matrices do not contribute because they are repeated in the join of the two matrices and are deleted in the symmetric difference. Thus, $A(S\Delta T)$ is a generator matrix for a self-orthogonal code.

We need the following result of MacWilliams [15].

Lemma 2. *The number of non-singular $k \times k$ symmetric matrices over $GF(2)$ is*

$$\prod_{i=1}^{k/2} (2^{k+1} - 2^{2i}) \text{ if } k \text{ is even, and}$$

$$\prod_{i=0}^{(k-1)/2} (2^k - 2^{2i}) \text{ if } k \text{ is odd.}$$

Note that enlarging the size of the matrix by one (from an even size to an odd size k) multiplies the number of non-singular matrices by the number $2^k - 1$ (in the term with $i = 0$) of non-zero vectors in the larger space.

There is an interesting relationship between self-orthogonal binary codes (all the codewords are pairwise orthogonal) and certain punctured Reed–Muller codes. A binary Reed–Muller code $RM(i,k)$ has parameters $[2^k, \binom{k}{0} + \cdots + \binom{k}{i}]$, where each position in a word corresponds to a vector in $GF(2)^2$), and each word corresponds to evaluations of a polynomial in k variables of a degree at most i over $GF(2)$. We consider punctured (projective) codes so that evaluations at the zero vector are neglected, trimming the code length $2^k - 1$.

Theorem 7. *Every $k \times n$ matrix A, $3 \leq k$ with no zero or repeated columns that satisfy $AA^t = 0$ (the rows generating a self-orthogonal code) can be formed from a set of planes under symmetric difference (giving a set of points) in the projective space $PG(k-1,2)$. Each symmetric difference corresponds to a unique word of the punctured Reed–Muller code $RM(k-3,k)$.*

Proof. Note that each plane consists of seven points, and if we write these points as the columns of a 3×7 generator matrix, the code is the simplex code $[7,3,4]$, which is self-

orthogonal, contained in its dual, the Hamming $[7,4,3]$ code. In the general case, we form a generator matrix G for the simplex code $[2^k-1,k]$, which is self-orthogonal for $k \geq 3$. We then append it to $\binom{k}{2}$ rows, each row corresponding to a polynomial $x_i x_j$, $1 \leq i < j \leq k$ and, for a column $c_1, \ldots, c_k$ of G, we insert the value $c_i c_j$ into the row with label $x_i x_j$. This makes a $(k + \binom{k}{2}) \times (2^k - 1)$ binary matrix H. This method of creating longer vectors is called Veronese mapping in geometry, so the columns of H are coordinates of points on a certain Veronesean variety. If we added the entire ones row to H, we would have a generator matrix for the quadratic punctured $RM(2,k)$ code. However, we do not follow that. We consider the dual code $C(H)^\perp$ to the code $C(H)$ generated by H.

If v is any word of $C(H)^\perp$ of weight n, it is easily checked that the corresponding columns where the ones are form a submatrix A of H that is self-orthogonal. What, however, is $C(H)^\perp$? It is the punctured $RM(k-3,k)$ made from polynomial functions in k variables of a degree at most $k-3$ (including the constant 1 function of degree 0).

The connection with planes of $P(k-1,2)$ follows from the Fundamental Theorem of Geometric Codes as found in [16,17]. In the binary case, the characteristic functions of the planes of the projective space $PG(k-1,2)$ generate a binary code, and a basis for this code is precisely the set of all functions from $GF(2)^k$ to $GF(2)$ given by the mappings $(x_1, \ldots, x_k) \mapsto p(x_1, \ldots, x_k)$, a polynomial of degree at most $k-3$ in the k variables. This code is called the subspace code $C_2^\perp$ in [17]. In general, the geometric codes and the subspace codes are classified for any finite field $GF(q)$, but a more general polynomial degree has to be specified that is related to the automorphisms of the field: mindeg for the geometric codes and maxdeg for their dual subspace codes. In the binary case (or any prime case), mindeg = maxdeg = deg, so it corresponds to Reed–Muller codes. □

Note that the matrix A in Theorem 7 may have a rank less than k, but, if it has full rank, then it is a generator matrix for a self-orthogonal code. The case where A has k rows but no columns happens when the symmetric difference of the planes is empty, such as when there are no planes.

