Polynomial aspects of codes, matroids and permutation groups

Peter J. Cameron
School of Mathematical Sciences
Queen Mary, University of London
Mile End Road
London E1 4NS
UK
p.j.cameron@qmul.ac.uk

Contents

Preface

The three subjects of the title (codes, matroids, and permutation groups) have many interconnections. In particular, in each case, there is a polynomial which captures a lot of information about the structure: we have the weight enumerator of a code, the Tutte polynomial (or rank polynomial) of a matroid, and the cycle index of a permutation group.

With any code is associated a matroid in a natural way. A celebrated theorem of Curtis Greene asserts that the weight enumerator of the code is a specialisation of the Tutte polynomial of the matroid. It is less well known that with any code is associated a permutation group, and the weight enumerator of the code is the same (up to normalisation) as the cycle index of the permutation group.

There is a class of permutation groups, the so-called *IBIS groups*, which are closely associated with matroids. More precisely, the IBIS groups are those for which the irredundant bases (in the sense of computational group theory) are the bases of a matroid. The permutation group associated with a code is an IBIS group, and the matroid associated to the group differs only inessentially from the matroid obtained directly from the code.

For some IBIS groups, the cycle index can be extracted from the Tutte polynomial of the matroid but not *vice versa*; for others, the Tutte polynomial can be obtained from the cycle index but not *vice versa*. This leads us to wonder whether there is a more general polynomial for IBIS groups which "includes" both the Tutte polynomial and the cycle index. Such a polynomial (the *Tutte cycle index*) is given in the last chapter of these notes (an expanded version of [5]).

Whether or not there is a more general concept extending both matroids and arbitrary permutation groups, and giving rise to a polynomial extending both the Tutte polynomial and the cycle index, I do not know; I cannot even speculate what such a concept might be.

The other theme of these notes is codes over $\mathbb{Z}_4$, the integers mod 4, where there have been some important recent developments. These codes fit naturally into the framework of permutation groups, but not so easily into the matroid framework. Carrie Rutherford has shown in her Ph.D. thesis [27] that we need a pair of matroids to describe such a code, and even then the correspondence is not exact; no natural matroid polynomial generalises the Lee weight enumerator. Moreover, the permutation group is not an IBIS group.

The remainder of the notes is concerned with developing the basics of codes, matroids and permutation groups, and their associated polynomials. For further background, see MacWilliams and Sloane [22] for codes, Oxley [25] or Welsh [30] for matroids, Cameron [4] or Dixon and Mortimer [13] for permutation groups, and Harary and Palmer [18] for the use of the cycle index in combinatorial enumeration. Another book by Welsh [31] gives further insights on polynomial aspects of codes and matroids. I refer to the Classification of Finite Simple Groups, but detailed knowledge of this is not required; see Gorenstein [15] for an overview.

These notes accompany a short course of lectures given at the Universitat Politecnica de Catalunya in Barcelona in March 2002. I have included a few exercises at the end of each chapter. I am grateful to the course participants for comments which have led to some improvements in the notes.

Peter J. Cameron
London, March 2002

CHAPTER 1

Codes

This chapter provides a very brief introduction to the theory of error-correcting codes. The highlight is the theorem of MacWilliams, asserting that the weight enumerator of a linear code determines that of its dual. The standard proof is algebraic, but we will see a combinatorial proof in Chapter 4.

1.1 Encoding and decoding

We begin with an example.

Suppose that we are transmitting information, in the form of a long string of binary digits, over a channel. There is a small probability, say 1 in 10^6, that a bit error occurs, that is, the received bit is not the same as the transmitted bit; errors in different bits are independent. In the course of sending, say, 1000 bits, the chance of an error is $1-(1-10^{-6})^{10^3}$, or about 1 in 1000, which may be unacceptably high.

Suppose that instead we adopt the following scheme. Break the data into blocks of four. Now for each 4-tuple $a=(a_1,a_2,a_3,a_4)$, we *encode* it by multiplying by the matrix

$$G=\begin{pmatrix}1&0&0&0&0&1&1\\0&1&0&0&1&0&1\\0&0&1&0&1&1&0\\0&0&0&1&1&1&1\end{pmatrix}.$$

(Arithmetic is performed in the *binary field* $\mathrm{GF}(2)=\mathbb{Z}_2$.) The first four bits of $c=aG$ are just the bits of a; the purpose of the other three bits is error correction.

We transmit the string c.

Suppose that a 7-tuple b is received. We calculate $s = bH$, where

$$H = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

If $s = 0$, we *assume* that b is the transmitted codeword. Otherwise, s is the base 2 representation of an integer i in the range $1, \ldots, 7$; we *assume* that there was a single bit error in position i, that is, we complement the ith entry of b. Then we read the first four bits of b and assume that these were the bits transmitted.

We will see shortly that our assumptions are correct provided that at most one error occurs to the bits of b. So the probability that the assumptions are wrong is $1-(1-10^{-6})^7 - 7\times 10^{-6}(1-10^{-6})^6$, which is about 2.1×10^{-11}. Now we have to send 250 blocks, so the error probability is about 1 in 190 000 000, much smaller than before!

It remains to justify our claims. First, by listing all the 7-tuples of the form aG, we find that each of them except 0 has at least three 1s. Moreover, since this set C is just the row space of G, it is closed under subtraction; so any two elements of C differ in at least three positions.This means that, if at most one error occurs, the resulting vector b is either in C (if no error occurs) or can be uniquely expressed in the form $c + e_i$, where $c \in C$ and e_i is the vector with 1 in position i and zero elsewhere. In the latter case, c was the sequence transmitted.

Now we can also check that $cH = 0$ for all $c \in C$. (For this, it is enough to show that $GH = 0$, since vectors in C have the form aG.) Then

$$(c+e_i)H = e_iH = i\text{th row of } H,$$

and H has the property that its ith row is the base 2 representation of i. So our claims about the correctness of the decoding procedure (assuming at most one error) are justified.

The price we pay for the much improved error correction capability of this scheme is slower transmission rate: instead of 1000 bits, we have to send 1750 bits through the channel. We say that the *rate* of the code is $4/7$.

To summarise: we encode the information (in blocks of four bits) as elements of the set C, and transmit these. The properties of C permit error correction. We call the set C a *code*, and its elements *codewords*.

The code C is an example of a *Hamming code*. The decoding method we described is called *syndrome decoding*.

1.2 Weights and weight enumerator

Let F be a set called the *alphabet* and n a positive integer. A *word* of length n over F is simply an n-tuple of elements of F; sometimes we write $a_1a_2\cdots a_n$ instead of $(a_1,a_2,\ldots,a_n)$. In the most important case here, F is a field; in this chapter, this is always assumed to be the case. A *code* is just a set of words, that is, a subset of F^n. We always require a code to have at least two words, since a code with one word would convey no information (since we would know for certain what message was sent). The words in a code are called *codewords*.

The code C is *linear* over the field F if it is a subspace of F^n. A linear code of length n and dimension k is referred to as an $[n,k]$ code.

From an algebraic point of view, a linear $[n,k]$ code is a k-dimensional subspace of an n-dimensional vector space *with a fixed basis*. It is this basis which makes coding theory richer than the elementary theory of a subspace of a vector space.

Let C be a $[n,k]$ code. We can describe C by a *generator matrix* G, a $k\times n$ matrix G whose rows form a basis for C, so that

$$C=\{aG : a\in F^k\}.$$

We can also describe C by a *parity check matrix* H, a $(n-k)\times n$ matrix such that C is the null space of $H^\top$, that is,

$$C=\{c\in F^n : cH^\top=0\}.$$

(This is the transpose of the matrix H of the preceding section.) The generator and parity check matrices for a given code are of course not unique.

The *dual code* $C^\perp$ of C is the set

$$C^\perp=\{x\in F^n : x\cdot c=0 \text{ for all } c\in C\},$$

where $\cdot$ denotes the standard inner product on F^n: that is,

$$a\cdot b=a_1b_1+a_2b_2+\cdots+a_nb_n.$$

Proposition 1.1 *A generator matrix for C is a parity check matrix for $C^\perp$, and* vice versa.

The *Hamming distance* $d(a,b)$ between words a and b is the number of coordinates where they differ:

$$d(a,b)=|\{i : 1\le i\le n, a_i\neq b_i\}|.$$

Let e be a positive integer. The code C is *e-error-correcting* if, for any word v, there is *at most one* codeword $c \in C$ for which $d(v,c) \leq e$. Thus, if C is used for transmitting information, and up to e errors occur during the transmission of a codeword, then the correct codeword can be recovered uniquely.

The *minimum distance* of C is the smallest distance between two different codewords. By the Triangle Inequality, if the minimum distance is at least $2e+1$, then C is e-error-correcting: for, if $d(v,c_1) \leq e$ and $d(v,c_2) \leq e$, then $d(c_1,c_2) \leq 2e$. Conversely, if the minimum distance is $2e$ or smaller, it is easy to find a word lying at distance e or smaller from two different codewords. So we have:

Proposition 1.2 *A code is e-error-correcting if and only if its minimum distance is at least* $2e+1$.

The *weight* $\mathrm{wt}(c)$ is the number of non-zero coordinates of c, that is, $\mathrm{wt}(c) = d(c,0)$, where 0 is the all-zero word. The *minimum weight* of C is the smallest weight of a non-zero codeword.

Proposition 1.3 *If C is linear, then its minimum distance is equal to its minimum weight.*

Proof Since $\mathrm{wt}(c) = d(c,0)$, every weight is a distance. Conversely, $d(c_1,c_2) = \mathrm{wt}(c_1-c_2)$; and, since C is linear, $c_1-c_2 \in C$; so every distance is a weight.

Thus, the minimum weight is one of the most significant parameters of a linear code. Indeed, if an $[n,k]$ code has minimum weight d, we sometimes describe it as an $[n,k,d]$ code.

If F is finite, the *weight enumerator* $W_C(X,Y)$ of the code C is the homogeneous polynomial

$$W_C(X,Y) = \sum_{c \in C} X^{n-\mathrm{wt}(c)} Y^{\mathrm{wt}(c)} = \sum_{i=0}^{n} A_i X^{n-i} Y^i,$$

where A_i is the number of words of weight i in C.

Two codes C, C' of length n over F are *monomial equivalent* if C' can be obtained from C by permuting the coordinates and multiplying coordinates by non-zero scalars. This is the natural equivalence relation on linear codes, and preserves dimension, weight enumerator, and most significant properties (including minimum weight).

What can be said about generator matrices of the same, or equivalent, codes? Elementary row operations on a matrix do not change its row space, and so leave the code unaltered. Column permutations, and multiplying columns by non-zero scalars, replace the code by an equivalent code. (The third type of elementary

column operation, adding a multiple of one column to another, does not preserve the structure of the code.) Thus equivalence classes of codes correspond to equivalence classes of matrices under these operations (i.e. arbitrary row operations, column permutations and scalar multiplications).

A simple example of a code is the binary *repetition code* of length n, consisting of the two words $(0,0,\ldots,0)$ and $(1,1,\ldots,1)$; its minimum weight is clearly n. Its dual is the binary *even-weight code* consisting of all words of even weight; its minimum weight is 2.

The Hamming code of the previous section is a $[7,4]$ binary linear code. If $a = 1100$, then $aG = 1100110$, a word of weight 4. Repeating for all 4-tuples a, we find that the code contains seven words of weight 3 and seven of weight 4, as well as the all-0 and all-1 words (with weight 0 and 7 respectively). So the weight enumerator is

$$X^7 + 7X^4Y^3 + 7X^3Y^4 + Y^7,$$

the minimum weight is 3, the minimum distance is also 3, and the code is 1-error-correcting (which should come as no surprise given the decoding procedure for it).

Further calculation shows that the dual code $C^\perp$ consists of the zero word and the seven words of weight 4 in C; its weight enumerator is $X^7 + 7X^3Y^4$, and its minimum weight is 4.

No brief account of codes would be complete without mention of the celebrated binary *Golay code*. This is a $[24,12,8]$ code with weight enumerator

$$X^{24} + 759X^{16}Y^8 + 2576X^{12}Y^{12} + 759X^8Y^{16} + Y^{24}.$$

This code is *self-dual*, that is, it is equal to its dual. Its automorphism group is the *Mathieu group* M_{24}.

1.3 MacWilliams' Theorem

From the weight enumerator of a code C, we can calculate the weight enumerator of the dual code $C^\perp$, using the theorem of MacWilliams:

Theorem 1.4 *Let C be an $[n,k]$ code over $\mathrm{GF}(q)$. Then the weight enumerators W_C and $W_{C^\perp}$ of C and its dual are related by*

$$W_{C^\perp}(X,Y) = \frac{1}{|C|} W_C(X + (q-1)Y, X - Y).$$

Proof We give here the classical proof, which is algebraic in nature. In Chapter 4, we will see a different, combinatorial proof.

We let χ be any non-trivial character of the additive group of $\mathrm{GF}(q)$ (that is, homomorphism from this group to the multiplicative group of complex numbers). If $q = p$ is prime, so that $\mathrm{GF}(q) = \mathbb{Z}_p$, then we can take $\chi(k) = \mathrm{e}^{2\pi ik/p}$. It is easily verified that

$$\sum_{x \in \mathrm{GF}(q)} \chi(x) = 0.$$

Now let $f(v) = X^{n-\mathrm{wt}(v)} Y^{\mathrm{wt}(v)}$ for $v \in \mathrm{GF}(q)^n$ (a term in the sum for the weight enumerator), and

$$g(u) = \sum_{v \in \mathrm{GF}(q)^n} \chi(u \cdot v) f(v)$$

for $u \in \mathrm{GF}(q)^n$. Then we have

$$\sum_{u \in C} g(u) = \sum_{v \in \mathrm{GF}(q)^n} f(v) \sum_{u \in C} \chi(u \cdot v).$$

We evaluate this sum in two ways. First, note that the inner sum on the right is equal to $|C|$ if $v \in C^\perp$, since $\chi(0) = 1$; and, for $v \notin C^\perp$, $\chi(u \cdot v)$ takes each value in $\mathrm{GF}(q)$ equally often, so the sum is zero. So the whole expression is $|C|$ times the sum of the terms $f(v)$ over $v \in C^\perp$, that is,

$$\sum_{v \in C} g(v) = |C| W_{C^\perp}(X, Y).$$

On the other hand, if we put

$$\delta(x) = \begin{cases} 0 & \text{if } x = 0, \\ 1 & \text{if } x \neq 0, \end{cases}$$

for $x \in \mathrm{GF}(q)$, then, with $u = (u_1, \ldots, u_n)$, we have

$$\begin{aligned} g(u) &= \sum_{v_1, \ldots, v_n \in \mathrm{GF}(q)} \prod_{i=1}^{n} X^{1-\delta(v_i)} Y^{\delta(v_i)} \chi(u_1 v_1 + \cdots + u_n v_n) \\ &= \prod_{i=1}^{n} \sum_{v \in \mathrm{GF}(q)} X^{1-\delta(v)} Y^{\delta(v)} \chi(u_i v). \end{aligned}$$

Now the inner sum here is equal to $X + (q-1)Y$ if $u_i = 0$, and to $X - Y$ if $u_i \neq 0$. So

$$g(u) = (X + (q-1)Y)^{n-\mathrm{wt}(u)} (X - Y)^{\mathrm{wt}(u)},$$

and $\sum_{u \in C} g(u) = W_C((X + (q-1)Y, X - Y)$. So we are done.

We saw that the weight enumerator of the Hamming code is $X^7+7X^4Y^3+7X^3Y^4+Y^7$. So the weight enumerator of the dual code is

$$\frac{1}{16}((X+Y)^7+7(X+Y)^4(X-Y)^3+7(X+Y)^3(X-Y)^4+(X-Y)^7=X^7+7X^3Y^4,$$

as we showed earlier.

Exercises

1.1. Let H be the $d\times 2^d-1$ matrix whose columns are the base 2 representations of the integers $1,\ldots,2^d-1$. Show that the $[2^d-1,2^d-d-1]$ binary code with parity check matrix H is 1-error-correcting, and devise a syndrome decoding method for it.

1.2. You are given 12 coins, one of which is known to be either lighter or heavier than all the others; you are also given a beam balance. Devise a scheme of three weighings which will identify the odd coin and determine if it is light or heavy; the coins weighed at each step should not depend on the results of previous weighings. What is the connection between this problem and error-correcting codes over $\mathbb{Z}_3=\{0,+1,-1\}$?

1.3. The *direct sum* $C_1\oplus C_2$ of two codes C_1 and C_2 is obtained by concatenating each word of C_1 with each word of C_2. Show that if C_i is a $[n_i,k_i,d_i]$ code for $i=1,2$, then $C_1\oplus C_2$ is a $[n_1+n_2,k_1+k_2,\min\{d_1,d_2\}]$ code. Show also that $W_{C_1\oplus C_2}(X,Y)=W_{C_1}(X,Y)W_{C_2}(X,Y)$. Show also how to construct

(a) a $[n_1+n_2,\min\{k_1,k_2\},d_1+d_2]$ code;

(b) a $[n_1n_2,k_2k_2,d_1d_2]$ code.

Why is there no general construction of a $[n_1+n_2,k_1+k_2,d_1+d_2]$ code?

1.4. A code (not necessarily linear) is said to be *systematic* in a given set of k coordinate positions if every k-tuple of symbols from the alphabet occurs in these positions in exactly one codeword. (Such a code contains q^k codewords, where k is the size of the alphabet.)

(a) Prove that a linear code is systematic in some set of coordinate positions.

(b) Prove that a code of length n which is systematic in every set of k coordinate positions has minimum distance $d=n-k+1$.

(A code with the property of (b) is called a *maximum distance separable* code, or *MDS code*.)

1.5. Let A be the binary code spanned by the word 01101001 and all words obtained by cyclic shifts of the first seven positions (fixing the last position). Show that A is a $[8,4,4]$ code. (This is an *extended Hamming code*.)

Let X be obtained from A by reversing the first seven positions (fixing the last position). Show that $A \cap X$ contains only the all-0 and all-1 words. Hence show that

$$G = \{(a+x, b+x, a+b+x) : a,b \in A, x \in X\}$$

is a $[24,12,8]$ code. (This is the *(extended) Golay code*.)

