Permutation codes

Peter J. Cameron

p.j.cameron@qmul.ac.uk

CGCS, Luminy, 3 May 2007

Subsets and permutations

There are many analogies between sets of subsets of $\{1, \ldots, n\}$ and sets of permutations of $\{1, \ldots, n\}$.

Subsets and permutations

There are many analogies between sets of subsets of $\{1, \ldots, n\}$ and sets of permutations of $\{1, \ldots, n\}$.

In both cases, the objects can be represented by lists of length n (with entries $\{0, 1\}$ for subsets or $\{1, \ldots, n\}$ for permutations).

Subsets and permutations

There are many analogies between sets of subsets of $\{1, \ldots, n\}$ and sets of permutations of $\{1, \ldots, n\}$.

In both cases, the objects can be represented by lists of length n (with entries $\{0, 1\}$ for subsets or $\{1, \ldots, n\}$ for permutations).

In each case, there is a metric structure (*Hamming distance* for the lists, $d(x, y)$ is the number of positions where x and y differ) and an algebraic structure (addition mod 2 or symmetric difference for subsets, composition for permutations).

Subsets and permutations

There are many analogies between sets of subsets of $\{1, \ldots, n\}$ and sets of permutations of $\{1, \ldots, n\}$.

In both cases, the objects can be represented by lists of length n (with entries $\{0, 1\}$ for subsets or $\{1, \ldots, n\}$ for permutations).

In each case, there is a metric structure (*Hamming distance* for the lists, $d(x, y)$ is the number of positions where x and y differ) and an algebraic structure (addition mod 2 or symmetric difference for subsets, composition for permutations).

It is a pleasure to present this paper at a conference for Michel Deza, who was a pioneer in the investigation of permutations from this point of view.

Algebraic substructures

The algebraic substructures are particularly interesting. For subsets, these are the linear codes over $\mathbb{F}_2$; for permutations, they are the permutation groups. If we are looking for extremal results, they are likely to be much stronger for these than for arbitrary families.

Algebraic substructures

The algebraic substructures are particularly interesting. For subsets, these are the linear codes over $\mathbb{F}_2$; for permutations, they are the permutation groups. If we are looking for extremal results, they are likely to be much stronger for these than for arbitrary families.

Here is a comparison of the two situations, showing corresponding concepts and parameters of a linear code C and a permutation group G.

Algebraic substructures

The algebraic substructures are particularly interesting. For subsets, these are the linear codes over $\mathbb{F}_2$; for permutations, they are the permutation groups. If we are looking for extremal results, they are likely to be much stronger for these than for arbitrary families.

Here is a comparison of the two situations, showing corresponding concepts and parameters of a linear code C and a permutation group G.

One of the most important parameters is the cardinality C or G.

Dimension and base size

A linear code C is a subspace of $\mathbb{F}_2{}^n$, and so has a dimension k. We have $|C| = 2^k$.

Dimension and base size

A linear code C is a subspace of $\mathbb{F}_2{}^n$, and so has a dimension k. We have $|C| = 2^k$.

In a permutation group G, a *base* is a sequence $i_1, \ldots, i_b$ of points whose pointwise stabiliser is the identity. Bases are important in computational group theory since an element of G is uniquely determined by its effect on a base. If b is the minimum base size for G, then

$$2^b \leq |G| \leq n^b.$$

Dimension and base size

A linear code C is a subspace of $\mathbb{F}_2{}^n$, and so has a dimension k. We have $|C| = 2^k$.

In a permutation group G, a *base* is a sequence $i_1, \ldots, i_b$ of points whose pointwise stabiliser is the identity. Bases are important in computational group theory since an element of G is uniquely determined by its effect on a base. If b is the minimum base size for G, then

$$2^b \leq |G| \leq n^b.$$

The bases of a linear code satisfy the matroid basis axioms; the bases of a permutation group do not, in general.

Minimum weight and minimum degree

In both cases, the minimum distance of the code or group (the minimum distance between distinct elements) is equal to the minimum weight (the minimum distance from zero or identity to another element). In the group case, the weight of G is n minus the number of fixed points of G.

