ENUMERATIVE COMBINATORICS AND CODING THEORY

L'Enseignement Mathématique, t. 40 (1994), p. 171-185

ENUMERATIVE COMBINATORICS AND CODING THEORY

by Shalom ELIAHOU [1])

ABSTRACT. Let f be a polynomial in n variables with non-negative integral coefficients. The enumeration of the values assumed by f on $\{1, -1\}^n$ is shown to be equivalent to the enumeration of the weights in some associated binary linear code L_f. We use this correspondence, together with the MacWilliams identity, to enumerate (1) Hadamard matrices of some fixed order, and (2) the proper 4-colorings of a graph, in terms of the weight distribution of suitable binary codes. Similar formulas could be obtained for other combinatorial objects.[2])

Let f be a polynomial in n variables with non-negative integral coefficients. We will address here the following

PROBLEM. Is there a point $p \in \{1, -1\}^n$ such that $f(p) = 0$? How many such binary zeros does f admit? More generally, what can be said about the *value enumerator* of f, which we define as

$$V_f(T) = \sum_{p \in \{\pm 1\}^n} T^{f(p)} \in \mathbf{N}[T, T^{-1}] \ ?$$

(Note that the coefficient of T^v in $V_f(T)$ is the number of binary points p such that $f(p) = v$, for $v \in \mathbf{Z}$.) Many classical combinatorial problems can be expressed in the above terms, for a suitable polynomial f.

In order to discuss these problems, we associate with f a binary linear code L_f, in such a way that the weight enumerator of L_f and the value enumerator of f faithfully reflect each other. We then invoke the MacWilliams identity from coding theory, to obtain formulas for the number of binary zeros

[1]) The author gratefully acknowledges support from the Fonds National Suisse de la Recherche Scientifique during part of the preparation of this paper.

[2]) MR classification primary 05A15, secondary 94B05, 05B20, 05C15.

of f, involving the value at -1 of a high order derivative of the weight enumerator of $L_f^{\perp}$.

The main result is a general enumeration formula, stated in Theorem 4. This formula is then applied to enumerate Hadamard matrices (Theorem 6), and the proper 4-colorings of a graph (Theorem 7), in terms of the weight enumerators of suitable binary codes. For Hadamard matrices, there seems to be no other enumeration formula in the literature.

As hinted in Section 1, similar enumeration formulas could be obtained for other combinatorial objects, such as conference matrices, Barker sequences, Golay complementary sequences, block designs, labelled regular graphs, proper graph q-colorings, and more.

1. Reduction to one special equation

Throughout the paper, we will consider a single polynomial f in n variables, with integral and non-negative coefficients, and with square-free monomials only.

This situation should be sufficiently general for most combinatorial applications. Indeed, many objects in combinatorics are defined as the solution set in $\{1, -1\}^n$, or in $\{0, 1\}^n$, of some system of polynomial equations

$$\{f_1(p) = \cdots = f_r(p) = 0\}$$

in n variables with integral coefficients. Think of Hadamard matrices, conference matrices, Barker sequences [B], Golay pairs [G, EKS], block designs, labelled regular and strongly regular graphs, proper graph colorings, etc.

By a suitable change of variables if necessary, we may and will restrict our attention to equations on $\{1, -1\}^n$. Since we are only interested in binary and hence real solutions, the system $\{f_1(p) = \cdots = f_r(p) = 0\}$ can be reduced to an equivalent single equation

$$\{f(p) = 0\},$$

where $f = f_1^2 + \cdots + f_r^2$.

Furthermore, we may assume that f has non-negative coefficients only. Indeed, suppose $f = f_1 - f_2$, where f_1, f_2 have non-negative coefficients. Introducing a new variable x_0, the system $\{f = 0\}$ is equivalent to the system $\{x_0 + 1 = 0, f_1 + x_0 f_2 = 0,\}$, which in turn is equivalent to the single equation

$$\{\hat{f} = 0\},$$

where $\hat{f} = (x_0 + 1)^2 + (f_1 + x_0 f_2)^2$. Of course, all coefficients of $\hat{f}$ are non-negative, as desired.

Finally, we may assume that all monomials in f are square-free, by removing any square in any monomial if necessary. The values assumed by f on binary points will not be altered, for if u, v are monomials, then $u^2 v - v$ takes the constant value 0 on $\{1, -1\}^n$.

2. Coding theory

To fix the notation and terminology, we briefly recall a few notions from coding theory. For more information, see [MS].