1.6. An *octad* in the Golay code is a set of eight coordinate positions supporting a codeword of weight 8. For any codeword $c \in G$, let $\pi(c)$ be the restriction of c to the positions of an octad. Prove that $\{\pi(c) : c \in G\}$ is the even-weight code E_8 of length 8. Now, for any subset X of E_8, let $\pi^{\leftarrow}(X)$ be the restriction to the complement of the octad of the set $\{c \in C : \pi(c) \in X\}$. Show that

(a) $\pi^{\leftarrow}(\{0\})$ is a $[16,5,8]$ code;

(b) $\pi^{\leftarrow}(E_8)$ is a $[16,11,4]$ code (each word occurring from two different codewords differing at all positions of the octad);

(c) If $X = \{00000000, 11000000, 10100000, 01100000\}$, then $\pi^{\leftarrow}(X)$ is a $[16,7,6]$ code;

(d) If $X = \{00000000, 11000000, 10100000, 10010000, 10001000, 10000100, 10000010, 10000001\}$, then $\pi^{\leftarrow}(X)$ is a nonlinear code consisting of 256 words of length 16 with minimum distance 6.

1.7. Prove that the Golay code, and the each of the codes constructed in (a), (b) and (d) of the preceding exercise, is of maximum possible cardinality for a binary code of its length and minimum distance. (Hint: Look up the Hamming and Plotkin bounds. Part (d) is more difficult!)

CHAPTER 2

Codes over $\mathbb{Z}_4$

The largest binary linear code with length 16 and minimum weight 6 has dimension 7, and thus has 128 codewords. However, this is beaten by a non-linear code, the *Nordstrom–Robinson code*, which has minimum distance 6 and has 256 codewords. (Both of these codes were constructed in Exercise 1.3.)

This code C has an additional property: for any codeword c and integer i with $0 \le i \le n$, the number of codewords c' satisfying $d(c,c') = i$ depends only on i and not on the chosen codeword $c \in C$. A code with this property is called *distance-invariant*. Another way of stating this property is as follows: for all $c \in C$, the weight enumerator of the code $C - c$ (the code c translated by $-c$) is the same. Any linear code C is distance-invariant, but it is rare for a non-linear code to have this property.

In the case of the Nordstrom–Robinson code, the weight enumerator is

$$X^{16} + 112X^{10}Y^6 + 30X^8Y^8 + 112X^6Y^{10} + Y^{16}.$$

This has an even more remarkable property. If there were a linear code C with this weight enumerator, then the MacWilliams theorem would show that $W_{C^\perp} = W_C$. For this reason, the code is called *formally self-dual*.

It turns out that the Nordstrom–Robinson code is the first member of two infinite families of non-linear codes, the *Kerdock codes* and *Preparata codes*. The nth codes K_n and P_n in each sequence have length 4^{n+1} and are distance-invariant, and their weight enumerators are related by the transformation of MacWilliams' Theorem. (They are said to be *formal duals*.)

For twenty years this observation defied explanation, until a paper by Hammons, Kumar, Calderbank, Sloane and Solé [19] presented the answer to the puzzle. We now describe this briefly.

2.1 The Gray map

The solution involves codes over the alphabet $\mathbb{Z}_4$, the integers mod 4. We regard the four elements of $\mathbb{Z}_4$ as being arranged around a circle, and define the distance d_L between two of them as the number of steps apart they are: for example, $d_L(1,3) = 2$, but $d_L(0,3) = 1$. Now we replace the Hamming distance between two words $a = (a_1, \ldots, a_n)$ and $b = (b_1, \ldots, b_n)$ of $\mathbb{Z}_4^n$ by the *Lee distance*, defined by

$$d_L(a,b) = \sum_{i=1}^{n} d_L(a_i, b_i).$$

Similarly the *Lee weight* of a is $\mathrm{wt}_L(a) = d_L(a,0)$.

Now, if C is a $\mathbb{Z}_4$-*linear* code, that is, an additive subgroup of $\mathbb{Z}_4^n$, then the *Lee weight enumerator* of C is given by

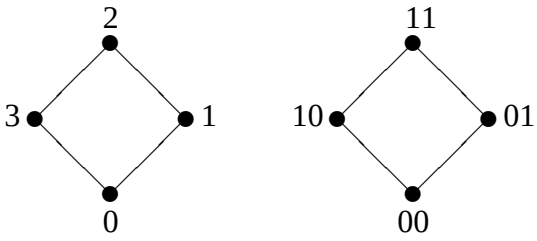
$$\mathrm{LW}_C(X,Y) = \sum_{c \in C} X^{2n - \mathrm{wt}_L(c)} Y^{\mathrm{wt}_L(c)}.$$

(Note that the maximum possible Lee weight of a word of length n is $2n$.)

It turns out that there is a version of MacWilliams' Theorem connecting the Lee weight enumerators of a $\mathbb{Z}_4$-linear code C and its dual $C^\perp$ (with respect to the natural inner product).

The set $\mathbb{Z}_4$, with the Lee metric d_L, is isometric to the set $\mathbb{Z}_2^2$ with the Hamming metric, under the *Gray map* γ, defined by

$$\gamma(0) = 00, \quad \gamma(1) = 01, \quad \gamma(2) = 11 \quad \gamma(3) = 10.$$



(More generally, a Gray map on the integers mod 2^n is a bijection to $\mathbb{Z}_2^n$ such that the images of consecutive integers lie at Hamming distance 1. Gray maps are used in analog-to-digital conversion.)

Now we extend the definition of the Gray map to map from $\mathbb{Z}_4^n$ to $\mathbb{Z}_2^{2n}$ by

$$\gamma(a_1, \ldots, a_n) = (\gamma(a_1), \ldots, \gamma(a_n)).$$

It is easily seen that γ is an isometry from $\mathbb{Z}_4^n$ (with the Lee metric) to $\mathbb{Z}_2^{2n}$ (with the Hamming metric).

The Gray map is non-linear, so the image of a $\mathbb{Z}_4$-linear code C is usually a non-linear binary code. But the isometry property shows that $\gamma(C)$ is necessarily distance-invariant, and that its weight enumerator is equal to the Lee weight enumerator of C. Thus, taking a $\mathbb{Z}_4$-linear code and its dual, and applying the Gray map, we obtain a pair of formally self-dual non-linear binary codes.

Hammons *et al.* show that, if this procedure is applied to the $\mathbb{Z}_4$ analogue of the extended Hamming codes and their duals, then the Preparata and Kerdock codes are obtained. Thus, the mystery is explained. (There is a small historical inaccuracy in this statement. They obtained, not the original Preparata codes, but another family of codes with the same weight enumerators.)

There is a more general weight enumerator associated with a $\mathbb{Z}_4$-linear code C. This is the *symmetrised weight enumerator* of C, defined as follows:

$$\mathrm{SW}_C(X,Y,Z) = \sum_{c\in C} X^{n_0(c)}Y^{n_2(c)}Z^{n_{13}(c)},$$

where $n_0(c)$ is the number of coordinates of C equal to zero; $n_2(c)$ the number of coordinates equal to 1; and $n_{13}(c)$ the number of coordinates equal to 1 or 3. Since these coordinates contribute respectively 0, 2, and 1 to the Lee weight, we have

$$\mathrm{LW}_C(X,Y) = \mathrm{SW}_C(X^2,Y^2,XY).$$

2.2 Chains of binary codes

Another approach to $\mathbb{Z}_4$-linear codes is via a representation as pairs of $\mathbb{Z}_2$-linear codes. Let C be a $\mathbb{Z}_4$-linear code. We construct binary codes C_1 and C_2 as follows. C_1 is obtained just by reading the words of C modulo 2; and C_2 is obtained by selecting the words of C in which all coordinates are even, and replacing the entries 0 and 2 mod 4 by 0 and 1 mod 2.

Theorem 2.1 *The pair (C_1,C_2) of binary codes associated with a $\mathbb{Z}_4$-linear codes C satisfies*

(a) $C_1 \subseteq C_2$;

(b) $|C| = |C_1|\cdot|C_2|$;

(c) $W_{C_1}(X,Y) = \mathrm{SW}_C(X,X,Y)/|C_2|$ *and* $W_{C_2}(X,Y) = \mathrm{SW}_C(X,Y,0)$.

Proof (a) If $v \in C$, then doubling v gives a word with all coordinates even; the corresponding word in C_2 is obtained by reading v mod 2. So $C_1 \subseteq C_2$.

(b) C_1 is the image of C under the natural homomorphism from $\mathbb{Z}_4^n$ to $\mathbb{Z}_2^n$, and C_2 is naturally bijective with the kernel of this map; so $|C| = |C_1| \cdot |C_2|$.

The proof of (c) is an exercise.

We call a pair (C_1, C_2) of binary linear codes with $C_1 \subseteq C_2$ a *chain* of binary codes.

Every chain of binary codes arises from a $\mathbb{Z}_4$-linear code in the manner of the theorem. For suppose that binary codes C_1 and C_2 are given with $C_1 \subseteq C_2$. Let

$$C = \{v_1 + 2v_2 : v_1 \in C_1, v_2 \in C_2\},$$

where the elements 0 and 1 of $\mathbb{Z}_2$ are identified with 0 and 1 in $\mathbb{Z}_4$ for this construction. Then the preceding construction applied to C recovers C_1 and C_2. So every *chain of codes* (that is, every pair (C_1, C_2) with $C_1 \subseteq C_2$) arises from a $\mathbb{Z}_4$-linear code.

However, the correspondence fails to be bijective, and many important properties are lost. Fore example, the two $\mathbb{Z}_4$-codes

$$\{000, 110, 220, 330\} \qquad \text{and} \qquad \{000, 112, 220, 332\}$$

give rise to the same pair of binary codes (with $C_1 = C_2 = \{000, 110\}$) but have different symmetrised weight enumerators (and so different Lee weight enumerators).

The problem of describing all $\mathbb{Z}_4$-linear codes arising from a given chain has not been solved. It resembles in some ways the "extension problem" in group theory.

Exercises

2.1. Prove that the Nordstrom–Robinson code as defined in Exercise 1.3 is distance-invariant and has the claimed weight enumerator.

2.2. Prove Theorem 2.1(c). Verify the conclusion directly for the two codes in the example following the theorem. Construct the images of these two codes under the Gray map.

2.3. Show that the $\mathbb{Z}_4$-linear code with generator matrix

$$\begin{pmatrix} 1 & 3 & 1 & 2 & 1 & 0 & 0 & 0 \\ 1 & 0 & 3 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 0 & 3 & 1 & 2 & 1 & 0 \\ 1 & 0 & 0 & 0 & 3 & 1 & 2 & 1 \end{pmatrix}$$

is equal to its dual and has Lee weight enumerator

$$X^{16} + 112X^{10}Y^6 + 30X^8Y^8 + 112X^6Y^{10} + Y^{16}.$$

(This is the code whose Gray map image is the Nordstrom–Robinson code.)

2.4. Prove that, for any $a, b \in \mathbb{Z}_4$, we have

$$\gamma(a+b) = \gamma(a) + \gamma(b) + (\gamma(a) + \gamma(-a)) * (\gamma(b) + \gamma(-b)),$$

where $*$ denotes componentwise product: $(a,b) * (c,d) = (ac, bd)$.

Hence prove that a (not necessarily linear) binary code C is equivalent to the Gray map image of a linear $\mathbb{Z}_4$ code if and only if there is a fixed-point-free involutory permutation σ of the coordinates such that, for all $u, v \in C$, we have

$$u + v + (u + u\sigma) * (v + v\sigma) \in C,$$

where $*$ is the componentwise product of binary vectors of arbitrary length.

(Define σ so that, if $u = \gamma(c)$, then $u\sigma = \gamma(-c)$; this permutation interchanges the two coordinates corresponding to each coordinate of the $\mathbb{Z}_4$ code.)

CHAPTER 3

Matroids

The notion of linear independence of a family of vectors in a vector space satisfies two simple conditions (namely, a subfamily of a linearly independent family is linearly independent, and the well-known *exchange property*), from which most of its familiar properties hold: the existence and constant size of bases, the rank and nullity theorem, etc. These properties crop up in various other situations. Indeed, the exchange property is credited to Steinitz who observed it for the notion of algebraic independence of elements in a field over an algebraically closed subfield. This leads to the concept of the transcendence degree of a field extension. Furthermore, subsets of the edge set of a graph which induce acyclic graphs (forests), and subfamilies of families of sets possessing systems of distinct representatives, also satisfy these conditions.

The underlying abstract structure was given the name "matroid" by Whitney (a generalisation of "matrix"). Tutte observed that a two-variable generalisation of the chromatic polynomial of a graph could also be extended to this setting; this is the *Tutte polynomial* of the matroid. In this chapter, we provide a brief introduction to these concepts.

3.1 The basics

Let E be a set. A *matroid* M on E is a pair $(E, \mathcal{J})$, where $\mathcal{J}$ is a non-empty family of subsets of E (called *independent sets*) with the properties

(a) if $I \in \mathcal{J}$ and $J \subseteq I$, then $J \in \mathcal{J}$;

(b) (the *exchange property*) if $I_1, I_2 \in \mathcal{J}$ and $|I_1| < |I_2|$, then there exists $e \in I_2 \setminus I_1$ such that $I_1 \cup \{e\} \in \mathcal{J}$.

As noted earlier, matroids were introduced by Whitney to axiomatise the notion of linear independence in a vector space. Indeed, if E is a family of vectors in a vector space V, and $\mathcal{J}$ is the set of linearly independent subsets of E, then $(E, \mathcal{J})$ is a matroid. Such a matroid is called a *vector matroid*.

Note that we speak of a family rather than a set of vectors here, since the same vector may occur more than once. (Any family containing a repeated vector is to be regarded as linearly dependent.) If we think of the vectors as the n columns of a matrix, we can regard the set E of elements of the matroid as the index set $\{1, 2, \ldots, n\}$ for the columns; then a subset I of E is independent if and only if the family of columns with indices in I is linearly independent.

More formally, a *representation* of a matroid $(E, \mathcal{J})$ over a field F is a map χ from E to an F-vector space with the property that a subset I of E belongs to $\mathcal{J}$ if and only if $\chi(I)$ is linearly independent. Two representations χ, χ' of M are *equivalent* if there is an invertible linear transformation of V whose composition with χ is χ'.

We will frequently meet the special case where E consists of all the vectors in an n-dimensional vector space over $\mathrm{GF}(q)$. This will be referred to as the *(complete) vector matroid*, and denoted by $V(n, q)$.

As referred to in the introduction, the following are also examples of matroids:

(a) Let E be a finite family of elements in a vector space, and $\mathcal{J}$ the set of *affine independent* subfamilies. (A family $(v_j : j \in J)$ is affine independent if the relation $\sum c_j v_j = 0$, where c_j are scalars with $\sum c_j = 0$, implies that $c_j = 0$ for all j.) Then $(E, \mathcal{J})$ is a matroid. Such a matroid is called *affine*.

(b) Let K be an algebraically closed field containing an algebraically closed subfield F. Let E be a finite family of elements of K, and $\mathcal{J}$ the set of all subfamilies of E which are algebraically independent over F. Then $(E, \mathcal{J})$ is a matroid. Such a matroid is called *algebraic*.

(c) Let $G = (V, E)$ be a finite graph (loops and multiple edges are allowed). Let $\mathcal{J}$ be the set of all subsets A of E for which the graph (V, A) is acyclic (that is, a forest). Then $(E, \mathcal{J})$ is a matroid. Such a matroid is called *graphic*, and is denoted by $M(G)$.

(d) Let $(X_e : e \in E)$ be a family of sets. Let $\mathcal{J}$ be the family of all subsets $I \subseteq E$ for which the subfamily $(X_e : e \in I)$ possesses a transversal (that is, there is a family $(x_e : e \in I)$ of distinct elements such that $x_e \in X_e$ for all $e \in I$). Then $(E, \mathcal{J})$ is a matroid. Such a matroid is called *transversal*.

It follows from the second axiom that all maximal independent sets in a matroid M have the same cardinality k, called the *rank* of M. These maximal independent sets are called the *bases* of M. It is possible to recognise when a family $\mathcal{B}$ of subsets of E consists of the bases of a matroid on E. This is the case if and only if

(a) no element of $\mathcal{B}$ properly contains another;

(b) if $B_1, B_2 \in \mathcal{B}$ and $y \in B_2 \setminus B_1$, then there exists $x \in B_1 \setminus B_2$ such that $B_1 \setminus \{x\} \cup \{y\} \in \mathcal{B}$. (This property is also referred to as the *exchange property*.)

We can extend the definition of rank to all subsets of E: the rank ρA of an arbitrary subset A of E is the cardinality of the largest independent set contained in A. It is also possible to recognise when a function ρ from the power set of a set E to the non-negative integers is the rank function of a matroid. (Again, the exchange property shows that any two maximal independent subsets of A have the same cardinality.)

The set of all complements of bases of M is the set of bases of another matroid M^* on E, called the *dual* of M. This is most easily proved by showing that conditions (a) and (b) above for a family $\mathcal{B}$ of sets imply the same condition for the family of complements.

A *flat* in a matroid $M = (E, \mathcal{I})$ is a subset F of E with the property that $\rho(F \cup \{x\}) = \rho F + 1$ for all $x \in E \setminus F$. If $\rho F = k$ and A is an independent subset of F of cardinality k, then $F = \{x \in E : \rho(A \cup \{x\}) = \rho A\}$. A flat whose rank is one less than that of E is called a *hyperplane*.

The flats of a matroid form a lattice (in which the meet operation is intersection), which is atomic and submodular; these properties of a lattice ensure that it arises as the lattice of flats of a matroid.

There are many other equivalent ways of defining matroids: via circuits, cocircuits, flats, hyperplanes, etc. We do not pursue this here but refer to the books [25] and [30].

