

Error-correcting Codes from Permutation Groups

Robert Bailey

Queen Mary, University of London

`r.f.bailey@qmul.ac.uk`

`http://www.maths.qmul.ac.uk/~rfb/`

2nd PoPMiNE, 28th June 2004

WHAT ARE ERROR-CORRECTING CODES?

- We want to transmit a message *accurately* along a “noisy channel”, where there may be interference.
- If the received message contains errors, we want to *decode* it to receive the original transmitted message.
- If the possible messages are sufficiently “different”, then this will be possible.
- What do we mean by “different”?
If messages are strings of the same length, then the *Hamming distance* between two strings is the number of places in which they differ.
- E.g. $d_H(001, 010) = 2$.

- Formally, a *code*, C , is a set of strings of symbols chosen from some alphabet.
- The *minimum distance* of C is

$$\min_{\substack{v, w \in C \\ v \neq w}} \{d_H(v, w)\}$$

Proposition

If C has minimum distance d , then C can correct

$$\left\lfloor \frac{d-1}{2} \right\rfloor$$

errors.

- “Good” codes have:
 1. a reasonably large number of codewords;
 2. a reasonably large minimum distance;
 3. a usable decoding algorithm.

SO WHERE ARE THE PERMUTATION GROUPS?

- Let G be a permutation group acting on Ω , where $|\Omega| = n$.
- We can write elements of G as ordered n -tuples of distinct symbols from Ω ,

$$\text{e.g. } 231794685 \in S_9$$

- Can define Hamming distance as before.
- However,

$$\begin{aligned} d_H(g, h) &= \#x \text{ where } x^g \neq x^h \\ &= n - |\text{Fix}(gh^{-1})| \end{aligned}$$

- Thus the minimum distance is

$$\min_{\substack{g \in G \\ g \neq 1}} \{n - |\text{Fix}(g)|\},$$

the *minimum degree* of G .

This parameter is not always easy to calculate, but for the following families it is straightforward:

- *Sharply k -transitive groups*
Minimum distance is $n - k + 1$
(no two elements can agree on k or more points)
- $GL(n, q)$ acting on $\mathbb{F}_q^n \setminus \{0\}$
Minimum distance is $q^n - q^{n-1}$
(fixed points sets are the vector subspaces of \mathbb{F}_q^n)
- $AGL(n, q)$ acting on \mathbb{F}_q^n
Minimum distance is $q^n - q^{n-1}$
(fixed points sets are the affine subspaces of \mathbb{F}_q^n)
- $C_m \wr S_n$ acting on $\{1, \dots, m\}^n$
Minimum distance is m
(fixed points occur in multiples of m)

WHAT ABOUT A DECODING ALGORITHM?

This uses algorithms from computational group theory.

Definition

A *base* for G is a sequence of points (x_1, \dots, x_b) from Ω such that its pointwise stabiliser is the identity.

- A consequence of this is that the action of $g \in G$ on a base *uniquely determines that element*.
- Clearly, if a received word contains e errors, then it must contain $n - e$ correct symbols.
- So, if these occur in positions labelled by a base, we can decode successfully.
- **PROBLEM:** We can't necessarily tell in which positions the errors are.

SOLUTION:

- We need a set of bases such that any combination of r error positions is disjoint from at least one base.
- We call this an *uncovering-by-bases*.
- E.g. for a sharply k -transitive group, ANY k points form a base.
- So we need a set \mathcal{U} of k -subsets of $\{1, \dots, n\}$ such that any r -subset is *disjoint* from at least one k -set.
- This is the complement of a *covering design*, so we call it an *uncovering*.

Example:

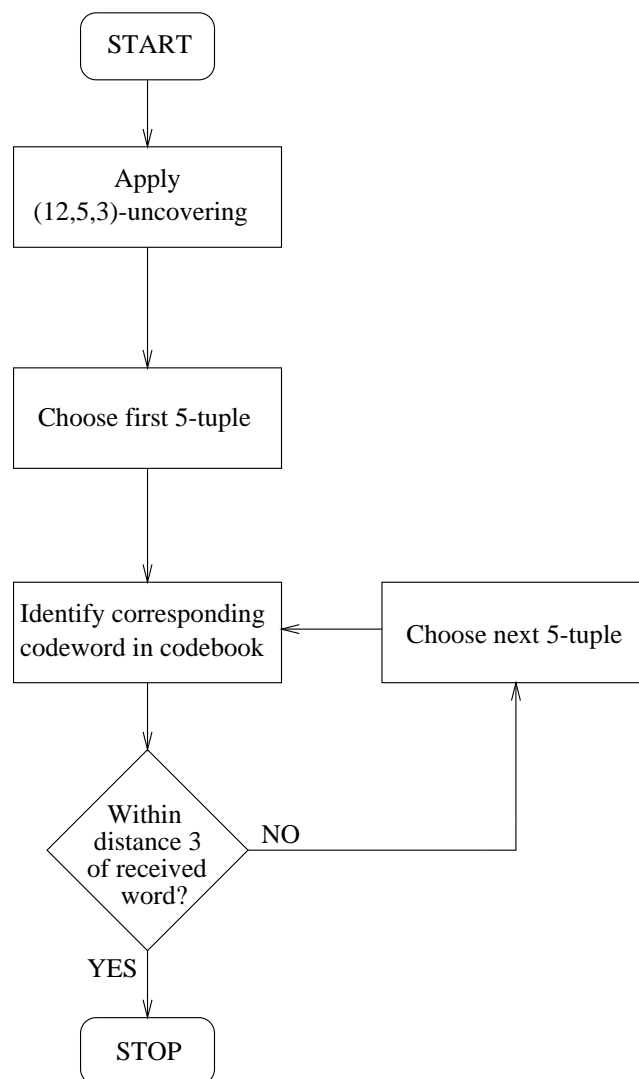
- Mathieu Group M_{12}
- Sharply 5-transitive, degree 12, order 95040
- Minimum degree 8
- Can correct $\left\lfloor \frac{8-1}{2} \right\rfloor = 3$ errors
- Need a $(12, 5, 3)$ -uncovering