Consider the vector space $\mathbf{F}_2^N$ with its canonical basis fixed. The *(Hamming) weight* $|z|$ of a vector $z \in \mathbf{F}_2^N$ is the number of non-zero coordinates of z. A *binary linear code* (or *code,* for short) is a vector subspace C of $\mathbf{F}_2^N$. The integer N is called the *length* of C. The *weight enumerator* of C is the polynomial

$$P_C(T) = \sum_{z \in C} T^{|z|}.$$

A *generator matrix* for C is a $k \times N$ matrix G over $\mathbf{F}_2$ whose rows span C. A *parity check matrix* for C is an $(N-k) \times N$ matrix H over $\mathbf{F}_2$ such that

$$C = \{z \in \mathbf{F}_2^N, \quad H \cdot z^{\top} = 0\}.$$

Equivalently, C is the kernel of the map $h : \mathbf{F}_2^N \to \mathbf{F}_2^{N-k}$ whose matrix in the standard bases is H. The *dual* of C is the space

$$C^{\perp} = \{y \in \mathbf{F}_2^N \mid y \cdot z = 0 \text{ for all } z \in C\},$$

where $y \cdot z$ denotes the usual dot product of y and z with value in $\mathbf{F}_2$. We have:

$$\dim C^{\perp} = N - \dim C, \text{ and } C^{\perp\perp} = C.$$

A binary matrix H is a generator matrix for C if and only if it is a parity check matrix for $C^{\perp}$. Finally, the weight enumerator of C determines the weight enumerator of its dual $C^{\perp}$ by the *MacWilliams identity* [M]:

$$P_{C^{\perp}}(T) = \frac{1}{|C|} \cdot (1+T)^N \cdot P_C\left(\frac{1-T}{1+T}\right).$$

3. The code associated with f

We will denote by M_n the set of square-free monomials in the variables $x_1, \ldots, x_n$. We consider M_n as a multiplicative group, by imposing the relations $x_i^2 = 1$ for all $i = 1, \ldots, n$. Note that M_n is then isomorphic to $\{1, -1\}^n$.

Let $f = u_1 + \cdots + u_N$ be a polynomial in $x_1, \ldots, x_n$ with non-negative integral coefficients, and with monomials $u_1, \ldots, u_N \in M_n$. The u_i need not be distinct, nor necessarily distinct from 1.

Definition.

1. The *matrix* $\Phi_f = (\Phi_{i,j})$ *associated with* f is the $n \times N$ matrix over $\mathbf{F}_2$, defined by

$$\Phi_{i,j} = \begin{cases} 1 & \text{if } x_i \text{ divides } u_j, \\ 0 & \text{if not .} \end{cases}$$

   Note that Φ_f is defined up to a permutation of its columns.

2. The *binary code* L_f *associated with* f is the subcode of $\mathbf{F}_2^N$ generated by the n rows of the matrix Φ_f.

   Note that the dual code $K_f := L_f^{\perp}$ sits in the following exact sequence:

$$0 \to K_f \to \mathbf{F}_2^N \xrightarrow{\Phi_f} \mathbf{F}_2^n .$$

Indeed, K_f admits Φ_f as a parity check matrix. Note also that any binary code C is of the form $C = L_f$ for some polynomial f with non-negative coefficients in $x_1, \ldots, x_n$. Indeed, such an f can be obtained as the sum of the monomials corresponding to the columns of any generator matrix for C.

We will now give a second description of the code L_f. With any $p \in \{1, -1\}^n$, we associate

- a subset $\nu_f(p)$ of $\{1, \ldots, N\}$, defined as

$$\nu_f(p) = \{i = 1, \ldots, N \mid u_i(p) = -1\}, \text{ and}$$

- a codeword $c_f(p)$ in $\mathbf{F}_2^N$, defined as

$$c_f(p) = \sum_{i \in \nu_f(p)} E_i ,$$

where $\{E_1, \ldots, E_N\}$ denotes the standard basis of $\mathbf{F}_2^N$.

We claim, among other things, that the image of the map

$$c_f : \{1, -1\}^n \to \mathbf{F}_2^N$$

is exactly the code L_f associated with f.