Let $M = (E, \mathcal{I})$ be a matroid of rank r, and let k be a non-negative integer with $k \leq r$. The *truncation* of M to rank k is the matroid on E whose family of independent sets is

$$\mathcal{I}_k = \{I \in \mathcal{I} : |I| \leq k\}.$$

The flats of the truncation are all the flats of rank less than k of the original matroid together with the whole set E.

We conclude with some simple examples of matroids. The *free matroid* on a finite set E is the matroid in which every subset of E is independent. If $|E| = n$, this matroid is denoted by F_n. The *uniform matroid* $U_{r,n}$, with $r \leq n$, is the truncation of the free matroid F_n to rank r; in other words, its independent sets are all the subsets of E of cardinality at most r.

3.2 Deletion and contraction

The roots of matroid theory in graph theory explain much of the terminology used. For example, the use of the letter E for the set of elements of a matroid arises from its use as the edge set of a graph. In this section, we will meet loops, deletion and contraction, all of which are more transparent for graphic matroids.

Let $M = (E, \mathcal{I})$ be a matroid. The element $e \in E$ is called a *loop* if $\{e\} \notin \mathcal{I}$, or equivalently, if $\rho\{e\} = 0$. In a graphic matroid, e is a loop if and only if it is a loop of the underlying graph. Thus, an element is a loop if and only if it is contained in no basis.

The element $e \in E$ is a *coloop* if it is a loop in the dual matroid M^*. Thus, e is a coloop if and only if it is contained in every basis of M; that is, $\rho(A \cup \{e\}) = \rho A + 1$ whenever $e \notin A$. In a graphic matroid, e is a coloop if and only if it is a *bridge*, an element whose removal increases by one the number of connected components.

Let e be an element which is not a coloop. The *deletion* of E is the matroid $M \backslash e$ on the set $E \setminus \{e\}$ in which a subset A is independent if and only if it is independent in M (and doesn't contain e). There is no compelling reason to forbid the deletion of coloops, but it makes the theory tidier – see the next paragraph. In a graphic matroid, deletion of e corresponds to deletion of the edge e from the graph.

Let e be an element which is not a loop. The *contraction* of e is the matroid M/e on the set $E \setminus \{e\}$ in which a set A is independent if and only if $A \cup \{e\}$ is independent in M. (Here it is clear that contracting a loop would make no sense, so our earlier restriction will preserve duality.) In a graphic matroid, contraction of e corresponds to contraction of the edge e, that is, identifying the vertices forming the two ends of e.

Proposition 3.1 *Let e be an element of the matroid M which is not a loop. Then e is not a coloop of M^*, and*

$$(M/e)^* = M^* \backslash e.$$

Deletion and contraction form the basic inductive method for studying matroids, as we will see.

3.3 Rank polynomial and Tutte polynomial

Let M be a matroid on the set E, having rank function ρ. The *Tutte polynomial* of M is most easily defined as follows:

$$T(M; x, y) = \sum_{A \subseteq E} (x-1)^{\rho E - \rho A} (y-1)^{|A| - \rho A}.$$

For example, the Tutte polynomial of the uniform matroid $U_{r,n}$ is

$$T(U_{r,n};x,y) = \sum_{i=0}^{r} \binom{n}{i} (x-1)^{r-i} + \sum_{i=r+1}^{n} \binom{n}{i} (y-1)^{i-r},$$

since a set A of cardinality $i \leq r$ satisfies $\rho E - \rho A = r - i$ and $|A| - \rho A = 0$, while a set A of cardinality $i \geq r+1$ satisfies $\rho E - \rho A = 0$ and $|A| - \rho A = i - r$.

The appearance of the terms $x-1$ and $y-1$ in the polynomial is a historical accident. Tutte defined his polynomial by a completely different method, depending on the choice of an ordering of the elements of the matroid, but giving a result independent of the ordering. Meanwhile, the *rank polynomial* of M was defined as

$$R(M;x,y) = \sum_{A \subseteq E} x^{\rho E - \rho A} y^{|A| - \rho A}.$$

Crapo [11] showed that in fact $T(M;x,y) = R(M;x-1,y-1)$.

A number of simple matroid invariants can be extracted from the Tutte polynomial, as the next result shows. The proof is an exercise.

Proposition 3.2 *Let M be a matroid on n elements.*

(a) The number of bases of M is equal to $T(M;1,1)$.

(b) The number of independent sets of M is equal to $T(M;2,1)$.

(c) The number of spanning sets of M is equal to $T(M;1,2)$.

(d) $T(M;2,2) = 2^n$.

Calculation of the Tutte polynomial is possible by an inductive method using deletion and contraction, as follows.

Theorem 3.3 *(a) $T(\emptyset;x,y) = 1$, where $\emptyset$ is the empty matroid.*

(b) If e is a loop, then $T(M;x,y) = yT(M\backslash e;x,y)$.

(c) If e is a coloop, then $T(M;x,y) = xT(M/e;x,y)$.

(d) If e is neither a loop nor a coloop, then

$$T(M;x,y) = T(M\backslash e;x,y) + T(M/e;x,y).$$

Proof (a) is trivial. For the other parts, we note that each subset A of M/e or $M\backslash e$ corresponds to a pair of subsets A and $A\cup\{e\}$ of M. let $M' = M\backslash e$ and $M'' = M/e$ (where appropriate), and use ρ_M, $\rho_{M'}$ and $\rho_{M''}$ for the rank functions of the three matroids M, M', M'', and $E' = E'' = E\setminus\{e\}$.

If e is a loop, then we have

$$\begin{aligned}&\rho_M E = \rho_{M'}E',\\ &\rho_M A = \rho_M A\cup\{e\} = \rho_{M'}A,\\ &|A\cup\{e\}| = |A|+1, |E| = |E'|+1.\end{aligned}$$

Thus the two terms in the sum for $T(M)$ are respectively 1 and $y-1$ times the term in $T(M')$ corresponding to A, and so (b) holds.

The other two parts are proved similarly.

As an illustration of the use of the inductive method, we consider the *chromatic polynomial* of a graph G, the polynomial P_G with the property that $P_G(k)$ is equal to the number of proper k-colourings of G.

Corollary 3.4 *Let $G = (V,E)$ be a graph. Then*

$$P_G(k) = (-1)^{\rho(G)}k^{\kappa(G)}T(M(G);1-k,0),$$

where $\kappa(G)$ is the number of connected components of G and $\rho(G)+\kappa(G)$ the number of vertices.

Proof The matroid $M(G)$ associated with G has rank $\rho E = n-\kappa(G)$, where n is the number of vertices. Let k be any positive integer.

The chromatic polynomial satisfies the following recursion:

(a) If G has n vertices and no edges, then $P_G(k) = k^n$.

(b) If G contains a loop, then $P_G(k) = 0$.

(c) If e is an edge which is not a loop, then

$$P_G(k) = P_{G\backslash e}(k) - P_{G/e}(k),$$

where $G\backslash e$ and G/e are the graphs obtained from G by deleting and contracting e, respectively.

Here (a) is clear since any vertex-colouring of the null graph is proper; and (b) is trivial. For (c), we note that, if e has vertices v and w, the proper colourings c of $G\backslash e$ can be divided into two classes:

(a) those with $c(v) \neq c(w)$, which yield proper colourings of G;

(b) those with $c(v) = c(w)$, which yield proper colourings of G/e.

Now we show by induction on the number of edges that

$$P_G(k) = (-1)^{\rho(G)} k^{\kappa(G)} T(M(G); 1-k, 0).$$

This is clear when there are no edges since $\rho(G) = 0$, $\kappa(G) = n$ and $T(M(G)) = 1$. It is also clear if there is a loop, since $T(M(G); x, 0) = 0$ in that case by part (b) of Theorem 3.3. If e is a coloop then deletion of e increases κ by 1 and decreases ρ by 1; also $P_{G\setminus e}(k) = kP_G(k)/(k-1)$, since a fraction $(k-1)/k$ of the colourings of $G\setminus e$ will have the ends of e of different colours. So the inductive step is a consequence of part (c) of Theorem 3.3.

Finally, if e is neither a loop nor a coloop, use (c) above and (d) of Theorem 3.3.

The Tutte polynomials of a matroid and its dual are very simply related:

Proposition 3.5

$$T(M^*; x, y) = T(M; y, x).$$

Proof Let A be a subset of E and let $E^* = E$ and $A^* = E \setminus A$. If ρ_M and ρ_{M^*} are the rank functions of M and M^* respectively, we have

$$\begin{aligned} |A^*| - \rho_{M^*}(A^*) &= \rho_M(E) - \rho_M(A), \\ \rho_{M^*}(E^*) - \rho_{M^*}(A^*) &= |A| - \rho_M(A). \end{aligned}$$

So the term in $T(M^*)$ arising from A^* is equal to the term in $T(M)$ arising from A but with x and y interchanged.

3.4 Perfect matroid designs

A *perfect matroid design*, or PMD, is a matroid having the property that the cardinality of a flat depends only on its rank. If the rank is r, and the cardinality of an i-flat is k_i for $i = 0, \ldots, r$ (with, of course, $k_r = n$, the total number of elements of the matroid), then we describe it as a PMD$(k_0, k_1, \ldots, k_r)$.

In a PMD$(k_0, k_1, \ldots, k_r)$, the number of loops is k_0; deleting the loops gives a PMD$(0, k_1 - k_0, \ldots, k_r - k_0)$. So usually nothing is lost by assuming that $k_0 = 0$.

In a PMD$(0,k_1,k_2,\ldots,k_r)$, each element is one of a family of k_1 parallel elements. Identifying these classes, we obtain a PMD$(0,1,k_2/k_1,\ldots,k_r/k_1)$. So again we often assume that $k_1 = 1$. This reduction is a bit more problematic, as we will see when we consider group actions.

Other operations on PMDs which yield PMDs are deletion, contraction, and truncation.

Not very many PMDs are known. The list below includes all PMDs with $k_0 = 0$ and $k_1 = 1$ which are not proper truncations.

(a) The free matroid on n elements (the matroid in which every set is independent) is a PMD$(0,1,\ldots,n)$.

(b) The complete vector matroid $V(n,q)$ is a PMD$(1,q,q^2,\ldots,q^n)$. (The elements of this matroid are the vectors in GF$(q)^n$, and independence is the usual notion of linear independence.) If we delete the zero vector and shrink each 1-dimensional subspace to a point, we obtain the *projective geometry* PG$(n-1,q)$, which is a PMD$(0,1,q+1,\ldots,(q^n-1)/(q-1))$.

(c) The affine geometry AG(n,q) is a PMD$(0,1,q,q^2,\ldots,q^n)$. (The elements of this matroid are the vectors in GF$(q)^n$, but independence is now the notion of affine independence defined earlier: vectors $v_1,\ldots,v_d$ are affine independent if there is no linear dependence

$$c_1v_1+\cdots+c_dv_d = 0$$

where $c_1+\cdots+c_d = 0$ and the c_i not all zero. (An equivalent condition for $d \geq 1$ is that the vectors $v_2-v_1,\ldots,v_d-v_1$ are linearly independent.)

(d) Let t,k,n be positive integers with $t < k < n$. A *Steiner system* $S(t,k,n)$ consists of a set X of n points, and a set $\mathcal{B}$ of subsets of S called blocks, such that any t points are contained in a unique block. From a Steiner system, we obtain a matroid on the set of points as follows: every set of cardinality at most t is independent; and a set of cardinality $t+1$ is independent if and only if it is not contained in a block. This is a PMD$(0,1,\ldots,t-1,k,n)$ in which the hyperplanes are the blocks.

(e) The points and lines of an affine space AG$(d,3)$ form a Steiner triple system (that is, a Steiner system $S(2,3,n)$) with the property that any three points not contained in a block lie in a unique subsystem with 9 points (an affine plane). Marshall Hall [17] discovered that there are other Steiner triple systems with this property. These are now called *Hall triple systems*. Such a system gives rise to a PMD$(0,1,3,9,n)$ of rank 4, where a 3-set is independent if it is not a block, and a 4-set is independent if it is not contained in an

affine plane. The number of points in a Hall triple system must be a power of 3.

See Deza [12] for a survey of perfect matroid designs.

The following theorem is due to Mphako [24].

Theorem 3.6 *Let M be a* PMD$(k_0,\dots,k_r)$. *Then the Tutte polynomial of M is determined by the numbers $k_0,\dots,k_r$.*

Proof It is enough to determine the number $a(m,i)$ of subsets of the domain which have cardinality m and rank i for all m and i: for

$$T(M;x,y)=\sum_{i=0}^{r}\sum_{m=i}^{n}a(m,i)(x-1)^{k-i}(y-1)^{m-i},$$

where $n=k_r$ is the number of points.

Let $s(i,j)$ be the number of i-flats containing a given j-flat for $j\le i$. Then

$$s(i,j)=\prod_{h=j}^{i-1}\frac{n-k_h}{k_i-k_h}.$$

For let $(x_1,\dots,x_j)$ be a basis for a j-flat F_j. The number of ways of choosing $x_{j+1},\dots,x_i$ so that $(x_1,\dots,x_i)$ is independent is the numerator of the above expression. Then this set spans an i-flat F_i containing F_j, and the number of ways of extending $(x_1,\dots,x_j)$ to a basis for F_i is the denominator.

Now we have

$$s(i,0)\binom{n_i}{m}=\sum_{j=0}^{i}a(m,j)s(i,j).$$

For the left-hand side counts the number of choices of an i-flat F_i and a subset of F_i of cardinality m. This subset has rank j for some $j\le i$, and spans a j-flat contained in F_i. So each m-set of rank j contributes $s(i,j)$ to the count.

This is a triangular system of equations for $a(m,j)$ with diagonal coefficients $s(i,i)=1$. We see that the $a(m,j)$ are indeed determined by $k_0,\dots,k_r$.

Exercises

3.1. Prove that algebraic matroids, graphic matroids, and transversal matroids do indeed satisfy the matroid axioms.

3.2. In this exercise, we prove that a graphic matroid is representable over any field.

Let $G = (V,E)$ be a graph, where $V = \{v_1,\ldots,v_n\}$ and $E = \{e_1,\ldots,e_m\}$. Choose arbitrarily an orientation of each edge e_i (that is, the edge e_i has an initial and a terminal vertex, which may be the same). Now construct an $n \times m$ matrix $A = (a_{ij})$ as follows:

$$a_{ij} = \begin{cases} +1 & \text{if } e_j \text{ is a non-loop with terminal vertex } v_i; \\ -1 & \text{if } e_j \text{ is a non-loop with initial vertex } v_i; \\ 0 & \text{otherwise.} \end{cases}$$

Prove that, given any cycle in the graph, the sum of the columns corresponding to the edges in the cycle (with signs ± 1 chosen appropriately) is zero. Prove also that if a set of edges contains no cycle, then there is a row containing a single non-zero entry in the corresponding columns.

Hence show that, for any field F, a set of columns of A is linearly independent over F if and only if the corresponding set of edges of G forms an acyclic subgraph.

3.3. What are the bases, the flats, the hyperplanes, and the rank function of the uniform matroid $U_{r,n}$? What is the dual of this matroid?

3.4. Prove that the matroid $U_{2,4}$ is not graphic.

3.5. Prove that every affine matroid can be represented as a vector matroid in a space of dimension one greater than the one affording the affine representation.

3.6. Let M be a graphic matroid arising from a connected graph $G = (V,E)$ on n vertices. Prove that the rank function is given by

$$\rho A = n - \kappa(A),$$

where $\kappa(A)$ is the number of connected components of the graph (V,A).

3.7. Let $M(G)$ be a graphic matroid, where the graph $G = (V,E)$ is connected. Show that a set $A \subseteq E$ is independent in $M(G)^*$ if the removal of A does not disconnect G.

3.8. Construct

(a) non-isomorphic graphs G_1, G_2 for which the graphic matroids are isomorphic;

(b) non-isomorphic graphic matroids $M(G_1), M(G_2)$ which have the same Tutte polynomial.

3.9. As we mentioned in Chapter 1, the binary *Golay code* is a $[24,12,8]$ code containing 759 words of weight 8. Prove that the 759 subsets of cardinality 8 of $\{1,\ldots,24\}$ which support codewords of weight 8 are the blocks of a Steiner system $S(5,8,24)$.

3.10. Show that the blocks of a Steiner system $S(t+1,2t,n)$ are the supports of words of minimum weight in a linear binary code if and only if the system has the *symmetric difference property*: if B_1 and B_2 are blocks for which $|B_1 \cap B_2| = t$, then their symmetric difference $B_1 \triangle B_2$ is a block.

Find examples with $t = 2$.

CHAPTER 4

Matroids and codes

There is a very close correspondence between linear codes, on one hand, and matroids (specifically, representations of matroids) on the other – the two types of structure correspond exactly, up to the natural definition of equivalence in each case. Among other things, this correspondence leads us to the theorem of Curtis Greene, showing that the weight enumerator of a code is a specialisation of the Tutte polynomial of the corresponding matroid. This then provides a combinatorial proof of MacWilliams' Theorem on the weight enumerators of dual codes.

As we already noted, Carrie Rutherford represented a $\mathbb{Z}_4$-linear code C by a chain of binary codes. She went on to associate a three-variable analogue of the Tutte polynomial to such a chain. This polynomial specialises to give various properties of C (though not its symmetrised weight enumerator). We describe this in the last section of the chapter.

4.1 The correspondence

Let A be a $k \times n$ matrix over a field F, satisfying the condition that the rows of A are linearly independent, so that the row space of A has dimension k.

There are two different structures that can be built from A.

First, the row space of A is an $[n,k]$ code over F, that is, a k-dimensional subspace of F^n. Now row operations on A simply change the basis for the code, leaving the actual code completely unaltered. Column permutations, and multiplications of columns by non-zero scalars, replace the code by a monomial equivalent code.

Second, there is a matroid M on the set $E = \{1,2,\ldots,n\}$, in which a set I is

independent if and only if the family of columns of A whose indices belong to I is linearly independent. (We cannot quite say that the elements of E are the columns and independence is linear independence, since E might have repeated columns.) More precisely, the function χ mapping i to the ith column is a representation of M over F. How do the elementary operations affect the matroid representation?