Minimum weight and minimum degree

In both cases, the minimum distance of the code or group (the minimum distance between distinct elements) is equal to the minimum weight (the minimum distance from zero or identity to another element). In the group case, the weight of G is n minus the number of fixed points of G.

The minimum weight d of a code determines its error-correction capability; it can correct up to $\lfloor (d-1)/2 \rfloor$ errors.

Minimum weight and minimum degree

In both cases, the minimum distance of the code or group (the minimum distance between distinct elements) is equal to the minimum weight (the minimum distance from zero or identity to another element). In the group case, the weight of G is n minus the number of fixed points of G.

The minimum weight d of a code determines its error-correction capability; it can correct up to $\lfloor (d-1)/2 \rfloor$ errors.

The minimum weight of a permutation group is usually called its minimum degree. This parameter has been studied since the time of Jordan.

Covering radius

A parameter which is in some sense dual to minimum distance is the covering radius, the maximum (over all words or permutations x) of the minimum distance from x to the code or group.

Covering radius

A parameter which is in some sense dual to minimum distance is the covering radius, the maximum (over all words or permutations x) of the minimum distance from x to the code or group.

This is also related to error correction: if more errors occur than the covering radius, then nearest-neighbour decoding will certainly be wrong!

Covering radius

A parameter which is in some sense dual to minimum distance is the covering radius, the maximum (over all words or permutations x) of the minimum distance from x to the code or group.

This is also related to error correction: if more errors occur than the covering radius, then nearest-neighbour decoding will certainly be wrong!

Much is known about this parameter for codes, but comparatively little for permutation groups.

A problem on covering radius

Let G be the 1-dimensional affine group over $\mathbb{F}_q$:

$$G = \{x \mapsto ax + b : a, b \in \mathbb{F}_q, a \neq 0\}.$$

What is the covering radius of G?

A problem on covering radius

Let G be the 1-dimensional affine group over $\mathbb{F}_q$:

$$G = \{x \mapsto ax + b : a, b \in \mathbb{F}_q, a \neq 0\}.$$

What is the covering radius of G?

Ian Wanless and I showed that the answer is $q - 2$ if q is even; $q - 3$ if q is odd and not congruent to 1 mod 6; and either $q - 3$ or $q - 4$ in the remaining case.

A problem on covering radius

Let G be the 1-dimensional affine group over $\mathbb{F}_q$:

$$G = \{x \mapsto ax + b : a, b \in \mathbb{F}_q, a \neq 0\}.$$

What is the covering radius of G?

Ian Wanless and I showed that the answer is $q - 2$ if q is even; $q - 3$ if q is odd and not congruent to 1 mod 6; and either $q - 3$ or $q - 4$ in the remaining case.
This has a geometric interpretation. The covering radius is $q - s$ if and only if there is a set Q of q points in the affine plane over $\mathbb{F}_q$ which meets every horizontal or vertical line in one point and any other line in at most s points, and s is the least such number.

A problem on covering radius

Let G be the 1-dimensional affine group over $\mathbb{F}_q$:

$$G = \{x \mapsto ax + b : a, b \in \mathbb{F}_q, a \neq 0\}.$$

What is the covering radius of G?

Ian Wanless and I showed that the answer is $q - 2$ if q is even; $q - 3$ if q is odd and not congruent to 1 mod 6; and either $q - 3$ or $q - 4$ in the remaining case.
This has a geometric interpretation. The covering radius is $q - s$ if and only if there is a set Q of q points in the affine plane over $\mathbb{F}_q$ which meets every horizontal or vertical line in one point and any other line in at most s points, and s is the least such number.

Problem

Find the covering radius in the remaining case.

Strength and degree of transitivity

Another parameter of a code is its strength (as an 'orthogonal array'), the largest number t such that, in any t coordinate positions, all possible t-tuples occur equally often as codewords.

Strength and degree of transitivity

Another parameter of a code is its strength (as an 'orthogonal array'), the largest number t such that, in any t coordinate positions, all possible t-tuples occur equally often as codewords.

Delsarte observed that the strength of a linear code is one less than the minimum weight of the dual code.

Strength and degree of transitivity

Another parameter of a code is its strength (as an 'orthogonal array'), the largest number t such that, in any t coordinate positions, all possible t-tuples occur equally often as codewords.