PROPOSITION 1. *Let* $f = u_1 + \cdots + u_N$ *with* $u_i \in M_n$ *for all* i. *Let* $c = c_f$ *and* $L = L_f$ *denote its associated map and code. Then we have:*

1. *The map* $c : \{1, -1\}^n \to \mathbf{F}_2^N$ *is a group homomorphism;*
2. $Im(c) = L$;
3. $Ker(c) = \{p \mid f(p) = N\}$;
4. *For every* $p \in \{1, -1\}^n$, *the weight of* $c(p)$ *is related to the value* $f(p)$ *by*

$$f(p) = N - 2 \mid c(p) \mid .$$

Proof.

1. Let $p, p' \in \{1, -1\}^n$. Then $\nu_f(pp')$ is obviously equal to the symmetric difference of $\nu_f(p)$ and $\nu_f(p')$, hence

$$c(pp') = c(p) + c(p') .$$

2. For every $i = 1, \ldots, n$, let $p_i \in \{1, -1\}^n$ denote the point which has a -1 at the i-th coordinate, and a 1 elsewhere. Since the n points $p_1, \ldots, p_n$ generate $\{1, -1\}^n$ as a group, their images under c generate $\mathrm{Im}(c)$. Now,

$$\begin{aligned} \nu_f(p_i) &= \{j \mid u_j(p_i) = -1\} \qquad \text{(by definition)} \\ &= \{j \mid x_i \text{ divides } u_j\} . \end{aligned}$$

Therefore, $c(p_i)$ coincides with the i-th row of the matrix Φ_f. Since these rows generate L by definition, the claim is proved.

3. If $p \in \{1, -1\}^n$, then

$$c(p) = 0 \Leftrightarrow \nu_f(p) = \emptyset \Leftrightarrow u_i(p) = 1 \ \forall i \Leftrightarrow f(p) = N .$$

4. Let $r = \mid c(p) \mid = \mid \nu_f(p) \mid$. Then

$$f(p) = \sum_{i=1}^{N} u_i(p) = (r)(-1) + (N - r)(1) = N - 2r . \qquad \square$$

We will now show that the value enumerator of f, defined as

$$V_f(T) = \sum_{p \in \{\pm 1\}^n} T^{f(p)} ,$$

is completely determined by the weight enumerator of L_f. (And of $L_f^{\perp}$ as well, by the MacWilliams identity.)

THEOREM 2. *Let* $f = u_1 + \cdots + u_N$ ($u_i \in M_n$ *for all* i) *and let* $L = L_f$ *be its associated code. Then*

$$V_f(T) = 2^{n-\dim L} \cdot T^N \cdot P_L\left(\frac{1}{T^2}\right).$$

Proof.

$$\begin{aligned} V_f(T) &= \sum_{p \in \{\pm 1\}^n} T^{f(p)} \\ &= \sum_p T^{N-2|c(p)|} \qquad \text{(by Proposition 1)} \\ &= T^N \sum_p \left(\tfrac{1}{T^2}\right)^{|c(p)|}. \end{aligned}$$

As p runs through $\{1, -1\}^n$, $c(p)$ runs through L by Proposition 1. Furthermore, since c is a homomorphism, the fiber $c^{-1}c(p)$ of $c(p)$ contains $|\operatorname{Ker} c|$ elements, for every p. Thus,

$$\begin{aligned} \sum_p \left(\tfrac{1}{T^2}\right)^{|c(p)|} &= |\operatorname{Ker} c| \cdot \sum_{z \in L} \left(\tfrac{1}{T^2}\right)^{|z|} \\ &= |\operatorname{Ker} c| \cdot P_L\left(\tfrac{1}{T^2}\right). \end{aligned}$$

As $\dim(\operatorname{Ker} c) = n - \dim(\operatorname{Im} c) = n - \dim(L)$, the claimed formula follows. □

Notation. For any $v \in \mathbf{Z}$, we will denote by $f^{-1}(v)$ the "binary fiber" of v, i.e. the set

$$f^{-1}(v) = \{p \in \{1, -1\}^n \mid f(p) = v\}.$$

Note that $f(p) \equiv N \bmod 2$ for every binary point p, for $1 \equiv -1 \bmod 2$.

COROLLARY 3. *For every* $v \in \mathbf{Z}$, *the cardinality of* $f^{-1}(v)$ *is equal to* $2^{n-\dim L}$ *times the number of codewords in* L *which are of weight* $(N - v)/2$.