We see that row operations on A don't change M but replace the representation χ by an equivalent representation. (Two representations are called *equivalent* if they differ by an invertible linear transformation of the embedding vector space.) On the other hand, column permutations and scalar multiplications replace M by an isomorphic matroid; effectively, permutations re-label the elements, while scalar multiplications have no effect at all.

So, if we call two matrices A and A' *CM-equivalent* if A' is obtained from A by a row operation and a monomial transformation of the columns, we see that CM-equivalence classes of matroids correspond bijectively to both monomial equivalence classes of linear codes, and equivalence classes of representations of matroids, under the natural notions of equivalence in each case.

Thus we expect information to transfer back and forth between code and matroid.

It is possible to go directly from the vector matroid to the code, without the intervening matrix, as follows.

Let $v_1, \ldots, v_n$ be vectors spanning the vector space V. The corresponding code is

$$\{(v_1 f, \ldots, v_n f) : f \in V^*\},$$

where V^* is the dual space of V, and vf is the image of v under f. THis is because the function giving the ith coordinate of a vector is an element of the dual space, and these functions form a basis for the dual space.

I leave as an exercise the problem of finding a matrix-free construction of the matroid from the code.

It is a simple exercise to show the following:

Proposition 4.1 *If the matroid M corresponds to the code C, then the dual matroid M^* corresponds to the dual code $C^\perp$.*

Proof If the matrix A happens to be in the form $[I_k\ B]$, where I_k is a $k \times k$ identity matrix and B is $k \times n-k$, then both the dual code and the dual matroid are represented by the matrix $[-B^\top\ I_{n-k}]$.

4.2 Greene's Theorem

The following theorem was proved by Greene [16].

Theorem 4.2 *Let C be a code over a field with q elements, and M the corresponding vector matroid. Then*

$$W_C(x,y) = y^{n-\dim(C)}(x-y)^{\dim(C)} T\left(M; \frac{x+(q-1)y}{x-y}, \frac{x}{y}\right).$$

Note that, if $X = (x+(q-1)y)/(x-y)$ and $Y = x/y$, then

$$(X-1)(Y-1) = q.$$

So the weight enumerator is an evaluation of the Tutte polynomial along a particular hyperbola in the "Tutte plane".

Proof The proof is by induction. For M, we have the "deletion-contraction rule" of Theorem 3.3.

The analogues of deletion and contraction of a matroid are the operations of *puncturing* and *shortening* a code.

To puncture a code at the ith position, we simply delete the ith coordinate from all codewords. To shorten it at the ith position, we take the subcode consisting of all codewords with zero in the ith position, and then delete this position. We denote by C' and C'' the codes obtained by puncturing and shortening C in a specified position. It is easy to see that puncturing and shortening correspond to deletion and contraction of the corresponding element of the matroid.

A loop in the M corresponds to a coordinate where all codewords have the entry 0. A coloop is a bit more complicated, but can be described as a coordinate such that (after row operations) the first entry in that column of the generator matrix is 1, while all other entries in that column or in the first row are 0.

If the element e of the matroid corresponds to the distinguished coordinate, we have the following recursive scheme for the weight enumerator:

(a) If C has length 0, then $W_C(X,Y) = 1$.

(b) If e is a loop, then $W_C(X,Y) = XW_{C'}(X,Y)$.

(c) If e is a coloop, then $W_C(X,Y) = (X+(q-1)Y)W_{C''}(X,Y)$.

(d) If e is neither a loop or a coloop, then

$$W_C(X,Y) = YW_{C'}(X,Y) + (X-Y)W_{C''}(X,Y).$$

Part (a) is obvious; part (b) holds because each word in C has one extra zero than the corresponding word in C'. Part (c) holds because each word w in C'' gives rise to q words in C (all possible entries occur in the added coordinate), of which one has the same weight as w and $q-1$ have weight one greater.

Finally, suppose that e is neither a loop nor a coloop. Let W_1 and W_2 be the sums of terms in W_C corresponding to words with zero, resp. non-zero, entry in position e. Then $W_C = W_1 + W_2$. We also have $W_{C'} = W_1/X + W_2/Y$, and $W_{C''} = W_1/X$. The assertion follows.

Now induction, together with Theorem 3.3, proves the theorem.

From Theorem 4.2 and Proposition 3.5, we can deduce MacWilliams' Theorem 1.4, which shows that the weight enumerator of the dual code $C^\perp$ can be calculated from that of C.

Theorem 4.3

$$W_{C^\perp}(x,y) = \frac{1}{|C|} W_C(x+(q-1)y, x-y).$$

Proof Since $C^\perp$ has dimension $n - \dim(C)$ and corresponds to the dual matroid M^*, we have

$$W_{C^\perp}(X,Y) = Y^{\dim(C)}(X-Y)^{n-\dim(C)} T\left(M; \frac{X}{Y}, \frac{X+(q-1)Y}{X-Y}\right).$$

On the other hand, we have

$$\begin{aligned}\frac{1}{|C|} & W_C(X+(q-1)Y, X-Y) \\ &= q^{-\dim(C)}(X-Y)^{n-\dim(C)}(qY)^{\dim(C)} T\left(M; \frac{qX}{qY}, \frac{X+(q-1)Y}{X-Y}\right).\end{aligned}$$

The two expressions are equal.

Note that this proof is entirely combinatorial, in contrast to the algebraic proof given in Chapter 1.

4.3 Is there a $\mathbb{Z}_4$ version?

There does not seem to be any way to produce a matroid which captures a $\mathbb{Z}_4$-linear code in the way that we have seen for linear codes over fields. However, we already saw that many features of a $\mathbb{Z}_4$-linear code C are captured by a pair (C_1, C_2) of binary codes. On this basis, Rutherford [27] considered certain

pairs of matroids, and attached to such pairs a three-variable analogue of the Tutte polynomial. This polynomial gives some features of C as specialisations.

The following section is entirely based on her work.

Let M_1 and M_2 be two matroids on the same set E. We say that (M_1, M_2) is a *matroid pair*, or that M_1 is a *quotient* of M_2, if there is a matroid N on a set $E \cup X$ such that $M_1 = N/X$ and $M_2 = N\backslash X$. (Deleting or contracting a set of points just means deleting or contracting the points one at a time.)

It can be shown that we may choose N and X so that $|X| = \rho_{M_2}(E) - \rho_{M_1}(E)$, and X is independent and E is spanning in N. Note that every set $A \subseteq E$ which is independent in M_1 is also independent in M_2. This condition however is not sufficient for (M_1, M_2) to be a matroid pair: see Exercise 4.3.

It is true that

(a) of (M_1, M_2) is a matroid pair, then so is (M_2^*, M_1^*);

(b) for any matroid M on n elements, (M, F_n) and (F_n^*, M) are matroid pairs.

Let (C_1, C_2) be a chain of binary codes, and M_1 and M_2 the associated matroids. Then (M_1, M_2) form a matroid pair. This is because we can find matrices A and B such that A and $\binom{A}{B}$ are generator matrices for C_1 and C_2 respectively; then we can take N to correspond to the code with generator matrix

$$\begin{pmatrix} O & A \\ I & B \end{pmatrix}.$$

In this case, we call the pair (M_1, M_2) a *matroid chain* over $\mathbb{Z}_2$.

Note that not every matroid pair is a matroid chain, even if the individual matroids are representable: see Exercise 4.3. Note also that, if (M_1, M_2) is a matroid chain over $\mathbb{Z}_2$, then so is (M_2^*, M_1^*).

Let $\mathcal{M} = (M_1, M_2)$ be a matroid pair. Rutherford defines the *generalised rank polynomial*, which I shall call for brevity the *Rutherford polynomial*, of the pair to be

$$G(\mathcal{M}; v, x, y) = \sum_{A \subseteq B \subseteq E} v^{|B|-|A|} x^{\rho_1 E - \rho_1 B} y^{|A| - \rho_2 A},$$

where ρ_1 and ρ_2 are the rank functions of M_1 and M_2. (In fact, we could use the analogue of the Tutte polynomial rather than the rank polynomial, by putting $v-1, x-1, y-1$ in place of v, x, y here, but the difference is inessential; I have chosen to follow Rutherford.)

The Rutherford polynomial has a number of interesting specialisations:

Theorem 4.4 *Let $G(\mathcal{M}; v, x, y)$ be the Rutherford polynomial of a matroid pair $\mathcal{M} = (M_1, M_2)$. Let $R(M_i; x, y)$ be the rank polynomial of M_i, for $i = 1, 2$.*

(a) $G(\mathcal{M};v,x,1) = (1+v)^{\rho_1 E} R(M_1; \frac{x}{1+v}, 1+v)$.

(b) $G(\mathcal{M};v,1,y) = (1+v)^{|E|-\rho_2 E} R(M_2; 1+v, \frac{y}{1+v})$.

(c) If $M_1 = M_2 = M$, *then* $G(\mathcal{M};0,x,y) = R(M;x,y)$.

(d) If $M_1 = F_n^*$ *and* $M_2 = M$, *then* $G(\mathcal{M};0,x,y) = R(M;1,y)$.

(e) If $M_2 = F_n$ *and* $M_1 = M$, *then* $G(\mathcal{M};0,x,y) = R(M;x,1)$.

However, the fact that a chain of binary codes does not even determine the symmetrised weight enumerator (or Lee weight enumerator) of the corresponding $\mathbb{Z}_4$-linear code shows that we cannot obtain these weight enumerators from the Rutherford polynomial by specialisation.

It seems likely that the Rutherford polynomial can be extended to chains of arbitrary length of codes over arbitrary fields.

Exercises

4.1. Describe the matroids corresponding to the Hamming code of Chapter 1 and its dual.

4.2. Show that the matroid associated to a linear code is uniform if and only if the code is MDS. (See Exercise 1.3.)

4.3. Find an example of two matroids M_1 and M_2 on a set E such that every independent set in M_1 is independent in M_2 but (M_1, M_2) is not a matroid pair.

4.4. Find an example of two matroids M_1 and M_2 on a set E such that both M_1 and M_2 are representable over $\mathbb{Z}_2$ and (M_1, M_2) is a matroid pair but not a matroid chain over $\mathbb{Z}_2$.

4.5. Let (M_1, M_2) be a matroid pair on E, and let ρ_i be the rank function of M_i for $i = 1, 2$. Prove that

$$0 \leq \rho_2 A - \rho_1 A \leq \rho_2 E - \rho_1 E$$

for any set $A \subseteq E$.

4.6. Calculate the weight enumerator of the code associated with a representation of $U_{3,n}$ over GF(q). Find examples with $n = q+1$.

CHAPTER 5

Permutation groups

In the second half of the notes, we introduce the last strand, permutation groups, and braid it together with codes and matroids.

Traditionally, permutation groups arise as automorphism groups of algebraic or combinatorial structures. The procedure here will be a bit different: the groups will be built from the algebraic structure of codes, and matroids will arise from the fixed point structure of permutation groups.

Before this, we give a brief account of permutation groups and their associated cycle index polynomials.

The treatment here is somewhat brief, since full accounts are available elsewhere. In addition to the classic treatments by Wielandt [32] and Passman [26], there are more recent books by Cameron [4] and Dixon and Mortimer [13].

5.1 Orbits and stabiliser

The set of all permutations of a set Ω is called the *symmetric group* on Ω. Usually we take Ω to be the set $\{1,\ldots,n\}$, and denote the symmetric group by S_n, for some positive integer n. The order of S_n is $n!$.

The convention of using Ω for the permutation domain and lower-case Greek letters for its elements was established by Wielandt in his book. We also use the convention that permutations act on the right, so that the image of α under the permutation g is denoted by αg. Thus, the result of applying the permutation g followed by h is written gh, and we have $\alpha(gh) = (\alpha g)h$.

As is well known, any permutation can be written as a product of disjoint cycles: we call this the *cycle decomposition*. For example, the permutation of

$\{1,\ldots,5\}$ which maps 1 to 4, 2 to 5, 3 to 1, 4 to 3, and 5 to 2 has cycle decomposition $(1,4,3)(2,5)$. The cycle decomposition is unique up to writing the cycles in a different order and starting them at different points: for example,

$$(1,4,3)(2,5) = (5,2)(3,1,4).$$

If we represent the permutation g as a function digraph, with edges $(\alpha,\alpha g)$ for all $\alpha \in \Omega$, the digraph has in-degree and out-degree 1 and so is a disjoint union of cycles; this is precisely the cycle decomposition.

The generalisation to permutation groups is the *orbit decomposition*, which we now discuss.

A *permutation group* G on a set Ω is a subgroup of the symmetric group on Ω; that is, it is a set of permutations closed under composition and inversion and containing the identity permutation. The *degree* of the permutation group G is $|\Omega|$.

Let G be a permutation group on Ω. Define a relation $\sim_G$ on Ω by the rule that $\alpha \sim_G \beta$ if there exists $g \in G$ with $\alpha g = \beta$. It is easy to see that $\sim_G$ is an equivalence relation; the reflexive, symmetric and transitive laws follow from the identity, inverse, and composition properties of G. The equivalence classes of $\sim_G$ are called the *orbits* of G,, and G is said to be *transitive* if there is a single orbit, *intransitive* otherwise.

Note that G is transitive if and only if, for all $\alpha,\beta \in \Omega$, there exists $g \in G$ which maps α to β.

The *stabiliser* of a point $\alpha \in \Omega$ is the subgroup

$$G_\alpha = \{g \in G : \alpha g = \alpha\}.$$

Now, if β is any point of Ω, then the set

$$X(\alpha,\beta) = \{g \in G;: \alpha g = \beta\}$$

is either empty (if α and β lie in different orbits) or a right coset of G_α. We see that the number of right cosets is equal to the size of the orbit. This is the *Orbit-Stabiliser Theorem*:

Theorem 5.1 *Let Δ be an orbit of the permutation group G, and α a point of Δ. Then*

$$|G_\alpha| \cdot |\Delta| = |G|.$$

But this is more than a counting result. Suppose that the group G acts as a permutation group on two different sets Ω_1 and Ω_2. We say that the actions are *isomorphic* if there is a bijection $\theta : \Omega_1 \to \Omega_2$ such that

$$(\alpha\theta)g = (\alpha g)\theta$$

for all $\alpha \in \Omega_1$ and $g \in G$. In other words, if we identify Ω_1 and Ω_2 according to the bijection θ, then the permutations corresponding to any group element are identical.

Now let H be a subgroup of G. The *coset space* $G : H$ is defined to be the set of right cosets of H in G. Now G acts as a permutation group on $G : H$ by the following rule: group element g acts as the permutation $Hx \mapsto Hxg$. (This is clearly well defined, independent of the choice of coset representative x.) Now the refined version of the Orbit-Stabiliser Theorem states:

Theorem 5.2 *Let Δ be an orbit of the permutation group G, and α a point of Δ. Then the actions of G on Δ and on the coset space $G : G_\alpha$ are isomorphic.*

The isomorphism is given by $\beta\theta = X(\alpha, \beta)$ in the earlier notation. The proof is an exercise.

It can also be shown that two coset spaces $G : H$ and $G : K$ provide isomorphic actions of G if and only if the subgroups H and K are conjugate, that is, $K = g^{-1}Hg$ for some $g \in G$.

Thus, to classify the transitive actions of G up to isomorphism, we list a set of representatives of the conjugacy classes of subgroups, and form the coset spaces. To classify all permutation actions, we take arbitrary disjoint unions of the transitive ones.

A permutation group G is *semiregular* if the stabiliser of any point is the identity. It is *regular* if it is semiregular and transitive. By Theorem 5.2, any regular action of G is isomorphic to the action of G on itself by right multiplication (with $\Omega = G$, where $g \in G$ induces the permutation $x \mapsto xg$).

Let G be a permutation group on Ω. Suppose that the set $\Delta \subseteq \Omega$ is invariant under G (that is, fixed setwise – this happens if and only if Δ is a union of orbits of G). Then G^Δ denotes the group of permutations of Δ induced by elements of G. It is a homomorphic image of G; the kernel of the homomorphism is the set of permutations which fix every point in Δ.

Now suppose that $(\Delta_i : i \in I)$ are the orbits of G. For each i, let $G_i = G^{\Delta_i}$. The permutation groups G_i are called the *transitive constituents* of G.

Then G is a subgroup of the *Cartesian product* $\prod_{i \in I} G_i$ of the subgroups G_i. Since we are only concerned with finite permutation groups, the set I is finite, and the Cartesian product is more usually referred to as the *direct product*, and written

$$G_1 \times \cdots \times G_r,$$

where $I = \{1, \ldots, r\}$. Note that G may not be equal to the direct product! In this sense, the orbit decomposition allows many questions about permutation groups to be "reduced" to questions about transitive groups, but there is a difficulty going back: a permutation group is not uniquely determined by its transitive constituents.

Now let Δ be any subset of Ω. Then G_Δ denotes the *setwise stabiliser* of Δ, the set of permutations which map the set Δ to itself; and $G_{(\Delta)}$ denotes the *pointwise stabiliser* of Δ, the set of permutations which fix every point of Ω, so that $G_{(\Delta)} = \bigcap_{\alpha\in\Delta} G_\alpha$).

Thus G_Δ^Δ is the permutation group induced on Δ by its setwise stabiliser in G, and is isomorphic to $G_\Delta/G_{(\Delta)}$. This group will be important in the final chapter. To avoid the double subscript, we denote it by $G[\Delta]$.

We conclude this section with another piece of terminology. Let G_i be a permutation group on Ω_i for $i = 1,2$. We say that G_1 and G_2 are *isomorphic as permutation groups* if there is a bijection $\theta : \Omega_1 \to \Omega_2$ and a group isomorphism $\phi : G_1 \to G_2$ such that

$$(\alpha g)\theta = (\alpha\theta)(g\phi)$$

for all $\alpha \in \Omega_1$ and $g \in G_1$. If two actions of the same group are isomorphic according to the earlier definition, then the induced permutation groups are isomorphic as permutation groups; but the converse is false, since we now permit an automorphism of G.