Corollary 4. *Every self-orthogonal binary code can be constructed from a set of planes of $PG(k-1,2)$ or, equivalently, from the punctured $RM(k-3,2)$ code.*

Since the all ones' words are in the punctured RM code, or, equivalently, the whole set of points of $PG(k-1,2)$ is the symmetric difference of the characteristic functions of a number of planes, we have the following:

Corollary 5. *There is a different kind of complement for any LCD code. Assuming that the columns of a (projective) generator matrix G are non-zero and are not repeated (equivalent to saying that the dual code has a distance of at least 3), we now form a new generator matrix from all the non-zero columns that do not appear in G. Thus, from a projective $[n,k]$ LCD code, we obtain a projective $[2^k-1-n,k']$ LCD code with $k' \leq k$. In most cases, $k' = k$.*

Theorem 8. *When k is even, $k \geq 4$, the number of sets of points of $PG(k-1,2)$ that correspond to projective LCD codes is given by:*

$$2^{2^k-1-k(k+1)/2} \prod_{i=1}^{k/2} (2^{k+1} - 2^{2i}).$$

When k is odd, $k \geq 3$, the number of sets of points of $PG(k-1,2)$ that correspond to projective LCD codes is given by:

$$2^{2^k-1-k(k+1)/2} \prod_{i=0}^{(k-1)/2} (2^k - 2^{2i}).$$

Proof. A binary simplex code is self-orthogonal if its dimension is at least 3. In this smallest case $k = 3$, a corresponding generator matrix, is the $3 \times 2^3 - 1 = 7$ matrix of all possible non-zero binary columns. This being the coordinates of the 7 points of $PG(2,2)$. This is a special plane in $PG(k-1,2)$, but using collineations (non-singular linear mappings), one sees that any plane π of $PG(k-1,2)$ (if $k \geq 3$) gives a generator matrix $A(\pi)$ for a self-orthogonal code. By using the closure of all planes under symmetric differences, we obtain the "subspace code" on the points of $PG(k-1,2)$, isomorphic to an elementary abelian group over $\mathbb{Z}_2$ of dimension $2^k - \binom{k}{0} - \binom{k}{1} - \binom{k}{2} = 2^k - (k^2+k+2)/2$, and each of these sets of points gives a self-orthogonal code, also counting the empty set which gives the trivial code $\{0\}$. Note that when $k = 3$, the smallest case, the subspace code has a dimension of 1 with two words, which is the whole plane $PG(2,2)$ and the empty set.

We check that every (projective) self-orthogonal code contained as a truncation by deleting columns from the simplex code of dimension k is obtained that way because the planes are the subsets of a minimal size 7. There are no self-orthogonal codes of a length less than 7 with no zero or repeated columns when considering all cases.

Each $k \times k$ binary symmetric matrix B will correspond to a coset of the above self-orthogonal code subgroup because we can always write $B = AA^t$ for some A; for example, forming an incidence matrix (vertices versus edges) of the graph with B as (vertex) adjacency matrix and appending appropriate unit columns to ensure that the number of ones in each row is even. Any set of points that gives B as its AA^t will then be the symmetric difference of a self-orthogonal set of points with a fixed set (e.g., corresponding to the columns of the (appended) incidence matrix just mentioned). Hence, we count the number of symmetric non-singular matrices B using Lemma 2 to finish the proof. □

In [18], it is shown that a (binary) LCD code with a non-singular symmetric matrix $AA^t = B$ with a 1 on the leading diagonal (equivalently, the code is "odd-like", having odd-weight words) has a normalised generator matrix with $AA^t = I$ (the rows then form an orthonormal basis). In the case of k being even, C can be "even-like", and then the generator matrix can be normalised so that AA^t is the symmetric permutation matrix with $k/2$ transpositions. If k is odd, then C must be odd-like, since a skew-symmetric matrix (with zero diagonal) of an odd size k is always singular (from the Pfaffian formula).