Proof. The cardinality of $f^{-1}(v)$ is equal to the coefficient b_v of T^v in the Laurent polynomial $V_f(T)$. By the theorem,

$$b_v = 2^{n-\dim L} \cdot a_w,$$

where $v = N - 2w$, and where a_w is the coefficient of T^w in $P_L(T)$. □

EXAMPLE 1: *the Hamming code.*

Let $f = (1 + x_1) \cdots (1 + x_n) - 1$. Developing f as a sum of monomials in M_n, we have

$$f = \sum_{u \in M_n \setminus \{1\}} u.$$

Thus the associated matrix Φ_f is the $n \times (2^n - 1)$ matrix over $\mathbf{F}_2$ whose columns are the elements of $\mathbf{F}_2^n$, except 0. By definition, this matrix is the parity check matrix of the Hamming code H_n. Thus, in our terminology, the Hamming code H_n is *the dual of the code* L_f *associated with* f.

The value enumerator of f is readily obtained. We have

$$f(p) = \begin{cases} 2^n - 1 & \text{if } p = (1, \ldots, 1), \\ -1 & \text{if not.} \end{cases}$$

Therefore,

$$V_f(T) = (2^n - 1)T^{-1} + T^{2^n - 1}.$$

Let $P_L(T)$ be the weight enumerator of $L = L_f$. By Theorem 2, we have

$$\begin{aligned} P_L(T^{-2}) &= T^{1-2^n} \cdot V_f(T) \\ &= T^{1-2^n} \cdot ((2^n - 1)T^{-1} + T^{2^n - 1}) \\ &= (2^n - 1)T^{-2^n} + 1, \end{aligned}$$

and hence $P_L(T) = 1 + (2^n - 1)T^{2^n - 1}$.

Finally, by the MacWilliams identity, the weight enumerator of the Hamming code $H_n = L^{\perp}$, is equal to

$$P_{H_n}(T) = \frac{1}{2^n}[(1+T)^{2^n-1} + (2^n - 1)(1+T)^{2^{n-1}-1}(1-T)^{2^{n-1}}].$$

EXAMPLE 2: *the Reed-Muller code* $\mathscr{R}(r, m)$.

Let m be a positive integer, and let $[m] := \{1, \ldots, m\}$. We consider 2^m variables $\{x_a\}$, indexed by the subsets $a \subset [m]$. If J is a set of subsets of $[m]$, we denote by u_J the monomial

$$u_J := \prod_{a \in J} x_a .$$

If $a \subset [m]$ and if $i \leqslant m$, we denote by $a^{(i)}$ the set of subsets of cardinality i in a. Now, given a non-negative integer $r \leqslant m$, we define the polynomial

$$f_{r,m} := \sum_{a \subset [m]} u_{a^{(0)}} \cdots u_{a^{(r)}} .$$

The code L_f associated with $f = f_{r,m}$, then, is known as the rth-order binary Reed-Muller code $\mathscr{R}(r, m)$. Checking the claim is left to the reader. The determination of the weight enumerator of $\mathscr{R}(r, m)$ is an open problem for $r \geqslant 3$. [MS, Chapter 15.]

4. On the least value of f

Let us consider again an integral polynomial $f = u_1 + \cdots + u_N$, with $u_i \in M_n$ for all i. By Theorem 2 and the MacWilliams identity, the cardinality of $f^{-1}(\upsilon)$ can be expressed in terms of the weight enumerator of the dual code $K_f = L_f^{\perp}$, for every υ in $\mathbf{Z}$.

In this section, we will obtain another such formula for $|f^{-1}(\upsilon)|$, provided υ is a lower bound for the range of f. These results could be applied to "count" the number of binary zeros of f, since $\upsilon = 0$ is a lower bound for the range of f^2, and f^2 has as many binary zeros as f does.

THEOREM 4. *Let $f = u_1 + \cdots + u_N$ with $u_i \in M_n$ for all i, and let $K := L_f^{\perp}$ be the dual of the code L_f associated with f, with weight enumerator $P_K(T)$. Assume that $\upsilon \in \mathbf{Z}$, $\upsilon \equiv N \bmod 2$, is a lower bound for f, i.e.*

$$\upsilon \leqslant f(p)$$

for all $p \in \{1, -1\}^n$. Then we have

$$|f^{-1}(\upsilon)| = \frac{1}{2^{\beta} \cdot \alpha!} \cdot P_K^{(\alpha)}(-1)$$

where

1. $\alpha = \alpha(\upsilon) = (N + \upsilon)/2$,
2. $\beta = \beta(\upsilon) = (N - \upsilon)/2 - n$, *and*
3. $P_K^{(\alpha)}(-1)$ *denotes the value at* -1 *of the α-th derivative of* $P_K(T)$.