Delsarte observed that the strength of a linear code is one less than the minimum weight of the dual code.

Analogously we have the degree of transitivity of a permutation group, the largest t for which the group acts transitively on t-tuples of distinct points. This is another parameter whose study goes back to the nineteenth century.

Strength and degree of transitivity

Another parameter of a code is its strength (as an 'orthogonal array'), the largest number t such that, in any t coordinate positions, all possible t-tuples occur equally often as codewords.

Delsarte observed that the strength of a linear code is one less than the minimum weight of the dual code.

Analogously we have the degree of transitivity of a permutation group, the largest t for which the group acts transitively on t-tuples of distinct points. This is another parameter whose study goes back to the nineteenth century.

Two differences between strength and degree of transitivity: first, there is no 'dual' permutation group, so Delsarte's result is not available; second, using the Classification of Finite Simple Groups, the degree of transitivity cannot be greater than 5 (apart from the symmetric and alternating groups).

Weight and support enumerators

The weight enumerator of a code is the generating function for the number of words of given weight. The analogous polynomial for a permutation group is the support enumerator. Often it is more natural to count fixed points instead, giving the fixed point enumerator.

Weight and support enumerators

The weight enumerator of a code is the generating function for the number of words of given weight. The analogous polynomial for a permutation group is the support enumerator. Often it is more natural to count fixed points instead, giving the fixed point enumerator.

These polynomials, suitably normalised, are the probability generating functions for the weight, or number of fixed points, of a randomly chosen element of the code or permutation group.

Weight and support enumerators

The weight enumerator of a code is the generating function for the number of words of given weight. The analogous polynomial for a permutation group is the support enumerator. Often it is more natural to count fixed points instead, giving the fixed point enumerator.

These polynomials, suitably normalised, are the probability generating functions for the weight, or number of fixed points, of a randomly chosen element of the code or permutation group.

Nigel Boston and others showed that, if $P_G(x)$ is the fixed point enumerator, normalised by dividing by $|G|$, and $F_G(x)$ is the exponential generating function for the number of orbits of the group on i-tuples of distinct points, then

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

Other polynomials

According to a theorem of Curtis Greene, the weight enumerator of a code C is a specialisation of the two-variable Tutte polynomial of the matroid whose bases are the bases for the code.

Other polynomials

According to a theorem of Curtis Greene, the weight enumerator of a code C is a specialisation of the two-variable Tutte polynomial of the matroid whose bases are the bases for the code.

Analogously, the fixed point enumerator of a permutation group is a specialisation of the n-variable cycle index of the group.

Other polynomials

According to a theorem of Curtis Greene, the weight enumerator of a code C is a specialisation of the two-variable Tutte polynomial of the matroid whose bases are the bases for the code.

Analogously, the fixed point enumerator of a permutation group is a specialisation of the n-variable cycle index of the group.

It is tempting to think that these two multivariate polynomials have a common generalisation, at least in some cases. There are some pointers in this direction.

IBIS groups

An irredundant base in a permutation group is a base $(i_1, \ldots, i_b)$ with the property that no base point is fixed by the stabiliser of its predecessors.

IBIS groups

An irredundant base in a permutation group is a base $(i_1, \ldots, i_b)$ with the property that no base point is fixed by the stabiliser of its predecessors.
A minimal base (and in particular a base of minimal cardinality) is irredundant, but not conversely. Irredundant bases can have different cardinalities.

IBIS groups

An irredundant base in a permutation group is a base $(i_1, \ldots, i_b)$ with the property that no base point is fixed by the stabiliser of its predecessors.
A minimal base (and in particular a base of minimal cardinality) is irredundant, but not conversely. Irredundant bases can have different cardinalities.
A permutation group is an IBIS group if it satisfies the following three conditions (which Dima Fon-Der-Flaas and I proved to be equivalent):

- all irredundant bases have the same cardinality;
- irredundant bases are preserved by re-ordering;
- irredundant bases satisfy the matroid basis axioms.

IBIS groups, continued

Problem

Determine all IBIS groups, or at least all the matroids which arise from IBIS groups.

IBIS groups, continued

Problem

Determine all IBIS groups, or at least all the matroids which arise from IBIS groups.