For example, let $G = C_2 \times C_2$ be generated by elements a and b. The actions given by

$$a = (1,2), \qquad b = (3,4)$$

and

$$a = (1,2)(3,4), \qquad b = (3,4)$$

are not isomorphic, but their images are isomorphic (indeed, identical) as permutation groups.

5.2 The Orbit-Counting Lemma

The Orbit-Counting Lemma (incorrectly called Burnside's Lemma in much of the literature of combinatorial enumeration) is a simple relationship between fixed points and orbits of a permutation group, which will be crucial in what follows.

Let G be a permutation group on Ω. For $g \in G$, let $\mathrm{fix}(g)$ denote the number of points of Ω fixed by g. Now the Orbit-Counting Lemma states:

Theorem 5.3 *The number of orbits of a permutation group G is equal to the average number of fixed points of its elements: that is, the number of orbits is*

$$\frac{1}{|G|}\sum_{g\in G} \mathrm{fix}(g).$$

Proof Construct a bipartite graph as follows. The vertex set is $\Omega \cup G$; there is an edge from $\alpha \in \Omega$ to $g \in G$ if and only if g fixes α.

To prove the theorem, we count the edges of the graph in two different ways. Clearly the vertex g lies on $\mathrm{fix}(g)$ edges, and so the number of edges is

$$\sum_{g \in G} \mathrm{fix}(g).$$

On the other hand, the vertex α lies on $|G_\alpha|$ edges. By the Orbit-Stabiliser Theorem 5.1, if Δ is the orbit containing α, then

$$|G_\alpha| \cdot |\Delta| = |G|,$$

so the number of edges containing a vertex in Δ is equal to $|G|$, and the total number of edges is $|G|$ times the number of orbits.

Equating these two numbers gives the result.

For example, the symmetric group S_4 contains one element with four fixed points; six elements (the transpositions) with two fixed points; eight elements (the 3-cycles) with one fixed point; and nine elements (the 4-cycles and the double transpositions) with no fixed points. So the number of orbits is

$$\frac{1}{24}(1 \cdot 4 + 6 \cdot 2 + 8 \cdot 1 + 9 \cdot 0) = 1.$$

Corollary 5.4 *If G is a transitive permutation group of degree $n > 1$, then G contains an element with no fixed points.*

Proof The average number of fixed points is one; the identity fixes more than one point; so some element fixes fewer than one.

This result is due to Jordan. Despite its simplicity, it has a variety of applications in number theory and topology: a recent paper of Serre [28] describes some of these.

In combinatorial enumeration, it is often the case that being able to count the members of a set and being able to choose one at random are closely related. This principle applies to the Orbit-Counting Lemma, as observed by Mark Jerrum [20].

Consider the following Markov chain, defined on the elements of Ω. In one step, we move from a point α to a randomly chosen neighbour of α in the bipartite graph of Theorem 5.3 (that is, an element $g \in G_\alpha$), and then to a randomly chosen neighbour β of g (that is, a fixed point of g).

Since it is possible to move from any point α to any point β in a single step (via the identity of G), the chain is irreducible and aperiodic; so there is a unique limiting distribution, to which it converges from any initial distribution. This distribution is easily seen to have the property that the probability of α is inversely proportional to the size of the orbit containing α; in other words, the limiting distribution is uniform on orbits.

It is important to know the mixing time of such a Markov chain, that is, how rapidly it approaches its limit, and in particular to characterise the permutation groups for which the chain is rapidly mixing. Very little is known about this!

5.3 Bases and strong generating sets

In practice, one needs a computer to investigate permutation groups. Even groups of moderate degree can be very large, and finding interesting subgroups by hand if we are given a set of permutations generating the group is a daunting task. On the other hand, there are very efficient algorithms for computing with permutation groups, and it is possible to study groups with degrees in the tens of thousands without too much trouble.

In this section, we take the first steps in computational permutation group theory. We are given a set S of permutations which generate a subgroup G of S_n, and we want to be able to do such things as find the order of G, choose a random element of G (from the uniform distribution), or test an element of S_n for membership in G.

The first thing we can do is to find the orbits of G. For consider the directed graph on Ω with edges $(\alpha, \alpha s)$ for all $\alpha \in \Omega$ and $s \in S$. The orbits are precisely the connected components of this graph. Moreover, for each point in the orbit of α, we can find a *witness*, an element of G (in the form of a word in the generators S) mapping α to β. These witnesses form a set X of coset representatives for G_α in G.

Next, a lemma of Schreier shows that, if generators of a group and coset representatives for a subgroup are known, then generators for the subgroup can be computed.

Now we apply this procedure recursively until the group is trivial. At this point, what we have found is the following:

(a) a *base* for G; that is, a sequence $(\alpha_1, \ldots, \alpha_r)$ of points of Ω whose pointwise stabiliser is the identity;

(b) for $i = 1, \ldots, r$, a set X_i of coset representatives for G_{i-1} in G_i, where G_i is the pointwise stabiliser of $(\alpha_1, \ldots, \alpha_i)$.

Now this information enables us to settle the above questions. We begin with

the membership test. Suppose that a permutation $g \in S_n$ is given. If G is the trivial group, we can decide immediately whether $g \in G$. Suppose not. We compute $\alpha_1 g$. If this is not in the G-orbit of α_1, then $g \notin G$, and we are done. Otherwise, there is a unique $x_1 \in X_1$ such that $\alpha_1 g = \alpha_1 x_1$. Now gx_1^{-1} fixes α_1, and we apply the test recursively to decide whether $gx_1^{-1} \in G_1$; for we have $g \in G$ if and only if $gx_1^{-1} \in G_1$ in this case.

If the test succeeds, then we will eventually find that

$$gx_1^{-1} \cdots x_r^{-1} = 1,$$

that is, $g = x_r \cdots x_1$, with $x_i \in X_i$ for $i = 1, \ldots, r$. This expression is unique, so

$$|G| = |X_r| \cdots |X_1|,$$

and we have found the order of G. This can also be seen by noting that $|X_i| = |G_{i-1} : G_i|$, and of course

$$|G| = |G_0 : G_1| \cdots |G_{r-1} : G_r|,$$

since $G_0 = G$ and $G_r = 1$.

The equation

$$G = X_r \cdots X_1$$

shows that the union of the sets $X_1, \ldots, X_r$ generates G; similarly, for any i, the set $X_{i+1} \cup \cdots \cup X_r$ generates G_i. The set $X = X_1 \cup \cdots \cup X_r$ is called a *strong generating set* for G.

The unique representation also shows that if we choose elements uniformly and independently at random from $X_r, \ldots, X_1$ and multiply them, we obtain a uniform random element of G.

We will have more to say about bases later, so we pursue the subject a little further here. First, we note another property of bases relevant to computational group theory. Any element $g \in G$ is determined uniquely by the image of a base B under g; for, if $Bg = Bh$, then $Bgh^{-1} = 1$, so that $gh^{-1} = 1$ (by definition of a base), and $g = h$. Thus, it is of interest to find the smallest possible base. Unfortunately, Kenneth Blaha [1] showed that this problem is NP-complete in general; but there are some things we can say.

When we are choosing a base, there is clearly no point in choosing a point α_i which is fixed by the stabiliser of its predecessors. So we call a base *irredundant* if no base point is fixed by the stabiliser of its predecessors. Usually we consider only irredundant bases.

Unlike vector spaces, permutation groups can have bases of different cardinalities. Consider, for example, the group C_2^n, acting with $n+1$ orbits as follows: for

each $i \leq n$, an orbit O_i of size 2 on which all generators except the ith act trivially; and an orbit O_0 of length 2^n on which the group acts regularly. Choose $\alpha_i \in O_i$ for each i. Then, for $1 \leq i \leq n$, there is an irredundant base of size i of the form $(\alpha_1, \ldots, \alpha_{i-1}, \alpha_0)$.

On the other hand, there are some restrictions:

Proposition 5.5 *The number r of elements in an irredundant base for a permutation group G of degree n satisfies*

$$\log|G|/\log n \leq r \leq \log|G|/\log 2.$$

Proof We have

$$|G| = |G_0 : G_1| \cdots |G_{r-1} : G_r|.$$

Each index $|G_{i-1} : G_i|$ is at least 2 (since the base is irredundant) and at most n (since it is the length of an orbit of G_{i-1}). So

$$2^r \leq |G| \leq n^r,$$

and we are done.

Our earlier example shows that in general no substantial improvement can be made.

5.4 Primitivity and multiple transitivity

Some transitive groups can be further "reduced".

Let G be a transitive permutation group on Ω. A *G-congruence* is an equivalence relation on Ω which is preserved by G. Its equivalence classes form a partition of Ω whose parts are permuted among themselves by G. The set of equivalence classes is called a *system of imprimitivity*, and the classes are *blocks of imprimitivity*.

A congruence (or the associated system or blocks of imprimitivity) is called *trivial* if either it is the relation of equality, or it is the "universal" relation $\Omega \times \Omega$. Every group preserves the trivial congruences. If there is a non-trivial G-congruence, then G is said to be *imprimitive*; otherwise it is *primitive*.

Note that we have defined these terms only for transitive permutation groups (see Exercise 5.5). Thus, all the equivalence classes of a G-congruence have the same size. In particular, any transitive permutation group of prime degree is primitive.

If G is imprimitive, let S be a system of imprimitivity, and B one of its blocks. From G, we construct two smaller permutation groups:

(a) $K = G^S$, the group of permutations of S induced by G;

(b) $H = G_B^B = G[B]$, the group of permutations of B induced by its setwise stabiliser in G.

Each of these groups is transitive, and it can be shown that G is isomorphic to a subgroup of the *wreath product* $H \operatorname{Wr} K$. Continuing this reduction if H or K is imprimitive, we end up with a sequence of primitive groups called the *primitive components* of G.

The next result gives some basic properties of primitive groups.

Proposition 5.6 *(a) A transitive permutation group G is primitive if and only if G_α is a maximal subgroup of G.*

(b) Let N be a non-trivial normal subgroup of the transitive group G. Then the orbits of N form a system of imprimitivity for G. In particular, if G is primitive, then any non-trivial normal subgroup of G is transitive.

Let t be a positive integer, at most $|\Omega|$. The permutation group G on Ω is said to be *t-transitive* if we can map any t-tuple of distinct elements of Ω to any other such t-tuple by some element of G. We say that G is *multiply transitive* if it is t-transitive for some $t > 1$.

The problem of determining the multiply transitive permutation groups goes back to the origins of group theory in the nineteenth century: Galois knew of the existence of 2-transitive groups $\mathrm{PSL}(2,p)$, and Mathieu constructed 5-transitive groups M_{12} and M_{24}. The condition of t-transitivity becomes stronger as t increases. The symmetric group S_n is n-transitive, and the alternating group A_n is $(n-2)$-transitive.

However, a definitive result had to wait for the Classification of Finite Simple Groups (CFSG), as we will see in the next section. Using this classification, all multiply transitive groups have been determined. In particular, the only 5-transitive groups apart from symmetric and alternating groups are the two Mathieu groups mentioned above.

5.5 Modern permutation group theory

The title of this section is taken from a talk by Michael Aschbacher to the London Mathematical Society in 2001. Aschbacher's theme was that many questions about finite permutation groups can be reduced to questions about almost simple groups (where a group is said to be *almost simple* if it is an extension of a non-abelian simple group by a subgroup of its outer automorphism group). Now the finite simple groups have been classified (though a complete proof has not yet

been published, so the proof of this claim is not open to scrutiny), and detailed properties of the known simple groups have been worked out, so such questions can often be settled.

The Classification of Finite Simple Groups, which we abbreviate to CFSG, is an enormously complicated theorem; the first complete published proof will cover many thousands of pages. So for several reasons it is prudent to label clearly a result proved using CFSG. See Gorenstein [15] for an introduction to the finite simple groups and to the proof of CFSG.

The reduction works as follows. We have seen a reduction from arbitrary permutation groups to transitive ones, and from transitive groups to primitive ones. Now let G be a primitive permutation group on Ω. We say that G is *non-basic* if there is an identification of Ω with F^n for some set F and some positive integer n, such that the following is true:

> each element of G has the form
>
> $$(a_1,\ldots,a_n) \mapsto (a_{1h}g_1,\ldots,a_{nh}g_n),$$
>
> where h is a permutation of $\{1,\ldots,n\}$, and $g_1,\ldots,g_n$ are permutations of F.

In other words, G preserves a non-trivial "power structure" on Ω. We say that G is *basic* if it is not non-basic.

This definition is similar in structure to that of transitive and primitive groups: a permutation group is transitive if it preserves no non-trivial subset of Ω, and a transitive group is primitive if it preserves no non-trivial partition.

Now part of the O'Nan–Scott Theorem is the following assertion:

Theorem 5.7 *A basic primitive permutation group is affine, diagonal, or almost simple.*

Here a permutation group is *affine* if (up to re-labelling the set Ω) it is a subgroup of the group

$$\{v \mapsto vA + c : A \in \mathrm{GL}(V), c \in V\}$$

of permutations of the finite vector space V and contains all the translations $v \mapsto v + c$. A *diagonal* group has a normal subgroup T^n, where T is a non-abelian simple group and $n \geq 2$, acting on the set of right cosets of the *diagonal subgroup*

$$D = \{(t,t,\ldots,t) : t \in T\}.$$

Almost simple groups were defined earlier.

Now we consider what kind of information about the finite simple groups is needed to understand basic permutation groups.

(a) If G is affine, then G is the semi-direct product of the translation group of V by an irreducible subgroup H of $\mathrm{GL}(V)$. A similar reduction theorem, due to Aschbacher, for irreducible linear groups shows that we can further reduce to the case where the centre $Z(H)$ of H consists of scalar transformations and $H/Z(H)$ is almost simple. Typically we now require properties about the irreducible projective representations of almost simple groups.

(b) If G is diagonal, then its properties can usually be derived from routine properties of simple groups.

(c) In the case where G is almost simple, we need to know about primitive permutation actions (equivalently, maximal subgroups) of almost simple groups.

Many results about primitive permutation groups have been proved by this method. We restrict ourselves to two applications. The first application is the classification of the 2-transitive groups. In this case, a very simple form of the O'Nan–Scott theorem (proved originally by Burnside) shows that a 2-transitive group is either affine or almost simple. We refer to Cameron [4] and Dixon and Mortimer [13] for the list of 2-transitive groups and for further details of the argument.

The second, more recent result is a composite theorem about almost simple primitive groups. The first part is due to Cameron and Kantor [8], the second to Liebeck and Shalev [21].

Theorem 5.8 *(CFSG) There are absolute constants c_1, c_2 with the following properties. Let G be an almost simple primitive permutation group of degree n. Suppose that G is not one of the following:*

(i) a symmetric or alternating group S_m or A_m, acting on the set of k-element subsets of $\{1,\ldots,m\}$ (with $n = \binom{m}{k}$);

(ii) a symmetric or alternating group S_m or A_m, acting on the set of partitions of $\{1,\ldots,m\}$ into l parts of size k, where $kl = m$;

(iii) a classical group, acting on an orbit of subspaces of its natural module.

Then

(a) $|G| \leq n^{c_1}$;

(b) G has a base of size at most c_2.

This theorem is in many ways typical of applications of CFSG to permutation group theory: a group is either "known" (in some more-or-less precise sense) or "small".

We saw that the size of an irredundant base for a permutation group G of degree n lies between $\log|G|/\log n$ and $\log|G|/\log 2$. For primitive groups, Laci Pyber has conjectured that the lower bound is approximately correct; more specifically, the minimal base size is at most $c\log|G|/\log n$, for some constant c. Part (b) of the above theorem is a result in the direction of this conjecture.

Exercises

5.1. Show that the number of ways of writing the cycle decomposition of a permutation $g \in S_n$ is equal to the order of the centraliser of g in S_n (the subgroup of elements commuting with g). Find a formula for this number.

5.2. The Orbit-counting Lemma asserts that the expected value of the number of fixed points of a random element of the permutation group G is equal to the number of orbits of G. What is the variance of this number?

5.3. Let G be a transitive permutation group on Ω. Let B be a non-empty subset of Ω with the property that, for all $g \in G$, either $Bg = B$ or $B \cap Bg = \emptyset$. Prove that B is a block of imprimitivity.

5.4. Suppose that $|\Omega| > 2$, and let G be a permutation group on Ω which preserves no non-trivial equivalence relation. Prove that G is transitive (and hence primitive).

5.5. Prove Proposition 5.6. Is it true that every block of imprimitivity for a transitive group G is an orbit of a normal subgroup of G?

5.6. Find a base and strong generating set for the permutation group on the set $\{1,2,3,4,5\}$ generated by $s = (1,2)(4,5)$ and $t = (2,3)(4,5)$. Hence find the order of this group, and determine whether it contains $(1,2,3)(4,5)$.

5.7. Prove that the permutation group G, of degree at least 2, is 2-transitive if and only if

$$\frac{1}{|G|}\sum_{g\in G} \mathrm{fix}(g)^2 = 2.$$

Generalise.

5.8. Find all systems of imprimitivity for $G = S_4$ acting on the set Ω of ordered pairs of distinct elements of $\{1,2,3,4\}$. Hence show that the primitive components of a transitive group are not uniquely determined.

5.9. A permutation group G is *sharply t-transitive* if, given any two ordered t-tuples of distinct elements of Ω, there is a *unique* element of G carrying the first pair to the second.

Prove that, in a sharply 2-transitive group G, the identity and the fixed-point-free permutations form a normal subgroup N. Show further that N is elementary abelian, and deduce that the degree of G is a prime power. Deduce that, if $t \geq 2$, then the degree of a sharply t-transitive group is of the form $p^r + t - 2$ for some prime power p^r.

Construct a sharply 2-transitive group of degree p^r for any prime power p^r.

5.10. Prove the following strengthening of Jordan's Theorem (Corollary 5.4), due to Cameron and Cohen [6]:

> Let G be a transitive permutation group of degree $n > 1$. Then at least a proportion $1/n$ of the elements of G are fixed-point-free. Equality holds if and only if G is sharply 2-transitive.

5.11. Prove that the permutations $(1,2)(3,4)(5,6)(7,8)(9,10)(11,12)$ and $(1,2,3)(4,5,7)(8,9,11)$ generate a sharply 5-transitive group of degree 12. (This is the *Mathieu group* M_{12}.)