Proof. Let $P_L(T) = \sum_{i=0}^{N} a_i T^i$ denote the weight enumerator of $L = L_f$, and let $\gamma = \gamma(\upsilon) = (N - \upsilon)/2$. By Corollary 3, we have $\deg P_L \leqslant \gamma$ since $f(p) \geqslant \upsilon$ for all p, and

$$|f^{-1}(\upsilon)| = 2^{n - \dim L} \cdot \alpha_{\gamma} . \tag{1}$$

Now, by the MacWilliams identity, the weight enumerator of K is given by

$$\begin{aligned} P_K(T) &= \tfrac{1}{|L|} \cdot \textstyle\sum_{i=0}^{\gamma} a_i (1+T)^{N-i}(1-T)^i \\ &= \tfrac{1}{|L|} \cdot (1+T)^{N-\gamma} \cdot \left(a_{\gamma}(1-T)^{\gamma} + (1+T)Q(T)\right) , \end{aligned}$$

where $Q(T)$ is some polynomial in T. Note that $N - \gamma = \alpha = (N + \upsilon)/2$.

To extract α_{γ} from the above expression, we derive α times, and evaluate at $T = -1$:

$$\begin{aligned} P_K^{(\alpha)}(-1) &= \tfrac{1}{|L|} \alpha!\, \alpha_{\gamma} 2^{\gamma} \\ &= \tfrac{1}{2^{\dim L}} \alpha!\, \alpha_{\gamma} 2^{N-\alpha} , \end{aligned}$$

and therefore

$$\alpha_\gamma = \frac{1}{2^{N-\alpha-\dim L}\,\alpha!} P_K^{(\alpha)}(-1) .$$

Multiplying both sides by $2^{n-\dim L}$, and plugging in equation (1), we obtain the claimed formula for $|f^{-1}(v)|$. □

COROLLARY 5. *Let* v_{min} *be the least value assumed by* f *on binary points. Then*

$$\frac{1}{2}(N + v_{min}) = \textit{the order of } -1 \textit{ as a root of } P_K(T) . \qquad \square$$

5. THE NUMBER OF HADAMARD MATRICES OF ORDER n

A *Hadamard matrix* is a square matrix H of order n with entries in $\{+1, -1\}$, satisfying the relation

$$H \cdot H^\top = nI_n .$$

($H^\top$ denotes the transpose of H, and I_n the identity matrix of order n.)

It is well known that the order of a Hadamard matrix can only be 1, 2 or a multiple of 4. Conversely, the existence of a Hadamard matrix of order n for every $n \equiv 0 \bmod 4$ is a longstanding conjecture, due to Jacques Hadamard [H]. The smallest open case currently occurs at $n = 428$. For a survey on Hadamard matrices, see [SY].

The theory exposed above yields a counting formula for Hadamard matrices of order n, in terms of the weight enumerator of a certain binary linear code of length $\binom{n}{2}^2$.

STEP 1. *Defining equations for Hadamard matrices.*

We represent binary matrices of order n as points $p = (p_{i,j}) \in \{1, -1\}^{n^2}$. Considering n^2 variables $\{x_{i,j}\}_{1 \leqslant i,j \leqslant n}$, let

$$g_{k,l} = \sum_{r=1}^{n} x_{k,r} x_{l,r} .$$

If $p = (p_{i,j})$ is a binary matrix, then $g_{k,l}(p)$ is the dot product of the k-th and l-th rows of p. Thus, a binary matrix p is Hadamard if and only if

$$g_{k,l}(p) = 0 \quad \text{for all } 1 \leqslant k < l \leqslant n .$$

STEP 2. *Reduction to a single equation.*

Let

$$g = \sum_{1 \leqslant k < l \leqslant n} g_{k,l}^2 .$$

By construction, we have the following properties:

(1) $g(p) \geqslant 0$ for every binary matrix p;

(2) $g(p) = 0$ if and only if p is Hadamard.

Developing the expression for g, we obtain:

$$\begin{aligned} g &= \textstyle\sum_{k<l} g_{k,l}^2 \\ &= \textstyle\sum_{k<l} (\sum_r x_{k,r} x_{l,r})^2 \\ &= \textstyle\sum_{k<l} (n + 2 \sum_{r<s} x_{k,r} x_{l,r} x_{k,s} x_{l,s}) \\ &= n \binom{n}{2} + 2f, \end{aligned}$$

where

$$f := \sum_{k<l} \sum_{r<s} x_{k,r} x_{l,r} x_{k,s} x_{l,s} .$$

(Of course, the above computation is performed modulo the relations $x_{i,j}^2 = 1$ for all i, j.)