Let us connect this with finite geometry. See [19] for the theory of polarities over finite fields. In general, a non-singular symmetric matrix B corresponds to an *orthogonal* polarity (having the absolute points of a quadric) and a non-singular skew-symmetric matrix corresponds to a *symplectic (or null)* polarity having every point absolute (lying on its image hyperplane). B is used to bijectively map a point (column vector) $x \in PG(k-1,2)$ to the hyperplane with dual (row) coordinates x^tB. Similarly, any hyperplane $y^t \mapsto B^{-1}y$. Repeating the polarity twice, $x \mapsto x^tB \mapsto B^{-1}B^tx = \pm x$ so that the polarity has order two (meaning it is an involution).

The (linear) polarities of $PG(k-1,q)$ (for any prime power q) are of these two standard types. There is also the *Hermitian* polarity, but this is non-linear depending on a square root subfield not existing in the binary case. The orthogonal (ordinary) type of polarity has the absolute points of a quadric, the set of points $\{x \mid x^tBx = 0\}$. This is non-degenerate in general, but, in the characteristic two case (e.g., binary), the quadric degenerates into a hyperplane (squared) if B has a non-zero main diagonal. If k is even, there can be symplectic polarities such that any point is absolute, lying on its image. Thus, $x^tBx = 0, \forall x$, meaning that B has zero diagonal. Note that the standard skew-symmetric condition for symplectic changes in characteristic two to symmetric and zero diagonal.

From the reduction to just two kinds of symmetric matrices (or linear polarities) we have this:

Proposition 1. *When k is even, $k \geq 4$, the number of sets of points of $PG(k-1,2)$ that correspond to non-equivalent LCD codes is at most:*

$$2^{2^k - k(k+1)/2}.$$

When k is odd, $k \geq 3$, the number of sets of points of $PG(k-1,2)$ that correspond to non-equivalent LCD codes is at most:

$$2^{2^k-1-k(k+1)/2}.$$

To reduce these numbers much more, one could investigate the groups of collineations of $PG(k-1,2)$ similar to [18] because sets of points that are mapped to each other by a collineation will give equivalent LCD codes.

The number of LCD codes has been counted in other ways, e.g., in [20], with mass formulae for $T_2(n,k)$. The number of non-equivalent LCD codes is harder but has also been calculated there for smaller cases.

Given a binary code with a $k \times n$ generator matrix A, we construct the matroid $M(C)$ which has the column vectors of A as a set of n points. A basis of $M(C)$ is a linearly independent set of k columns over the binary field $GF(2)$. The property of no zero or repeated columns in A means that $M(C)$ is a *simple* matroid. Note that the matroid is the same no matter what generator matrix is used because multiplying on the left of A by a non-singular matrix does not change the dependence properties of the columns of A.

$M(C)$ is called a *representable* matroid since it corresponds to a set of n points in binary projective space $PG(k-1,2)$, where each column of C forms the homogeneous coordinates for the corresponding point. Any non-zero codeword of C can be seen in the geometry of $M(C)$ by considering a hyperplane of $PG(k-1,2)$. Then, the positions with non-zero values (ones) in the codeword are those points of $M(C)$ that are not in that hyperplane. Thus, the distance of C is given by the complement in $M(C)$ of the largest intersection of C with a hyperplane. This gives a word of the smallest Hamming weight in the code.

Theorem 9. *A binary code C is an LCD code if and only if the number of bases of $M(C)$ is odd.*

Proof. This follows directly from the fact that for a binary matrix A, AA^t is the sum of the $k \times k$ subdeterminants of A, as seen when taking square roots in Corollary 1. □

The number of bases of a matroid has been well studied, and, in particular, it is related to the Tutte polynomial invariant. Here, we discuss some of the ideas which can be found in [21].