CHAPTER 6

Cycle index

The cycle index is a polynomial associated with a permutation group. Unlike the polynomials we considered earlier for codes and matroid, it has many variables (possibly as many as the degree of the permutation group). To clarify the process of substituting into a multivariate polynomial F in indeterminates $s_1, \ldots, s_n$, we use the notation

$$F(s_i \leftarrow t_i)$$

for the result of substituting the term t_i for s_i for $i = 1, \ldots, n$.

The cycle index is basic in the theory of combinatorial enumeration pioneered by Redfield and Pólya. We refer to Harary and Palmer [18] for a more detailed account.

6.1 Definition

Let G be a permutation group on a set Ω, where $|\Omega| = n$. For each element $g \in G$, we can decompose the permutation g into a product of disjoint cycles; let $c_i(g)$ be the number of i-cycles occurring in this decomposition. Now the *cycle index* of G is the polynomial $Z(G)$ in indeterminates $s_1, \ldots, s_n$ given by

$$Z(G) = \frac{1}{|G|} \sum_{g \in G} s_1^{c_1(g)} \cdots s_n^{c_n(g)}.$$

This can be regarded as a multivariate probability generating function for the cycle structure of a random element of G (chosen from the uniform distribution). In particular,

$$P_G(x) = Z(G)(s_1 \leftarrow x, s_i \leftarrow 1 \text{ for } i > 1)$$

is the probability generating function for the number of fixed points of a random element of G, so that substituting $x \leftarrow 0$ gives the proportion of derangements in G. In other words,

$$P_G(x) = \frac{1}{|G|} \sum_{g \in G} x^{c_1(g)}.$$

The number $c_1(g)$ is the number of fixed points of g, which we called $\mathrm{fix}(g)$ in Chapter 5; the function $g \mapsto c_1(g)$ is the *permutation character* of G.

Let us work two examples. First, let G be the symmetric group of degree 4. Each partition of 4 is the cycle type of some element of G, and it is not hard to count the number of elements corresponding to each partition:

Partition	4	31	22	211	1111
Number	6	8	3	6	1

So

$$Z(G) = \frac{1}{24}(6s_4 + 8s_1s_3 + 3s_2^2 + 6s_2s_1^2 + s_1^4).$$

Now let us take the same group S_4 acting on the set of 2-element subsets of $\{1,2,3,4\}$. We simply need to find for each shape of permutation the cycle structure on the set of pairs; we obtain the following:

On points	4	31	22	211	1111
On pairs	42	33	2211	2211	111111

So

$$Z(G) = \frac{1}{24}(6s_2s_4 + 8s_3^2 + 9s_1^2s_2^2 + s_1^6).$$

6.2 The cycle index theorem

The cycle index is an important tool in combinatorial enumeration. Typically, we have a collection of "figures" decorating some set (e.g. colours of the faces of a regular polyhedron), and we are interested in counting the number of configurations up to some notion of symmetry (given by a group of automorphisms of the set). More formally, let A be a set of "figures", each of which has a non-negative integer "weight". The number of figures may be infinite, but we assume that there are only finitely many figures of any given weight. The *figure-counting series* of A is the formal power series

$$A(t) = \sum_{n \geq 0} a_n t^n,$$

where a_n is the number of figures of weight n; that is, it is just the generating function for figures by weight.

Now let Ω be a finite set. A function $f : \Omega \to A$ has a *weight* given by

$$\mathrm{wt}(f) = \sum_{\alpha \in \Omega} \mathrm{wt}(f(\alpha)).$$

If G is a permutation group on Ω, then there is a natural action of G on the set of functions, given by the rule

$$f^g(\alpha) = f(\alpha g^{-1}).$$

(The inverse is required to make this a good definition of an action.) Clearly this action preserves the weight of a function. The *function-counting series* is the formal power series

$$B(t) = \sum_{n \geq 0} b_n t^n,$$

where b_n is the number of G-orbits on the set of functions of weight n. Now the *Cycle Index Theorem* states:

Theorem 6.1 *With the above notation,*

$$B(t) = Z(G; s_i \leftarrow A(t^i)).$$

Proof Here is a sketch of the proof: fill in the details as an exercise.

The generating function for all functions, disregarding the group action, is $A(t)^n$, since the coefficient of t^m in $A(t)^n$ is equal to the sum of $a_{i_1} \cdots a_{i_n}$ over all expressions $i_1 + \cdots + i_n = m$. Note that $A(t)^n = s_1^n(s_1 \leftarrow A(t))$, and s_1^n is the cycle index of the trivial group.

For any permutation g, let $z(g) = s_1^{c_1(g)} \cdots s_n^{c_n(g)}$. A function is fixed by the permutation g if and only if it is constant on each cycle in the cycle decomposition of g. The weight of such a function is the sum of the products of cycle length and weight of the figure at a point of the cycle. Hence the generating function for the number of functions fixed by g is

$$A(t)^{c_1(g)} \cdots A(t^n)^{c_n(g)} = z(g; s_i \leftarrow A(t^i)).$$

Now the result follows from the Orbit-Counting Lemma (Theorem 5.3) and the definition of cycle index, on averaging over G.

Here is a typical application of the theorem. How many graphs are there on 4 vertices with any given number (from 0 to 6) of edges, up to isomorphism? We take Ω to be the set of all 2-elements of the vertex set $\{1,2,3,4\}$. To each element $\{i, j\}$ of Ω we attach either an edge or a non-edge. Taking edges to have weight 1 and non-edges to have weight 0, the weight of the function is just the total number

of edges in the corresponding graph. Moreover, two graphs are isomorphic if and only if there is a permutation of $\{1,2,3,4\}$ carrying the first function to the second. So, taking the figure-counting series to be $A(t) = 1+t$, and the group S_4 acting on 2-sets (whose cycle index we calculated in the previous section), we find the generating function for graphs on four vertices (enumerated by edges) to be

$$\begin{aligned} & \frac{1}{24}(6(1+t^2)(1+t^4)+8(1+t^3)^2+9(1+t)^2(1+t^2)^2+(1+t)^6) \\ = \;& 1+t+2t^2+3t^3+2t^4+t^5+t^6. \end{aligned}$$

6.3 Some other counting results

Let G be a permutation group on a set Ω. Many counting problems related to G, other than those described in the Cycle Index Theorem, can be solved by specialisations of the cycle index. Here are some examples.

(a) Let F_n be the number of orbits of G acting on the set of all n-tuples of distinct elements of Ω. We consider the *exponential generating function*

$$F_G(t) = \sum_{n\geq 0} \frac{F_n t^n}{n!}$$

for the sequence (F_n). Now we have

$$F_G(t) = Z(G)(s_1 \leftarrow x+1, s_i \leftarrow 1 \text{ for } i > 1).$$

(b) If instead we want the total number F_n^* of orbits of G on n-tuples (with repeats allowed), then it can be calculated as

$$F_n^* = \sum_{k=1}^{n} S(n,k) F_k,$$

where $S(n,k)$ is the *Stirling number of the second kind*, the number of partitions of an n-element set into k parts.

(c) Let f_n be the number of orbits of G acting on the set of all n-element subsets of Ω. Then the *ordinary generating function*

$$f_G(t) = \sum_{n\geq 0} f_n t^n$$

is given by the specialisation

$$f_G(t) = Z(G)(s_i \leftarrow t^i + 1).$$

(d) The *Parker vector* of G is the vector $(p_1, p_2, \ldots)$, where p_k is the number of orbits of G on the set of k-cycles occurring in the cycle decompositions of its elements (and G acts on these k-cycles by conjugation). The Parker vector was introduced by Parker in the context of computational Galois theory, and was studied by Gewurz [14]. It is given by

$$p_k = k[(\partial/\partial s_k)Z(G)](s_i \leftarrow 1).$$

Many of these sequences play an important role in combinatorial enumeration. See [3] for more details about (a)–(c).

6.4 The Shift Theorem

Let G be a permutation group on Ω. For any subset Δ of Ω, we defined $G[\Delta]$ to be the group of permutations of Δ induced by elements of G fixing Δ pointwise. Thus, $G[\Delta]$ is the quotient of the setwise stabiliser of Δ by its pointwise stabiliser.

We let $\mathcal{P}\Omega/G$ denote the set of G-orbits on the power set of Ω; by abuse of notation, this will also be used for a set of orbit representatives.

Now the following result (the *Shift Theorem*) holds:

Theorem 6.2 *For any finite permutation group G on Ω,*

$$\sum_{\Delta \in \mathcal{P}\Omega/G} Z(G[\Delta]) = Z(G; s_i \leftarrow s_i + 1).$$

Proof Rather than a proof of this theorem (which is just elementary but complicated double counting), I will try to explain why it has to hold. (This explanation would be a proof if we knew that the cycle index is the unique polynomial for which the Cycle Index Theorem holds.) Suppose that we have a set A^* of figures containing one distinguished figure $*$ of weight zero. Let $A^*(t)$ be its figure-counting series, and $A(t)$ the figure-counting series of $A = A^* \setminus \{*\}$. Then $A(t) = A^*(t) - 1$, and so the function-counting series is

$$B(t) = Z(G; s_i \leftarrow A(t^i) + 1).$$

Now this can be calculated in another way. Any function f is determined by giving the set $\Delta = \{\alpha \in \Omega : f(\alpha) \neq *\}$ and then the function $f' : \Delta \to A$ given by its restriction to Δ. Two functions lie in the same orbit of G if and only if

(a) the sets Δ lie in the same orbit of G; and

(b) assuming that we have translated by an element of G to make these two sets equal, the functions f' lie in the same orbit of $G[\Delta]$.

So the function-counting series is given by

$$B(t) = \sum_{\Delta \in \mathcal{P}\Omega/G} Z(G[\Delta], s_i \leftarrow A(t^i)).$$

So the two polynomials in the theorem yield the same result for every substitution $s_i \leftarrow A(t^i)$, for any figure-counting series $A(t)$.

This theorem may not seem to have very much use as it stands. One use to which it was put in [3] was to extend the definition of cycle index to certain infinite permutation groups, namely, those which are *oligomorphic*. (A permutation group is said to be oligomorphic if the number f_n of orbits on n-element subsets is finite for all natural numbers n.) The point is that the cycle index of an infinite permutation group cannot be defined directly, since permutations may have infinitely many cycles of some length; but the right-hand side of the Shift Theorem is well-defined for any oligomorphic group, if we interpret $\mathcal{P}\Omega/G$ to be a set of representatives for the orbits on *finite* sets.

Our interest in the theorem is a bit different. A corollary of it is the following result, first observed by Boston *et al.* [2]. Recall that $P_G(x)$ is the probability generating function for fixed points of random elements of G, while $F_G(t)$ is the exponential generating function for the number of orbits of G on n-tuples of distinct elements.

Corollary 6.3 *For any finite permutation group G, we have*

$$F_G(t) = P_G(t+1).$$

Proof We know that

$$P_G(x) = Z(G; s_1 \leftarrow x, s_i \leftarrow 1 \text{ for } i > 1).$$

Also, a set Δ of cardinality n can be labelled in $n!$ different ways; these fall into $n!/|G(\Delta)|$ orbits under G. So we have

$$\begin{aligned} F_G(t) &= \sum_{n \geq 0} \sum_{\Delta \in \mathcal{P}\Omega/G, |\Delta|=n} \frac{n!}{|G(\Delta)|} \frac{t^n}{n!} \\ &= \sum_{\Delta \in \mathcal{P}\Omega/G} Z(G(\Delta), s_1 \leftarrow t, s_i \leftarrow 0 \text{ for } i > 1) \\ &= Z(G; s_1 \leftarrow t+1, s_i \leftarrow 1 \text{ for } i > 1), \end{aligned}$$

the last equality coming from the Shift Theorem. So the result holds.

However, the original proof by Boston *et al.* [2] is more direct. Let $c_1(g)$ denote the number of fixed points of the element g. Then the number of ordered j-tuples of distinct elements it fixes is

$$c_1(g)(c_1(g)-1)\cdots(c_1(g)-j+1).$$

By the Orbit-Counting Lemma,

$$F_j = \frac{1}{|G|}\sum_{g\in G} c_1(g)(c_1(g)-1)\cdots(c_1(g)-j+1).$$

Multiplying by $t^j/j!$ and reversing the order of summation,

$$\begin{aligned} F_G(t) &= \frac{1}{|G|}\sum_{g\in G}\sum_{j=0}^{n}\binom{c_1(g)}{j}t^j \\ &= \frac{1}{|G|}\sum_{g\in G}(t+1)^{c_1(g)} \\ &= P_G(t+1), \end{aligned}$$

since $P_G(x) = \sum_{g\in G} x^{c_1(g)}/|G|$.

Exercises

6.1. Let g be a permutation of Ω, and suppose that the order of g is m. Show that

$$\mathrm{fix}(g^k) = \sum_{l|k} l s_l(g)$$

for all k dividing m, and deduce that

$$s_k(g) = \frac{1}{k}\sum_{l|k}\mu(k/l)\,\mathrm{fix}(g^l)$$

for all k dividing m, where μ is the Möbius function.

6.2. Let G be a permutation group on two sets Ω_1 and Ω_1. Let $\mathrm{fix}_1(g)$ and $\mathrm{fix}_2(g)$ denote the numbers of fixed points of G in Ω_1 and Ω_2 respectively. Suppose that $\mathrm{fix}_1(g) = \mathrm{fix}_2(g)$ for all $g \in G$. Prove that the cycle indices of the two permutation groups G^{Ω_1} and G^{Ω_2} are equal. (Hint: use the preceding exercise.)

6.3. Let G be the group of rotations of a cube.

(a) Prove that G is isomorphic to the symmetric group S_4.

(b) Compute the cycle index of G, acting on the set of faces of the cube.

(c) Is this action of G isomorphic to the action of S_4 on the set of 2-element subsets of $\{1,\ldots,4\}$?

6.4. Find the generating function for colourings of the faces of the cube red and blue, enumerated by the number of red faces.

6.5. Use the Cycle Index Theorem to enumerate

(a) graphs on four vertices having at most one loop at each vertex but no multiple edges, by number of edges;

(b) graphs on four vertices having at most two edges between each pair of distinct vertices but no loops, by number of edges.

6.6. Verify the Shift Theorem for the permutation group S_4 (in its natural action on four points).

6.7. Use the Corollary to the Shift Theorem to calulcate the function $P_G(x)$, where G is the symmetric group S_n. Deduce that the probability that a random permutation has no fixed points tends to $1/\mathrm{e}$ as $n \to \infty$.

6.8. Prove that

$$\sum_{n\geq 0} Z(S_n) = \exp\left(\sum_{i=1}^{\infty} \frac{s_i}{i}\right).$$

CHAPTER 7

Codes and permutation groups

This chapter describes the link between codes and permutation groups. From any linear code, we construct a permutation group, whose cycle index is essentially the weight enumerator of the code. If we start instead with a $\mathbb{Z}_4$-linear code, the cycle index of the group is the symmetrised weight enumerator of the code. Essentially, we "inflate" each coordinate of the code into a copy of the alphabet.

We begin with a technical result concerning a similar operation of "inflating" a matroid, which will be relevant in Chapter 8.

7.1 Inflating a matroid

How does the Tutte polynomial of a matroid change if a single element is replaced by q parallel elements? This can be described explicitly in terms of the Tutte polynomials of the deletion and contraction with respect to that element. However, we need to know what happens if *every* element of the matroid is replaced by a set of q parallel elements, and here the answer is much simpler.

To be more precise, we define the q-fold *inflation* of a matroid M on the set E to be the matroid on the set $E \times Q$, where Q is a q-element set, whose independent sets are as follows: for each independent set $A \subseteq E$, and each function $f : A \to Q$, the set $\{(a, f(a)) : a \in A\}$ of $E \times Q$ is independent; and these are the only independent sets.

Proposition 7.1 *If M_q is a q-fold inflation of M, then*

$$T(M_q;x,y) = \left(\frac{y^q-1}{y-1}\right)^{\rho(E)} T\left(M; \frac{xy-x-y+y^q}{y^q-1}, y^q\right).$$

Proof To each subset $A \subseteq E$, there are 2^q-1 subsets of $E \times Q$ whose projection onto E is A. For any such subset, the rank (in M_q) is equal to the rank of A in Q. The contribution to the Tutte polynomial from such sets is given by

$$\begin{aligned}
& (x-1)^{\rho E-\rho A}(y-1)^{|A|-\rho A}\prod_{i=1}^{|A|}\left(\sum_{j=1}^{q}\binom{q}{j}(y-1)^{j-1}\right) \\
= \;& (x-1)^{\rho E-\rho A}(y-1)^{|A|-\rho A}(y^q-1)^{|A|} \\
= \;& (x-1)^{\rho E-\rho A}\left(\frac{y^q-1}{y-1}\right)^{\rho A}(y^q-1)^{|A|-\rho A} \\
= \;& \left(\frac{y^q-1}{y-1}\right)^{\rho(E)}\left(\frac{(x-1)(y-1)}{y^q-1}\right)^{\rho E-\rho A}(y^q-1)^{|A|-\rho A}.
\end{aligned}$$

Summing over $A \subseteq E$, we obtain

$$\begin{aligned}
T(M_q;x,y) &= \left(\frac{y^q-1}{y-1}\right)^{\rho E}\sum_{A\subseteq E}\left(\frac{(x-1)(y-1)}{y^q-1}\right)^{\rho E-\rho A}(y^q-1)^{|A|-\rho A} \\
&= \left(\frac{y^q-1}{y-1}\right)^{\rho E} T\left(M; \frac{(x-1)(y-1)}{y^q-1}+1, y^q\right).
\end{aligned}$$

7.2 The connection

Let C be a linear $[n,k]$ code over $\mathrm{GF}(q)$. We construct from C a permutation group whose cycle index is (more-or-less) the weight enumerator of C.