The following properties of $f = \frac{1}{2}\left(g - n\binom{n}{2}\right)$ derive instantly from those of g:

(1) $f(p) \geqslant -\frac{1}{2} n \binom{n}{2}$ for every binary matrix p;

(2) $f(p) = -\frac{1}{2} n \binom{n}{2}$ if and only if p is Hadamard.

STEP 3. *The code associated with* f.

Let $K_n := L_f^{\perp}$ denote the dual of the binary code L_f associated with f, as defined in Section 3. Explicitly, we consider the map

$$\begin{aligned} \phi_n : \quad \mathbf{F}_2^{\binom{n}{2}^2} &\to \mathbf{F}_2^{n^2} \\ E(k, l; r, s) &\mapsto e_{k,r} + e_{l,r} + e_{k,s} + e_{l,s} , \end{aligned}$$

where $\{E(k, l; r, s)\}_{1 \leqslant k < l \leqslant n,\, 1 \leqslant r < s \leqslant n}$ and $\{e_{i,j}\}_{1 \leqslant i, j \leqslant n}$ denote the standard bases of the left and right spaces, respectively; by construction then, $K_n = \mathrm{Ker}(\phi_n)$.

As a direct consequence of Theorem 4 and of the above-mentioned properties of f, we obtain the

THEOREM 6. *Let* K_n *(n even) be the code of length* $\binom{n}{2}^2$ *defined as the kernel of the above map* $\phi_n : \mathbf{F}_2^{\binom{n}{2}^2} \to \mathbf{F}_2^{n^2}$. *Let* $P_n(T)$ *denote the weight enumerator of* K_n. *Then the number* $h(n)$ *of Hadamard matrices of order* n *is given by*

$$h(n) = \frac{1}{2^{\beta(n)}\,\alpha(n)!} \cdot P_n^{(\alpha(n))}(-1)\,,$$

where

1. $\alpha(n) = n^2(n-1)(n-2)/8$;
2. $\beta(n) = n^3(n-1)/8 - n^2$;
3. $P_n^{(\alpha(n))}(-1)$ *denotes the* $\alpha(n)$*-th derivative of* $P_n(T)$, *evaluated at* -1.

Proof. In the formula of Theorem 4, replace:

- N, the length of the code, by $\binom{n}{2}^2$;
- v, a lower bound for the values of f, by $-\frac{1}{2}n\binom{n}{2}$; and
- n, the number of variables in f, by n^2. □

Thus, the determination of the weight enumerator of K_n is an important problem. We will give below, without proof, the number of codewords of weight 3, 4 and 5 of K_n. (Of course, there are no words of weight 1 or 2 in K_n.) But the problem can be generalized a little bit, as follows. Consider the map

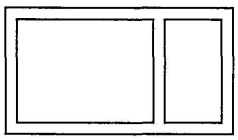
$$\begin{aligned} \phi_{m,n}: \quad \mathbf{F}_2^{\binom{m}{2}\binom{n}{2}} &\to \mathbf{F}_2^{mn} \\ E(k,l;r,s) &\mapsto e_{k,r} + e_{l,r} + e_{k,s} + e_{l,s}\,, \end{aligned}$$

where now, the indices $k < l$ range from 1 to m instead of 1 to n. We denote by $K_{m,n}$ the kernel of $\phi_{m,n}$.

Let $\Gamma = \{1, \ldots, m\} \times \{1, \ldots, n\}$. We can think of the vector basis $e_{i,j}$ as the point on row i and column j in the grid Γ, and of $E(k,l;r,s)$ as the rectangle determined by rows k, l and columns r, s in Γ. The image of $E(k,l;r,s)$ under $\phi_{m,n}$, then, is the formal sum of its four corners.

Thus, an element of weight w in $K_{m,n}$ can be pictured as a set of w rectangles in the grid Γ, such that every point in the grid coincides with an *even* number of corners of the rectangles in the set.