The Tutte polynomial of a matroid M is defined as

$$T_M(x,y) := \Sigma_S \, (x-1)^{z(S)}(y-1)^{n(S)},$$

where $z(S)$ is the rank $r(M)$ of the whole matroid minus the rank $r(S)$ of the subset S, and where $n(S)$, the nullity of S, is the number of points $|S|$ in S minus $r(S)$. Note that $0 \leq z(S) \leq r$, and $0 \leq n(S) \leq |S|$. Furthermore, $x(S) = 0$ if and only if S is a basis of M, while $z(S) = r(M)$ if and only if S is a set of loops of M (points of rank 0). $n(S) = 0$ if and only if S is a subset of a basis of M (an "independent" set). $n(S) = |S|$ if and only if S is a set (perhaps empty) of loops of M.

Another important property for coding theory is that the Tutte polynomial $T(M^\perp)$ of the dual matroid satisfies $T_{M^\perp}(x,y) = T_M(y,x)$. If the matroid M comes from the columns of a generator matrix of a linear code C, then the dual matroid comes from a generator matrix for the dual code $C^\perp$. In general, the dual $M^\perp$ of a matroid M is defined on the same (ordered) set of points, with the bases being the complements $M \setminus B$ of the bases B in the original matroid. The dual concept of a loop is an isthmus, which is a point of M contained in every basis of M. Thus, adding a loop to a set S never increases its rank, while adding an isthmus always increases the rank of S by one.

The Tutte polynomial can be calculated recursively by the operations of matrioid deletion and contraction, which for generator matrices correspond essentially to the oper-

ations of deleting a column or deleting a row of a generator matrix. The main relation is the following:

$$T_M(x,y) = T_{M\backslash e}(x,y) + T_{M/e}(x,y),$$

where e is a point (neither a loop nor isthmus) of M, $M\backslash e$ is the deletion of e from M, while M/e is the projection (or contraction) of M from e. The recursion in the case where f is a loop is

$$T_M(x,y) = yT_{M\backslash f}(x,y) = yT_{M/f}(x,y),$$

and, if g is an isthmus, then

$$T_M(x,y) = xT_{M\backslash g}(x,y) = xT_{M/g}(x,y).$$

Of course, with all these recursions, it turns out that the Tutte polynomial in the general case has a high complexity of calculation. However, we note that it is known that the evaluation at the point $(1,1)$ gives the number of bases of M.

Theorem 10. *$T_M(1,1)$ is the number of bases of the matroid M.*

Corollary 6. *The matroid $M(C)$ of a binary code has an odd $T_{M(C)}(1,1)$ if and only if C is an LCD.*

Since the number of maximal non-singular submatrices determines whether a generator matrix gives an LCD code or not, we have the following easy fact which also comes from the evaluation of the Tutte invariant as above.

Corollary 7. *Given any (pivot) position with the value 1 of a $k \times n$ LCD generator matrix A, either we can delete the column with the pivot to obtain a smaller LCD generator matrix or we can use row operations to make all zeros below the pivot, and then the minor of A will be a $k-1 \times n-1$ LCD generator matrix. One of these cases will occur but not both.*

Proof. This follows immediately from the permanent function description of $|AA^t|$ in Theorem 3 because the (odd) number of permutations in A is counted by going through the pivot or not. □

5. Conclusions

We have elucidated some facets of the theory of LCD codes, counting them in various ways and connecting them with other types of codes: self-orthogonal, Reed–Muller, subspace and geometric. Invariants, or more generally, quasi-invariants, have been shown to determine the binary LCD codes; notably, the permanent of a generator matrix and an evaluation of the Tutte polynomial of an associated matroid. We have also explained how the classification of polarities leads to two basic kinds of LCD code in the characteristic two case.

Author Contributions: Conceptualization, D.G.G.; validation, H.S. and P.S.; methodology, D.G.G. and S.U.R.; investigation, A.N.A., H.S.A. and S.U.R.; writing—original draft preparation, D.G.G.; resources, A.N.A.; data curation, S.U.R.; writing—review and editing, S.U.R.; project administration, A.N.A.; funding acquisition: A.N.A. All authors have read and agreed to the published version of the manuscript.