The group we construct is the additive group of C. We let it act on the set $\{1,\ldots,n\}\times\mathrm{GF}(q)$ (a set of cardinality nq) in the following way: the codeword $(a_1,\ldots,a_n)$ acts as the permutation

$$(i,x) \mapsto (i, x+a_i)$$

of the set $\{1,\ldots,n\}\times\mathrm{GF}(q)$. The group $G(C)$ is an elementary abelian group of order q^k.

Proposition 7.2 $\dfrac{1}{|C|}W_C(X,Y) = Z(G; s_1 \leftarrow X^{1/q}, s_p \leftarrow Y^{p/q})$, *where q is a power of the prime number p.*

Proof Consider the group element $w = (a_1, \ldots, a_n)$. If $a_i \neq 0$, then the q points of the form (i,x) for $x \in GF(q)$ are all fixed by this element; if $i \neq 0$, they are permuted in q/p cycles of length p, each of the form

$$(i,x) \mapsto (i,x+a_1) \mapsto (i,x+2a_i) \mapsto \cdots \mapsto (i,x+pa_i) = (i,x),$$

the last equation holding because $\mathrm{GF}(q)$ has characteristic p. So this element contributes $s_1^{q(n-\mathrm{wt}(w))} s_p^{(q/p)\,\mathrm{wt}(w)}$ to the sum in the formula for the cycle index, and $X^{n-\mathrm{wt}(w)} Y^{\mathrm{wt}(w)}$ to the weight enumerator of C. The result follows.

7.3 More generally ...

The construction of a permutation group from a code does not require the code to be linear, only for it to form an additive group. So the procedure works much more generally. What is the coding-theoretic equivalent of the cycle index of the group?

Proposition 7.3 *Let C be an additive $\mathbb{Z}_4$-code, with symmetrised weight enumerator S_C. Then*

$$\frac{1}{|C|} S_C(X,Y,Z) = Z(G; s_1 \leftarrow X^{1/4}, s_2 \leftarrow Y^{1/2}, s_4 \leftarrow Z).$$

Proof The proof is almost identical to that of Proposition 7.2. The permutation corresponding to the codeword $w = (a_1, \ldots, a_n)$ acts on the set $\{(i,x) : x \in \mathbb{Z}_4\}$ as four fixed points if $a_i = 0$; as two 2-cycles if $a_i = 2$; or as one 4-cycle if $a_i = 1$ or $a_i = 3$.

More generally, let C be any subgroup of the direct product $A_1 \times A_2 \times \cdots \times A_n$, where $A_1, \ldots, A_n$ are groups of order q. Then C acts on the set $\bigcup_{i=1}^{n} A_i$ (disjoint union) in the obvious way. The cycle index of the corresponding permutation group is a kind of generalised symmetrised weight enumerator of C, a multivariate polynomial which counts the number of codewords whose projection onto A_i has order m_i, where $m_1, \ldots, m_n$ are divisors of q. I will not pursue this further.

Exercises

7.1. Calculate the Tutte polynomial of a q-fold expansion of the free matroid F_r and of its dual, and show that both of these matroids are graphic.

Remark The dual of the above matroid was used recently by Alan Sokal [29] to show that the zeros of chromatic polynomials of planar graphs are dense in the complex plane outside the circle with centre and radius 1.

7.2. True or false? A permutation group with cycle index involving only s_1 and s_2 arises as $G(C)$ for some linear code over a field of characteristic 2.

7.3. Let G be an abelian permutation group, having n orbits each of length q. Show that G i associated with a code of length n over alphabets $A_1, \ldots, A_n$, each of which is an abelian group of order q.

7.4. Let $(v_i : i \in I)$ be vectors spanning a vector space V over $F = \mathrm{GF}(q)$. Verify the following description of the group of the code associated with the vector matroid defined by these vectors:

- the domain is $\Omega = I \times F$;
- the group is V^*;
- the action is given by
$$f : (i,a) = (i, a + v_i f)$$
for $i \in I$, $a \in F$ and $f \in V^*$.

CHAPTER 8

IBIS groups

In this chapter, we consider a special class of permutation groups introduced by Cameron and Fon-Der-Flaass [7], which have a very close connection with matroids, in the sense that the bases for the permutation group form the bases of a matroid. These include the groups we associated with linear codes, for which the weight enumerator of the code is essentially the same as the cycle index of the group. They also include the *base-transitive groups*, for which the associated matroids are perfect matroid designs. We conclude by proposing a more general polynomial which includes both Tutte polynomial and cycle index.

This is surely not the end of the story. For an arbitrary permutation group, the irredundant bases do not constitute a matroid. Perhaps there is some more general structure, for which the analogue of the Tutte polynomial of a matroid can be defined.

This chapter is my reason for preparing these notes. The original version is in the paper [5].

8.1 Matroids and IBIS families

The basic idea which connects matroids to permutation groups works in much greater generality.

Let I be an index set, and let $(X_i : i \in I)$ be a family of subsets of a set A. For any $J \subseteq I$, let

$$X_J = \bigcap_{j \in J} X_j.$$

By convention, we put $X_\emptyset = A$.

The subset J of I is called a *base* if $X_J = X_I$. Moreover, if J is ordered, say $J = (j_1, \ldots, j_k)$, then we say that J is *irredundant* if, for each m with $1 \leq k$, we have

$$X_{j_m} \not\supseteq X_{\{j_1,\ldots,j_{m-1}\}},$$

or, in other words, $X_{\{j_1,\ldots,j_m\}} \subset X_{\{j_1,\ldots,j_{m-1}\}}$. Note that any ordered base can be made irredundant by dropping those indices for which this condition fails.

Theorem 8.1 *The following conditions on a family $(X_i : i \in I)$ of sets are equivalent:*

(a) All irredundant bases have the same number of elements.

(b) The irredundant bases are preserved by re-ordering.

(c) The irredundant bases are the bases of a matroid on I.

Proof Suppose that condition (a) holds, and let J be an (ordered) irredundant base and J' be obtained by re-ordering J. Clearly J' is a base, so we can obtain an irredundant base by possibly dropping some elements. But, if any elements are dropped, then the resulting base would be smaller than J. So (b) holds.

Next, suppose that (b) holds. We have to verify the matroid base axioms, that is, no base contains another, and the exchange axiom holds. The first condition is clear: if $J \subset K$, we can order K so that the elements of J come first; then the irredundance of K is contradicted.

Let J and K be irredundant bases, and suppose that $j \in J \setminus K$. Order $J \cup K \setminus \{j\}$ so that the elements of $J \setminus \{j\}$ come first. This is a base, and so we can obtain an irredundant base by dropping some elements of K. We have to show that only one element of K remains; so suppose not, and let k be the first element of K to appear. Then the ordered sequence consisting of the elements of $J \setminus \{j\}$, then k, then j is an irredundant base, but if the last two elements are swapped, it is no longer irredundant, contradicting (b).

Finally, (c) trivially implies (a).

A family of sets satisfying the conditions of this theorem is called an *IBIS family*. (This term is an acronym for "Irredundant Bases of Invariant Size", reflecting condition (a).)

Every matroid can be represented by an IBIS family. For let M be a matroid on E, and let A be the family of hyperplanes of M. For $e \in E$, let X_e be the set of hyperplanes containing e. It is now a simple exercise to prove that $(X_e : e \in E)$ is an IBIS family, whose associated matroid is M. This is a bit surprising: we think of the exchange axiom as an essential part of the definition of a matroid; but here we see that it follows from the constancy of base size.

8.2 IBIS groups

Let G be a permutation group on Ω, We say that G is an *IBIS permutation group*, or *IBIS group* for short, if the family $(G_\alpha : \alpha \in \Omega)$ of point stabilisers is an IBIS family of subsets of G.

Remark The family of point stabilisers in a permutation group is closed under conjugation. Conversely, if $(G_i : i \in I)$ is any IBIS family of subgroups of the group G which is closed under conjugation, then G_I is a normal subgroup of G and G/G_I is isomorphic to an IBIS permutation group. I do not know anything about IBIS families of subgroups which are not closed under conjugation.

In the case when G is a permutation group on Ω and $(G_\alpha : \alpha \in \Omega)$ is the family of point stabilisers, we see that G_I is just the pointwise stabiliser of I, for $I \subseteq \Omega$. Hence the notions of a base and an irredundant base for the family coincide with those we met in Chapter 5: a *base* is a sequence of points whose stabiliser is the identity, and it is *irredundant* if no point in the sequence is fixed by the stabiliser of its predecessors.

So we can say more succinctly: the permutation group G on Ω is an *IBIS group* if its irredundant bases all have the same cardinality. The irredundant bases of such a group G are the bases of a matroid on the set Ω, and clearly G acts as a group of automorphisms of this matroid. We define the *rank* of an IBIS group to be the common cardinality of its irredundant bases (that is, the rank of the associated matroid).

We now give some examples of IBIS groups. First we note that adding or removing global fixed points of a permutation group doesn't change the IBIS property or the rank; so, where necessary, we assume that there are none. (A global fixed point of an IBIS group is a loop of the associated matroid.)

Any non-identity semiregular permutation group (one in which the stabiliser of any point is the identity) is an IBIS group of rank 1, and conversely (apart from global fixed points). Also, the stabiliser of a point in an IBIS group is an IBIS group, with rank one less than that of the original. (This is the analogue of deletion for IBIS groups. There is no natural analogue of contraction.)

Let t be a non-negative integer, and let G be a t-transitive permutation group in which the stabiliser of any $t+1$ points is the identity (but the stabiliser of t points is not the identity. Such groups have had a lot of attention in the literature, though there appears to be no general name for them. I will call them *t-Frobenius groups*: this extends the terminology *Frobenius groups* for permutation groups satisfying this condition with $t = 1$. (A 0-Frobenius group is just a semiregular permutation group.)

Any t-Frobenius group is an IBIS group, and the associated matroid is the uniform matroid $U_{t+1,n}$. The converse is also true:

Theorem 8.2 *Let G be an IBIS group of rank $t+1$, whose associated matroid is the uniform matroid $U_{t+1,n}$. Then G is a t-Frobenius group.*

Proof We have to show that such a group is t-transitive. The proof is by induction on t. When $t = 0$, there is nothing to show; we start the induction with the case $t = 1$. An exercise in Wielandt's book [32] shows that, if G is a permutation group in which all 2-point stabilisers are trivial, then either G is semiregular, or G has one orbit on which it acts s a Frobenius group, and the action on all other orbits is regular. In our case, there cannot be any regular orbits, since these would give bases of cardinality 1. So G is a Frobenius group.

Now suppose that the result holds for $t-1$, and let G be an IBIS group of rank $t+1$ with associated matroid $U_{t+1,n}$. Then the point stabiliser G_α acts on the remaining points as an IBIS group with matroid $U_{t,n-1}$. By induction, G_α is $(t-1)$-transitive; so G is t-transitive, as required.

For Frobenius groups, we have good information about the structure, based on *Frobenius' Theorem*:

Theorem 8.3 *Let G be a Frobenius group. Then the identity and the fixed-point-free elements form a subgroup N of G, which is regular and normal in G.*

The subgroup N is called the *Frobenius kernel* of G. It follows that G is the semidirect product of N and a point stabiliser G_α (which is called a *Frobenius complement*). Moreover, Thompson proved that the Frobenius kernel is nilpotent, and Zassenhaus proved that the structure of a Frobenius complement is very restricted: in particular, it has at most one non-abelian composition factor (this being isomorphic to the smallest non-abelian simple group A_5). See Passman [26] for an account of this work (which preceded CFSG).

A 2-Frobenius group is usually called a *Zassenhaus group*. These groups were completely determined by Zasenhaus, Feit, Ito, and Suzuki (also before CFSG); such a group either is soluble or has minimal normal subgroup isomorphic to $\mathrm{PSL}(2,q)$ or $\mathrm{Sz}(q)$ for some prime power q. (The *Suzuki groups* $\mathrm{Sz}(q)$ were discovered by Suzuki in the course of this determination.) An account of this is also found in Passman [26]. From this, it is possible to determine the t-Frobenius groups for all larger values of t.

Hence we can conclude that all IBIS groups whose associated matroid is $U_{r,n}$ for $r > 2$ are known. However, the situation is very different for matroids which are inflations of uniform matroids (see Exercise 8.7 and the following remark); so the standard procedure in matroid theory of collapsing parallel elements to a single element cannot be applied here.

Let V be a vector space of dimension n over the field $\mathrm{GF}(q)$. The general linear group $\mathrm{GL}(n,q)$ acts on V as an IBIS group; the associated matroid is the complete vector matroid $V(n,q)$. To see this, we observe first that the pointwise stabiliser of any set of vectors fixes pointwise the subspace they span; and then, given any proper subspace, there is a non-identity linear transformation fixing this subspace pointwise.

Any subgroup of $\mathrm{GL}(n,q)$ with this last property is an IBIS group with the same associated matroid. One example is the *symplectic group*, the group of linear transformations preserving a non-degenerate alternating bilinear form B on V. For any hyperplane has the form $a^\perp = \{v \in V : B(a,v) = 0\}$ for some non-zero vector a; this hyperplane is fixed pointwise by the *symplectic transvection*

$$x \mapsto x + B(x,a)a.$$

The group we associated with a linear code in the last chapter is an IBIS group. We discuss this in the next section.

The 5-transitive Mathieu group M_{24} is an IBIS group of rank 7. The associated matroid is not the familiar one whose hyperplanes are the blocks of the associated Steiner triple system $S(5,8,24)$ defined in Exercise 3.4 (this matroid has rank 6), nor is it the matroid associated with the extended Golay code mentioned in Chapter 1 (this matroid has rank 12).

8.3 Groups from codes

Let $G(C)$ be the permutation group that we associated earlier to a $[n,k]$ code C over $\mathrm{GF}(q)$. Recall that $G(C)$ is the additive group of C, and acts on the set $\{1,\ldots,n\} \times \mathrm{GF}(q)$ by the rule

$$(a_1,\ldots,a_n) : (i,x) \mapsto (i,x+a_i).$$

This group is an IBIS group of rank k. For a set $\{(i_1,x_1),\ldots,(i_k,x_k)\}$ is a base for $G(C)$ if and only if the only codeword with zeros in positions $i_1,\ldots,i_k$ is the zero word; this is equivalent to saying that the columns of a generator matrix of C with indices $i_1,\ldots,i_k$ are linearly independent; so any irredundant base for $G(C)$ has rank k.

Now the matroid associated with C has the property that a set $\{i_1,\ldots,i_l\}$ is independent if and only if the corresponding columns of a generator matrix for C are linearly independent. Thus the matroid associated with $G(C)$ is the q-fold inflation of the matroid $M(C)$ of the code C.

Proposition 7.1 shows that we can pass between $T(M(C))$ and $T(M(C)_q)$. Greene's theorem shows that these polynomials determine the weight enumerator

of C, and hence the cycle index of G. But the weight enumerator of C does not determine the Tutte polynomial of $M(C)$, since we can have codes with the same weight enumerator but different Tutte polynomials.

So in this case, the Tutte polynomial carries more information than the cycle index. Sometimes, however, it is the other way around, as we will see.

8.4 Flat actions

The action of an IBIS group on its associated matroid has the following very strong property:

($*$) The pointwise stabiliser of any set of points fixes pointwise the flat spanned by the set.

For let B be a subset of A minimal with respect to having the same pointwise stabiliser. A point α not fixed by the stabiliser of B can be adjoined to B, and the result extended to an irredundant base; so α is independent of B.

An action of a group on a matroid will be called *flat* if condition ($*$) holds.

Any permutation group has a flat action on the free matroid; and any linear group (that is, any subgroup of $\mathrm{GL}(n,q)$) has a flat action on the vector matroid $V(n,q)$.

If a group has a flat action on a perfect matroid design, then an analogue of the Shift Theorem holds: there is a linear relation between the numbers of orbits of the group on independent tuples of points and the probabilities that a random group element has a flat of given rank as its fixed point set. We prove this by showing that a linear relation holds between numbers of orbits on independent tuples and numbers of orbits on arbitrary tuples; then we can invoke the original Shift Theorem corollary.

Theorem 8.4 *Let M be a $\mathrm{PMD}(k_0,\ldots,k_r)$, with $k_r = n$. Then there are numbers $b(m,i)$, for $0 \le m \le n$ and $0 \le i \le r$, depending only on $k_0,\ldots,k_r$, such that the following is true: If a group G has a flat action on M and has x_i orbits on independent i-tuples and y_m orbits on m-tuples of distinct elements, then*

$$y_m = \sum_{i=0}^{r} b(m,i)x_i$$

for $m = 0,\ldots,n$.

Proof By the Orbit-Counting Lemma, it suffices to show that such a linear relation holds between the number of linearly independent i-tuples fixed by an arbitrary element $g \in G$ and the total number of m-tuples of distinct elements fixed

by g. Since the fixed points of G form a flat, it suffices to establish such a relation between the numbers of tuples in any flat of M.

So let F be an s-flat containing x_i independent i-tuples and y_m m-tuples of distinct elements. Then

$$\begin{aligned} x_i &= \prod_{j=0}^{i-1}(k_s - k_j) = X_i(k_s), \\ y_m &= \prod_{t=0}^{m-1}(k_s - t) = Y_m(k_s), \end{aligned}$$

where X_i and Y_i are polynomials of degree i, independent of s. It follows immediately that the theorem holds for $m \leq r$, with $(b(m,i))$ the transition matrix between the two sequences of polynomials.

For $m > r$, let $F_m(x)$ be the unique monic polynomial of degree m having roots $k_0, \ldots, k_r$ and no term in x^l for $r+1 \leq l \leq m-1$. Using F_m, we can express k_i^m (and hence $Y_m(k_i)$) as a linear combination of $1, k_i, \ldots, k_i^r$ (and hence of $X_0(k_i), \ldots, X_r(k_i)$). This concludes the proof.