For example, all elements of weight 3 in $K_{m,n}$ can be represented (up to proper size and location) by the following picture:



or its vertical analogue. This picture represents a codeword of the form

$$E(k, l; r_1, r_2) + E(k, l; r_1, r_3) + E(k, l; r_2, r_3) .$$

Thus, the number of codewords of weight 3 in $K_{m,n}$ is equal to

$$w_3(K_{m,n}) = \binom{m}{2}\binom{n}{3} + \binom{m}{3}\binom{n}{2} .$$

Similarly, one can show that

$$w_4(K_{m,n}) = 3\binom{m}{2}\binom{n}{4} + 9\binom{m}{3}\binom{n}{3} + 3\binom{m}{4}\binom{n}{2} ;$$

$$w_5(K_{m,n}) = 12\binom{m}{2}\binom{n}{5} + 72\binom{m}{3}\binom{n}{4} + 72\binom{m}{4}\binom{n}{3} + 12\binom{m}{2}\binom{n}{5} + 9\binom{m}{3}\binom{n}{3} .$$

As a last remark, note that an upper bound for the weights in the associated code L_f is given by $\frac{1}{8}n^3(n-1)$, and that this bound is actually attained for some n if and only if there exists a Hadamard matrix of order n. This follows from, say, Corollary 3.

6. The number of proper 4-colorings of a graph

Let $G = (V, E)$ be a simple graph (no loops, no multiple edges) with vertex set V and edge set E. We will identify V with $\{1, \ldots, n\}$, and denote the cardinality of E by e.

A 4-*coloring* of G is the assignment to every vertex of one among four fixed colors; such a coloring is *proper* if the colors assigned to the end vertices of any edge are distinct. For a survey on the 4-colorings of planar graphs, see [SK].

We will count the number of proper 4-colorings of G, in terms of the weight enumerator of a certain code of length $3e$.

Step 1. *The defining equations for proper 4-colorings.*

As our palette of colors, we will choose the 4-set $\{1, -1\}^2$. The space of all 4-colorings of G can thus be identified with $\{1, -1\}^{2n}$, for example as follows:

Convention. If $p = (p_1, \ldots, p_{2n}) \in \{1, -1\}^{2n}$, then the pair (p_i, p_{i+n}) represents the color assigned to vertex i, for $i = 1, \ldots, n$.

Consider now $2n$ variables $x_1, \ldots, x_{2n}$, and define

$$g_{i,j} := (1 + x_i x_j)(1 + x_{i+n} x_{j+n})$$

for $1 \leqslant i, j \leqslant n$.

If p is a 4-coloring of G, then $g_{i,j}(p) = 0$ if and only if the colors assigned to vertices i and j are distinct.

Thus, a 4-coloring p of G is proper if and only if

$$g_{i,j}(p) = 0 \quad \text{for all } (i,j) \in E .$$

STEP 2. *Reduction to a single equation.*

Let

$$g := \sum_{(i,j) \in E} g_{i,j}^2 .$$

By construction, g satisfies the following properties:

(1) $g(p) \geqslant 0$ for all 4-colorings p;

(2) $g(p) = 0$ if and only if p is proper.

Developing the expression for g, we obtain:

$$\begin{aligned} g &= \textstyle\sum_E g_{i,j}^2 \\ &= \textstyle\sum_E (1 + x_i x_j + x_{i+n} x_{j+n} + x_i x_j x_{i+n} x_{j+n})^2 \\ &= 4 \textstyle\sum_E (1 + x_i x_j + x_{i+n} x_{j+n} + x_i x_j x_{i+n} x_{j+n}) \\ &= 4e + 4f, \end{aligned}$$

where

$$f := \sum_E (x_i x_j + x_{i+n} x_{j+n} + x_i x_j x_{i+n} x_{j+n}) .$$

(Here again, the computation was performed modulo $x_i^2 = 1$ for all i.)

Obviously, f satisfies the following properties:

(1) $f(p) \geqslant -e$ for all 4-colorings p;

(2) $f(p) = -e$ if and only if p is proper.

STEP 3. *The code associated with* f.

Let $K_G := L_f^{\perp}$ be the dual of the code L_f associated with f. To describe it, we consider the map

$$\begin{aligned} \phi_G : \mathbf{F}_2^{3e} &\to \mathbf{F}_2^{2n} \\ E_{i,j} &\mapsto e_i + e_j \\ E_{i+n,j+n} &\mapsto e_{i+n} + e_{j+n} \\ E_{i+2n,j+2n} &\mapsto e_i + e_j + e_{i+n} + e_{j+n} \quad ((i,j) \in E) \end{aligned}$$

where $\{E_{i,j}, E_{i+n,j+n}, E_{i+2n,j+2n}\}_{(i,j)\in E}$ and $\{e_i\}_{i\in V}$ denote the standard bases of the left and right spaces, respectively. By construction, $K_G = \mathrm{Ker}(\phi_G)$.