Note that from any binary linear code we get an IBIS group, whose matroid is obtained by 'doubling' each element of the matroid of the code. Clearly we can't hope to determine all of these! But perhaps the primitive IBIS groups can be determined.

Extremal permutation theory

This theory, much of it due to Michel Deza and his co-authors, takes results of extremal set theory and finds analogues for permutations.

Extremal permutation theory

This theory, much of it due to Michel Deza and his co-authors, takes results of extremal set theory and finds analogues for permutations.

For a simple example, the distances between distinct permutations lie in the set $\{2, 3, \ldots, n\}$. If A is a subset of this set, we let $F_A(n)$ be the maximum cardinality of a set of permutations such that all distances lie in the set A.
An argument using the fact that the metric space admits a transitive group of isometries shows that, if A and B are sets with $A \cup B = \{2, \ldots, n\}$, then

$$F_A(n) \cdot F_B(n) \leq n! \cdot F_{A \cap B}(n).$$

In particular, if also $A \cap B = \emptyset$ then $F_A(n) F_B(n) \leq n!$.

The coding problem

Let $F_{\geq d}(n)$ denote the maximum number of permutations which are pairwise at distance at least d. An analogue of the Singleton bound holds:

$$F_{\geq n-t+1}(n) \leq n(n-1)\cdots(n-t+1).$$

The coding problem

Let $F_{\geq d}(n)$ denote the maximum number of permutations which are pairwise at distance at least d. An analogue of the Singleton bound holds:

$$F_{\geq n-t+1}(n) \leq n(n-1)\cdots(n-t+1).$$

Equality holds if and only if there is a sharply t-transitive set of permutations (any t-tuple of distinct points can be carried to any other by a unique permutation in the set).

The coding problem

Let $F_{\geq d}(n)$ denote the maximum number of permutations which are pairwise at distance at least d. An analogue of the Singleton bound holds:

$$F_{\geq n-t+1}(n) \leq n(n-1)\cdots(n-t+1).$$

Equality holds if and only if there is a sharply t-transitive set of permutations (any t-tuple of distinct points can be carried to any other by a unique permutation in the set).

The existence of sharply t-transitive sets of permutations for $t = 1, 2, 3$ is equivalent to that of certain geometric objects: Latin squares, affine planes, inversive planes respectively. So they always exist for $t = 1$; but for $t = 2$ it is a very hard problem!

Analogue of EKR

Let $F_{\leq d}(n)$ denote the maximum number of permutations which are pairwise at distance at most d, i.e. any two agreeing in at least $n - d$ points.

Analogue of EKR

Let $F_{\leq d}(n)$ denote the maximum number of permutations which are pairwise at distance at most d, i.e. any two agreeing in at least $n - d$ points.

Problem

If $n \geq n_0(t)$, then $F_{\leq n-t}(n) \leq (n-t)!$. Moreover, a set which attains this bound is a coset of the stabiliser of t points in the symmetric group.

Analogue of EKR

Let $F_{\leq d}(n)$ denote the maximum number of permutations which are pairwise at distance at most d, i.e. any two agreeing in at least $n - d$ points.

Problem

If $n \geq n_0(t)$, then $F_{\leq n-t}(n) \leq (n-t)!$. Moreover, a set which attains this bound is a coset of the stabiliser of t points in the symmetric group.

This is true for $t = 1$: the bound comes from the fact that $F_{\{n\}}(n) = n$, and the structure of sets meeting the bound uses the fact that Latin squares exist in profusion. This method will not easily generalise.

Analogue of EKR

Let $F_{\leq d}(n)$ denote the maximum number of permutations which are pairwise at distance at most d, i.e. any two agreeing in at least $n - d$ points.

Problem

If $n \geq n_0(t)$, then $F_{\leq n-t}(n) \leq (n-t)!$. Moreover, a set which attains this bound is a coset of the stabiliser of t points in the symmetric group.

This is true for $t = 1$: the bound comes from the fact that $F_{\{n\}}(n) = n$, and the structure of sets meeting the bound uses the fact that Latin squares exist in profusion. This method will not easily generalise.

For $t = 2$, we know that $F_{\leq n-2}(n) = (n-2)!$ if there exists a projective plane of order n (that is, we know it if n is a prime power!). New methods are needed!