Remark It is also interesting to consider the numbers z_m of orbits of G on arbitrary m-tuples. As we mentioned in Section 6.3, for any permutation group G, we have

$$z_m = \sum_{k=1}^{m} S(m,k) y_k,$$

where the numbers $S(m,k)$ are the Stirling numbers of the second kind (so that $S(m,k)$ is the number of partitions of an m-set with k parts). Hence, by the standard inversion for the Stirling numbers, we have

$$y_m = \sum_{k=1}^{m} s(m,k) z_k,$$

where the numbers $s(m,k)$ are the (signed) Stirling numbers of the first kind (so that $(-1)^{m-k}s(m,k)$ is the number of permutations of an m-set having k cycles). Thus we can easily move back and forth between these two sequences.

In the case of the free matroid, every set is independent, and so $x_i = y_i$, and the matrix $(b(m,i))$ is the identity.

For the complete vector matroid $V(n,q)$, we have

$$z_m = \sum_{i=0}^{m} \begin{bmatrix} m \\ i \end{bmatrix}_q x_i,$$

where the numbers $\begin{bmatrix} m \\ i \end{bmatrix}_q$ are the *Gaussian coefficients*, so that $\begin{bmatrix} m \\ i \end{bmatrix}_q$ is the number of i-flats in $V(m,q)$. Hence the matrix $(b(m,i))$ is the composite of the matrices of Gaussian coefficients and Stirling numbers.

All this can be found in Cameron and Taylor [10].

Remark The exponential generating function for $y_0, \ldots, y_n$ is $P_G(x+1)$, by the corollary to the Shift Theorem. So the numbers $x_0, \ldots, x_r$ determine $P_G(x)$.

Now the number of fixed points of an element of G is equal to the cardinality of a flat, that is, in the set $\{k_1, \ldots, k_r\}$; so the other coefficients of $P_G(x)$ are all zero. If the coefficient of x^{k_i} in $P_G(x)$ is p_i, then we have a linear map connecting the sequences $(p_0, \ldots, p_r)$ and $(x_0, \ldots, x_r)$.

In the case of the free matroid, this map is given by Corollary 6.3: we have $x_i/i! = \sum_{j=0}^{i} \binom{i}{j} p_j$. In the case of the complete vector matroid, it is the q-analogue of this, involving the Gaussian coefficients. In each case there is a standard method to invert the matrix. (See Cameron and Majid [9] for a connection between inversion of the q-analogue and affine braided groups.)

I do not know a convenient formula for this matrix or its inverse in the case of a general PMD.

8.5 Base-transitive groups

If G is a permutation group which permutes its ordered (irredundant) bases transitively, then clearly all the irredundant bases have the same size, and so G is an IBIS group. Moreover, since G also permutes the ordered independent sets of size i transitively for all i, the associated matroid is a perfect matroid design.

Such groups have been given the somewhat unfortunate name of "geometric groups". Here I will simply call them *base-transitive permutation groups*, or *base-transitive groups* for short. The base-transitive groups of rank greater than 1 were determined by Maund [23], using CFSG; those of sufficiently large rank by Zil'ber [33] by geometric methods not requiring the Classification. Base-transitive groups of rank 1 are just regular permutation groups (possibly with some global fixed points).

Theorem 8.5 *For a base-transitive group G, the p.g.f. $P_G(x)$ and the Tutte polynomial of the associated matroid determine each other, and each is determined by knowledge of the numbers of fixed points of elements of G.*

Proof A permutation group G is base-transitive if and only if the stabiliser of any sequence of points acts transitively on the points that it doesn't fix (if any). Thus the fixed points of every element form a flat. Also, by Corollary 5.4, every flat is the fixed point set of some element. So the numbers of fixed points of the elements of G determine the cardinalities of flats, and hence the Tutte polynomial of the matroid, by Theorem 3.6.

Theorem 8.4 shows that the numbers $k_0,\ldots,k_r$ of fixed points of elements in a base-transitive group determine the function $P_G(x)$, since the numbers $x_0,\ldots,x_r$ are all equal to 1.

To obtain $P_G(x)$ directly from the Tutte polynomial, we show the following:

$$P_G(x+1)=\sum_{m=0}^{n}\left(\sum_{i=0}^{r}\frac{a(m,i)}{R(i)}\right)x^m,$$

where $n=k_r$ is the number of points, and $R(i)$ is the number of independent i-tuples in the matroid; as in Theorem 3.6, $a(m,i)$ is the number of m-sets of rank i.

To prove this, we note that each m-set can be ordered in $m!$ different ways. If the rank of the m-set is i, the resulting sequence has stabiliser of order $\prod_{j=i}^{r-1}(n-k_j)$, and so lies in an orbit of size $\prod_{j=0}^{i-1}(n-k_j)=R(i)$. Thus, the number of orbits on such tuples is $a(m,i)m!/R(i)$. We obtain the total number of orbits on m-tuples by summing over i, and so we find that the exponential generating function is the right-hand side of the displayed equation. But this e.g.f. is $P_G(x+1)$, by the corollary to the Shift Theorem.

Even for a regular permutation group, knowledge of the fixed point numbers does not determine the cycle index; the latter also contains information about the number of group elements of each given order. A regular permutation group is base-transitive. So we see that the cycle index contains more information than the Tutte polynomial in this case.

8.6 Some examples

Unfortunately, the cycle index does not in general tell us whether a permutation group is base-transitive. The simplest counterexample consists of the two permutation groups of degree 6,

$$G_1=\langle(1,2)(3,4),(1,3)(2,4)\rangle,\quad G_2=\langle(1,2)(3,4),(1,2)(5,6)\rangle.$$

The first is base-transitive; the second is an IBIS group of rank 2 (indeed, it is the group arising from the binary even-weight code of length 3), but not base-transitive. A simple modification of this example shows that the cycle index does not determine whether the IBIS property holds.

Suppose we are given the cycle index of one of these groups, namely $Z(G)=\frac{1}{4}(s_1^6+3s_1^2s_2^2)$, or simply the p.g.f. for fixed points, namely $P_G(x)=\frac{1}{4}(x^6+3x^2)$.

(a) If we are told that the group is base-transitive, then we know that its matroid is a PMD$(2,6)$, and so we can compute that its Tutte polynomial is $y^2(y^3+y^2+y+x)$.

(b) If we are told that the group arises from a linear code C, then we can deduce that $W_C(X,Y) = X^3 + 3XY^2$. In general the Tutte polynomial is not computable from the weight enumerator, but in this case the code must be the even-weight code and so the Tutte polynomial of the code matroid is $x^2 + x + y$. Now Proposition 7.1 shows that the Tutte polynomial of the group matroid is $y^4 + 2y^3 + 3y^2 + y + 3xy + x^2 + x$.

This matroid on 6 elements in case (b) arises from two different base-transitive groups of order 24, each isomorphic to the symmetric group S_4. These were both considered in Chapter 6; one is the action of S_4 on the set of 2-element subsets of $\{1,\ldots,4\}$, and the other is the action of the rotation group of the cube on the set of faces. Using any of several methods we've seen, it follows that, for any such group G, we have $P_G(x) = \frac{1}{24}(x^6 + 9x^2 + 14)$. However, the stabiliser of a point is the Klein group of order 4 in the first case and is the cyclic group in the other, so the two groups have different cycle index. (See Exercise 6.4.)

8.7 The Tutte cycle index

As we have seen, for some IBIS groups the Tutte polynomial can be obtained from the cycle index but not *vice versa*, while for others it is the opposite way round. Is there a polynomial from which both the Tutte polynomial and the cycle index can be obtained? In this section we construct such a polynomial.

Following the definition of the Tutte polynomial, we try for a sum, over subsets, of "local" terms. First, some terminology and observations. Let G be a permutation group on Ω. For any subset Δ of Ω, G_Δ and $G_{(\Delta)}$ are the setwise and pointwise stabilisers of Δ, and $G[\Delta]$ the permutation group induced on Δ by its setwise stabiliser (so that $G[\Delta] \cong G_\Delta/G_{(\Delta)}$). Let $b(G)$ denote the minimum size of a base for G. (This is the rank of the associated matroid if G is an IBIS group.)

Now we have

(a) $$\sum_{\Delta \in \mathcal{P}\Omega/G} Z(G[\Delta]) = Z(G; s_i \leftarrow s_i + 1 \text{ for } i = 1,\ldots,n),$$
where $\mathcal{P}\Omega/G$ denotes a set of orbit representatives for G on the power set of Ω. This is Theorem 6.2, the Shift Theorem.

(b) If G is an IBIS group, then the fixed point set of $G_{(\Delta)}$ is the flat spanned by Δ; so $G_{(\Delta)}$ is an IBIS group, and $\rho(\Delta) = b(G) - b(G_{(\Delta)})$. In fact, $G_{(\Delta)}$ is also an IBIS group, but its base size may be smaller than $\rho(\Delta)$.

Now we define the *Tutte cycle index* of G to be the polynomial in $u, v, s_1, \ldots, s_n$ given by

$$ZT(G) = \frac{1}{|G|} \sum_{\Delta \subseteq \Omega} u^{|G_\Delta|} v^{b(G_{(\Delta)})} Z(G[\Delta]).$$

One obvious flaw in this definition is that the factors $u^{|G_\Delta|}$ and $v^{b(G_{(\Delta)})}$ are not really "local". Nevertheless, we have the properties we are looking for:

Theorem 8.6 *Let G be an IBIS permutation group, with associated matroid M.*

(a) $\left(\dfrac{\partial}{\partial u}ZT(G)\right)(u \leftarrow 1, v \leftarrow 1) = Z(G; s_i \leftarrow s_i + 1 \text{ for } i = 1,\ldots,n).$

(b) $|G|ZT(G; u \leftarrow 1, s_i \leftarrow t^i \text{ for } i = 1,\ldots,n) = t^{b(G)}T(M; x \leftarrow v/t + 1, y \leftarrow t + 1).$

Proof (a) The G-orbit of the subset Δ has cardinality $|G|/|G_\Delta|$. Dividing by this number has the same effect as choosing one representative set from each orbit. Now apply point (a) before the Theorem.

(b) Point (b) before the Proposition shows that $\rho(\Delta) = b(G) - b(G_{(\Delta)})$; in particular, $\rho(\Omega) = b(G)$. Also, substituting t^i for s_i in $Z(H)$ gives t^n, where n is the degree of the permutation group H. So the left-hand side is

$$\sum_{\Delta\subseteq\Omega} v^{\rho(\Omega)-\rho(\Delta)}\, t^{|\Delta|}.$$

The rest is just manipulation.

Exercises

8.1. Let M be a matroid on E, and A the set of hyperplanes of H. For $e \in E$, let X_e be the set of hyperplanes containing e. Prove that $(X_e : e \in E)$ is an IBIS family whose associated matroid is M.

8.2. Show that any family of subgroups, all of index p, in an elementary abelian p-group is an IBIS family. Describe the associated matroid by means of the dual group.

8.3. Let G be an IBIS group of permutations of Ω.

(a) Let Δ be an orbit of G. Prove that both the permutation group induced on Δ and the kernel of the action of G on Δ are IBIS groups.

(b) Prove that the stabiliser of a point is an IBIS group.

(c) Prove that the group induced on a flat by its setwise stabiliser is an IBIS group.

8.4. Let G_i be an IBIS group of permutations of Ω_i, for $i \in I$. Prove that the direct product of the groups G_i, acting on the disjoint union of the sets Ω_i, is an IBIS group.

8.5. Let p be a prime. Let G be an elementary abelian p-group, and $(H_i : i \in I)$ a family of subgroups of G. If each subgroup H_i has index p in G, prove that $(H_i : i \in I)$ is an IBIS family. Show that this may not be true if not all the subgroups have index p.

8.6. Prove that, if G is a permutation group in which all two-point stabilisers are trivial, then either G is semiregular, or G acts as a Frobenius group on one of its orbits and regularly on all the others. (Hint: Use Frobenius' Theorem.)

8.7. Prove that M_{24} is an IBIS group of rank 7.

8.8. Calculate the Tutte cycle indices for the group S_4, in each of its transitive actions as an IBIS group on 6 points. Hence calculate the Tutte polynomial and cycle index in each case.

8.9. Say that a base for a permutation group G is *strongly irredundant* if the removal of any point results in a sequence which is no longer a base. Show that a base is strongly irredundant if and only if any ordering of it is irredundant. Give an example of a permutation group which has strongly irredundant bases of different sizes.

8.10. Show that a sharply t-transitive group (other than the symmetric group S_t of degree t) is a base-transitive group associated with a uniform matroid $U_{t,n}$, and conversely. Why is the free matroid $F_n = U_{n,n}$ not associated with the symmetric group S_n?

8.11. Let G be a base-transitive group associated with a q-fold inflation of the free matroid F_n (with $q > 1$. Show that G is the wreath product of a regular group H of order q with S_n.

8.12. Show that a group G, acting on the coset space $G : H$, is a Frobenius group if and only if $H \cap H^g = 1$ for all $g \notin H$.

A subgroup H of G is called a *TI-subgroup* if $H \cap H^g = 1$ for all $g \notin N_G(H)$, where $N_G(H) = \{g \in G : H^g = H\}$ is the *normaliser* of H in G. Show that, if H is a non-normal TI-subgroup of G, then G acting on $G : H$ is an IBIS group, for which the associated matroid is an inflation of a uniform matroid of rank 2.

Remark TI-subgroups are very common: for example, any subgroup of prime order is a TI-subgroup. This shows that the procedure of identifying parallel elements can have dramatic effects in the case of an IBIS group.

Bibliography

[1] K. D. Blaha, Minimal bases for permutation groups: the greedy approximation, *J. Algorithms* **13** (1992), 297–306.

[2] N. Boston, W. Dabrowski, T. Foguel, P. J. Gies, J. Leavitt, D. T. Ose and D. A. Jackson, The proportion of fixed-point-free elements of a transitive permutation group, *Commun. Algebra* **21** (1993), 3259–3275.

[3] P. J. Cameron, *Oligomorphic Permutation Groups*, London Math. Soc Lecture Notes **152**, Cambridge University Press, Cambridge, 1990.

[4] P. J. Cameron, *Permutation Groups*, London Math. Soc. Student Texts **45**, Cambridge University Press, Cambridge, 1999.

[5] P. J. Cameron, Cycle index, weight enumerator and Tutte polynomial, *Electronic J. Combinatorics* **9** (2002), #N2 (10pp), available from `http://www.combinatorics.org`

[6] P. J. Cameron and A. M. Cohen, On the number of fixed point free elements of a permutation group, *Discrete Math.* **106/107** (1992), 135–138.

[7] P. J. Cameron and D. G. Fon-Der-Flaass, Bases for permutation groups and matroids, *Europ. J. Combinatorics* **16** (1995), 537–544.

[8] P. J. Cameron and W. M. Kantor, Random permutations: Some group-theoretic aspects, *Combinatorics, Probability and Computing* **2** (1993), 257–262.

[9] P. J. Cameron and S. Majid, Braided line and counting fixed points of $\mathrm{GL}(d,\mathbb{F}_q)$, preprint available from
`http://lanl.arXiv.org/abs/math/0112258`

[10] P. J. Cameron and D. E. Taylor, Stirling numbers and affine equivalence, *Ars Combinatoria* **20B** (1985), 3–14.

[11] H. H. Crapo, The Tutte polynomial, *Aequationes Math.* **3** (1969), 211–229.

[12] M. Deza, Perfect matroid designs, *Encycl. Math. Appl.* **40** (1992), 54–72.

[13] J. D. Dixon and B. Mortimer, *Permutation Groups*, Springer, New York, 1996.

[14] D. Gewurz, *Sui vettori di Parker e concetti correlati*, Ph.D. thesis, Università di Roma "La Sapienza", 1999.

[15] D. Gorenstein, *Finite Simple Groups: An Introduction to their Classification*, Plenum Press, New York, 1982.

[16] C. Greene, Weight enumeration and the geometry of linear codes, *Studia Appl. Math.* **55** (1976), 119–128.

[17] M. Hall, Jr., Automorphisms of Steiner triple systems, *IBM J. Research Develop.* **4** (1960), 460–472.

[18] F. Harary and E. M. Palmer, *Graphical Enumeration*, Academic Press, New York, 1973.

[19] A. R. Hammons Jr., P. V. Kumar, A. R. Calderbank, N. J. A. Sloane and P. Solé, The Z_4-linearity of Kerdock, Preparata, Goethals and related codes, *IEEE Trans. Inform. Theory* **40** (1994), 301–319.

[20] M. R. Jerrum, Computational Pólya theory, pp. 103–118 in *Surveys in Combinatorics, 1995* (P. Rowlinson, ed.), London Math. Soc. Lecture Notes **218**, Cambridge University Press, Cambridge, 1995.

[21] M. W. Liebeck and A. Shalev, Simple groups, permutation groups, and probability, *J. Amer. Math. Soc.* **12** (1999), 497–520.

[22] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*, North-Holland, Amsterdam, 1977.

[23] T. C. Maund, D.Phil. thesis, University of Oxford, 1989.

[24] E. G. Mphako, Tutte polynomials of perfect matroid designs, *Combinatorics, Probability and Computing* **9** (2000), 363–367.

[25] J. G. Oxley, *Matroid Theory*, Oxford University Press, Oxford, 1992.

[26] D. S. Passman, *Permutation Groups*, Benjamin, New York, 1968.

[27] C. G. Rutherford, *Matroids, codes and their polynomial links*, Ph.D. thesis, University of London, 2001.

[28] J.-P. Serre, On a theorem of Jordan, *Mathematical Medley* (Singapore Mathematical Society), to appear. Available from `http://www.math.nus.edu.sg/~chanhh/JordanMar11.pdf`

[29] A. D. Sokal, Bounds on the complex zeros of (Di)chromatic polynomials and Potts-model partition functions, *Combinatorics, Probability and Computing* **10** (2001), 41–77.

[30] D. J. A. Welsh, *Matroid Theory*, Academic Press, London, 1976.

[31] D. J. A. Welsh, *Complexity: Knots, Colourings and Counting*, London Mathematical Society Lecture Notes **186**, Cambridge University Press, Cambridge, 1993.

[32] H. Wielandt, *Finite Permutation Groups*, Academic Press, New York, 1964.

[33] B. I. Zil'ber, The structure of models of uncountably categorical theories, pp. 359–368 in *Proc. Internat. Congr. Math.* Vol. 1 (Warsaw 1983).

Index