Here again, as a direct consequence of Theorem 4 and of the stated properties of f, we have:

THEOREM 7. *Let* K_G *be the code of length* $3e$ *defined above, and let* $P_G(T)$ *denote its weight enumerator. Then, the number* $\chi_G(4)$ *of proper* 4*-colorings of* G, *is given by*

$$\chi_G(4) = \frac{1}{4^{e-n}e!} P_G^{(e)}(-1) ,$$

where $P_G^{(e)}(-1)$ *denotes the value at* -1 *of the e-th derivative of* $P_G(T)$.

Proof. Apply the formula of Theorem 4 by replacing

- N, the length of the code, by $3e$;
- v, a lower bound for the range of f, by $-e$;
- n, the number of variables in f, by $2n$. □

Note that there are some obvious elements of weight 3 in K_G, namely

$$E_{i,j} + E_{i+n,j+n} + E_{i+2n,j+2n} \quad ((i,j) \in E) .$$

More interestingly, every cycle of length r in G gives rise to at least 3 elements of weight r in K_G. Indeed, if $(i_1, \ldots, i_r)$ is a cycle, then

$$E_{i_1+kn,i_2+kn} + E_{i_2+kn,i_3+kn} + \cdots + E_{i_r+kn,i_1+kn} \in K_G ,$$

for $k = 0, 1, 2$.

One can go a little bit further. A somewhat technical computation shows that $P_G(T)$ can be decomposed as follows:

$$P_G(T) = (1+T^3)^e P_{\tilde{K}}\left(\frac{T}{1-T+T^2}\right) ,$$

where $\tilde{K}$ is the $\mathbf{F}_4$-code defined by the exact sequence

$$0 \to \tilde{K} \to \mathbf{F}_4^e \to \mathbf{F}_4^n ,$$

with weight enumerator $P_{\tilde{K}}(T)$. (The map on the right sends the basis vector corresponding to an edge, to the formal sum of its two endvertices.)

With the above formula, we find that $P_G^{(e)}(-1) = 3^e e!\ P_{\tilde{K}}\left(-\frac{1}{3}\right)$. Plugging this into Theorem 7, we obtain

$$\chi_G(4) = \frac{3^e}{4^{e-n}} P_{\tilde{K}}\left(-\frac{1}{3}\right) .$$

In particular, G is 4-colorable if and only if $P_{\tilde{K}}\left(-\frac{1}{3}\right) \neq 0$. See also [E, Theorem 5.7], for a different formulation and proof of this formula.

REFERENCES

[B] BARKER, R.H. Group synchronizing of binary digital systems. In *Communication Theory,* W. Jackson, Ed., London: Butterworth, 1953, 273-287.

[E] ELIAHOU, S. An algebraic criterion for a graph to be four-colourable. *Aportaciones Matemáticas, Notas de Investigación 6* (1992), 3-27.

[EKS] ELIAHOU, S., M. KERVAIRE and B. SAFFARI. On Golay polynomial pairs. *Adv. Appl. Math. 12* (1991), 235-292.

[G] GOLAY, M.J.E. Static multislit spectrometry and its application to the panoramic display of infrared spectra. *J. Opt. Soc. Amer. 41* (1951), 468-472.

[H] HADAMARD, J. Résolution d'une question relative aux déterminants. *Bull. Sci. Math. (2) 17* (1893), 240-246.

[M] MACWILLIAMS, F.J. A theorem on the distribution of weights in a systematic code. *Bell Syst. Tech. J. 42* (1963), 79-94.

[MS] MACWILLIAMS, F.J. and N.J.A. SLOANE. *The Theory of Error-Correcting Codes.* North-Holland, 1977.

[SK] SAATY, T.L. and P.C. KAINEN. *The Four-Color Problem, Assaults and Conquest.* McGraw-Hill, New York, 1977.

[SY] SEBERRY, J. and M. YAMADA. Hadamard matrices, Sequences, and Block Designs. In *Contemporary Design Theory: A Collection of Surveys,* J.H. Dinitz and D.R. Stinson, Eds., Wiley-Interscience, New York, 1992, 431-560.

(Reçu le 16 novembre 1993)

Shalom Eliahou

Section de Mathématiques
Université de Genève
C.P. 240
1211 Genève 24, Switzerland
e-mail: shalom@sc2a.unige.ch