Funding: The authors extend their appreciation to the Deputyship for Research and Innovation, Ministry of Education in Saudi Arabia for funding this research work through project number 1129.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. MacWilliams, J.; Sloane, N.J.A. *The Theory of Error-Correcting Codes*; Elsevier: Amsterdam, The Netherlands, 1977.
2. van Lint, J.H. *Introduction to Coding Theory*; Springer: Berlin/Heidelberg, Germany, 1965.
3. Alahmadi, A.; Güneri, C.; Özkaya, B.; Shoaib, H.; Solé, P. On linear complementary-dual multinegacirculant codes. *Cryptogr. Commun.* **2020**, *12*, 101–113. [CrossRef]
4. Massey, J.L. Linear codes with complementary duals. *Discrete Math.* **1992**, *106/107*, 337–342. [CrossRef]
5. Harada, M.; Saito, K. Binary linear complementary dual codes. *Cryptogr. Commun.* **2019**, *11*, 677–696. [CrossRef]
6. Sendrier, N. On the dimension of the hull. *SIAM J. Discrete Math.* **1997**, *10*, 282–293. [CrossRef]
7. Sendrier, N. Linear codes with complementary duals meet the Gilbert-Varshamov bound. *Discrete Math.* **2004**, *285*, 345–347. [CrossRef]
8. Bringer, J.; Carlet, C.; Chabanne, H.; Guilley, S.; Maghrebi, H. Orthogonal direct sum masking. In Proceedings of the Workshop in Information Security Theory and Practice (WISTP 2014), Heraklion, Greece, 30 June–2 July 2014; Lecture Notes in Computer Science 8501; Naccache, D., Sauveron, D., Eds.; Springer: Berlin/Heidelberg, Germany, 2014; pp. 40–56.
9. Carlet, C.; Guilley, S. Statistical properties of side-channel and fault injection attacks using coding theory. *Cryptogr. Commun.* **2018**, *10*, 909–933. [CrossRef]
10. Overbeck, R.; Sendrier, N. Code-based cryptography. In *Post-Quantum Cryptography*; Bernstein, D.J., Buchmann, J., Dahmen, E., Eds.; Springer: Berlin/Heidelberg, Germany, 2009.
11. Radić, M.; Sušanj, R. On determinants of rectangular matrices which have Laplace's expansion along the rows. *Glasnik Mat.* **2012**, *47*, 175–180. [CrossRef]
12. Glynn, D.G. An invariant for matrices and sets of points in prime characteristic. *Des. Codes Cryptogr.* **2011**, *58*, 155–172. [CrossRef]
13. Glynn, D.G. The permanent of a square matrix. *Europ. J. Combin.* **2010**, *31*, 1887–1891. [CrossRef]
14. Hodge, W.V.D.; Pedoe, D. *Methods of Algebraic Geometry*; Cambridge University Press: Cambridge, UK, 1968; Volume 1.
15. MacWilliams, J. Orthogonal matrices over finite fields. *American Math. Monthly* **1969**, *76*, 152–164. [CrossRef]
16. Glynn, D.G.; Hirschfeld, J.W.P. On the classification of geometric codes by polynomial functions. *Des. Codes Cryptogr.* **1995**, *6*, 189–204. [CrossRef]
17. Glynn, D.G. On the orthogonality of geometric codes. *Des. Codes Cryptogr.* **2004**, *31*, 43–50. [CrossRef]
18. Carlet, C.; Mesnager, S.; Tang, C.; Qi, Y. New characterization and parameterization of LCD codes. *IEEE Trans. Inf. Theory* **2018**, *65*, 39–49.
19. Hirschfeld, J.W.P. *Projective Geometries over Finite Fields*; Oxford University Press: Oxford, UK, 1979.
20. Araya, M.; Harada, J. On the classification of linear complementary dual codes. *Discrete Math.* **2019**, *342*, 270–278.
21. Oxley, J.G. *Matroid Theory*, 2nd ed.; Oxford University Press: Oxford, UK; New York, NY, USA; Tokyo, Japan, 2006.