Example:

1	2	3	4	5
1	2	6	11	12
1	3	7	8	9
1	4	6	7	10
1	5	8	9	11
2	4	8	9	12
2	5	7	10	11
3	4	7	11	12
3	5	6	10	12
3	6	8	9	11
6	7	8	9	10

Once we have such an uncovering, for each 5-set we determine the *unique* group element agreeing with received word in these 5 positions.

If distance is ≤ 3 , then stop. Otherwise, go to next 5-set and repeat.



Example:

Suppose we transmit

$$g = 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ 10\ 11\ 12$$

and receive

$$w = 6\ 2\ 1\ 4\ 6\ 6\ 7\ 8\ 9\ 10\ 11\ 12$$

w has errors in positions 1, 3 and 5.

As the algorithm works through the uncovering, it outputs:

Error (repeated symbol);

Error (repeated symbol);

6 3 1 4 12 2 7 8 9 5 10 11, which is distance 6 from w and is rejected;

Error (repeated symbol);

Error (repeated symbol);

1 2 3 4 5 6 7 8 9 10 11 12, which is distance 3 from w and is accepted.

For which groups do we know the structure of the bases?

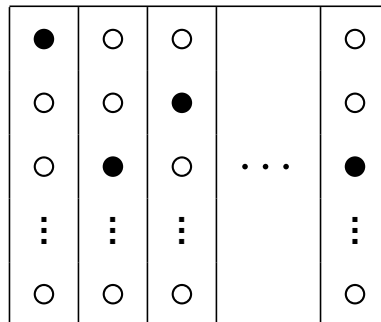
It helps if the bases are “well-behaved” in some way. The following list of groups are all *base-transitive*, that is G acts transitively on its irredundant bases.

- *Sharply k -transitive groups*
Any k points form a base.
- $GL(n, q)$ acting on $\mathbb{F}_q^n \setminus \{0\}$
A base for the group is a basis for the vector space.
- $AGL(n, q)$ acting on \mathbb{F}_q^n
A base for the group is an affine basis for \mathbb{F}_q^n .
- $C_m \wr S_n$ acting on $\{1, \dots, m\}^n$

The base-transitive groups were classified by Maund (D.Phil. thesis).

More on $C_m \wr S_n$

- A base consists of a single point from each copy of $\{1, \dots, m\}$.
- We call such a base a *transversal*.



- To apply our decoding algorithm, we need a set of transversals such that any possible set of $\lfloor \frac{m-1}{2} \rfloor$ error positions is avoided by at least one transversal.

We'll consider the case $m = 5$, where we can correct 2 errors.

- If both errors occur in the same column, then three disjoint transversals will do the job.
- If the errors occur in two distinct columns, we start with two disjoint transversals $\mathbf{x} = \{x_1, \dots, x_n\}$ and $\mathbf{y} = \{y_1, \dots, y_n\}$, and construct further transversals from them.
- We label these by vectors as follows:
If, in a new transversal, x_i occurs in the i^{th} position, label this by $+1$, else put -1 .
- So we now need a set of row vectors with entries $+1, -1$, such that in any pair of columns, each of the pairs $(+1, +1)$, $(+1, -1)$, $(-1, +1)$ and $(-1, -1)$ occurs.
- Where can we find such a set of vectors?

HADAMARD MATRICES

Definition

A *Hadamard matrix* of order n is an $n \times n$ matrix H , with entries $+1, -1$, which satisfies

$$H.H^T = n.I$$

- This is equivalent to saying that any two columns (or rows) are orthogonal.
- Example:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

- It is always possible to *normalise* a Hadamard matrix, so that the first row and column are all $+1$ s.

- A consequence of the orthogonality is that, in any pair of columns (other than the first) the pairs $(+1, +1)$, $(+1, -1)$, $(-1, +1)$ and $(-1, -1)$ must all occur (in fact, the same number of times).
- So, this gives us the set of row vectors we require.
- This in turn gives us a means of constructing our transversals.

EXCITING NEWS!

- A Hadamard matrix must have order 1, 2 or a multiple of 4.
- It is conjectured that they exist for *all* multiples of 4.
- Prior to last week, the smallest multiple of 4 for which no Hadamard matrix was known to exist was 428. But one has now been found: see

<http://www.ipm.ac.ir/>