At the other end . . .

We considered $F_{\leq s}(n)$ for s close to n. What if s is small?

At the other end . . .

We considered $F_{\leq s}(n)$ for s close to n. What if s is small?

For fixed s, the examples in the preceding slide have size $s!$, independent of n. One can do better. If s is even, then the ball of radius $s/2$ about any permutation has all distances at most s, and has cardinality

$$|B_{s/2}(g)| = \sum_{i=0}^{s/2} \binom{n}{i} d(i) \sim f(s) n^{s/2},$$

where $d(i)$ is the number of derangements of an i-set. There is a similar construction for s odd.

At the other end . . .

We considered $F_{\leq s}(n)$ for s close to n. What if s is small?

For fixed s, the examples in the preceding slide have size $s!$, independent of n. One can do better. If s is even, then the ball of radius $s/2$ about any permutation has all distances at most s, and has cardinality

$$|B_{s/2}(g)| = \sum_{i=0}^{s/2} \binom{n}{i} d(i) \sim f(s) n^{s/2},$$

where $d(i)$ is the number of derangements of an i-set. There is a similar construction for s odd.

Deza and Frankl showed that, if $n \geq n_1(s)$, then these sets have maximum size, and are the only sets which do so.

At the other end . . .

We considered $F_{\leq s}(n)$ for s close to n. What if s is small?

For fixed s, the examples in the preceding slide have size $s!$, independent of n. One can do better. If s is even, then the ball of radius $s/2$ about any permutation has all distances at most s, and has cardinality

$$|B_{s/2}(g)| = \sum_{i=0}^{s/2} \binom{n}{i} d(i) \sim f(s) n^{s/2},$$

where $d(i)$ is the number of derangements of an i-set. There is a similar construction for s odd.

Deza and Frankl showed that, if $n \geq n_1(s)$, then these sets have maximum size, and are the only sets which do so.

What happens in the middle of the range, where both s and $n - s$ are large?

What about groups?

Better bounds are known for groups. These results are ‘classical’.

What about groups?

Better bounds are known for groups. These results are 'classical'.

All the sharply t-transitive groups for $t > 1$ are known. The determination is by Jordan for $t \geq 4$ and Zassenhaus for $t = 2$ and $t = 3$.

What about groups?

Better bounds are known for groups. These results are 'classical'.

All the sharply t-transitive groups for $t > 1$ are known. The determination is by Jordan for $t \geq 4$ and Zassenhaus for $t = 2$ and $t = 3$.

If $F_A^\circ(n)$ denotes the largest cardinality of a permutation group whose distances lie in the set A, then Blichfeldt showed that

$$F_A^\circ(n) \text{ divides } \prod_{a \in A} a.$$

A group meeting this bound is called a sharp permutation group. The sharp groups have been determined in several cases, but the general determination is not yet complete.

Permutation groups as codes

To conclude I would like to discuss some recent work by Robert Bailey on another topic introduced by Michel Deza and others, concerning the possibility of using a permutation group as an error-correcting code.

Permutation groups as codes

To conclude I would like to discuss some recent work by Robert Bailey on another topic introduced by Michel Deza and others, concerning the possibility of using a permutation group as an error-correcting code.

Let G be a permutation group of degree n which has minimal degree m. We have seen that G can correct up to e errors, where $e = \lfloor (m-1)/2 \rfloor$.

Permutation groups as codes

To conclude I would like to discuss some recent work by Robert Bailey on another topic introduced by Michel Deza and others, concerning the possibility of using a permutation group as an error-correcting code.

Let G be a permutation group of degree n which has minimal degree m. We have seen that G can correct up to e errors, where $e = \lfloor (m-1)/2 \rfloor$.

Suppose that we use G as a code over the alphabet $\{1, \ldots, n\}$. Let $(i_1, \ldots, i_b)$ be a base. If we knew that the entries in the received word in these positions were correct, then we could calculate the transmitted word uniquely using techniques of computational group theory.

Uncovering by bases

A set $\mathcal{B}$ of bases for G is said to be an uncovering by bases (or UBB) if, for every set E of points of cardinality $e = \lfloor (m-1)/2 \rfloor$, there is a base $B \in \mathcal{B}$ such that $E \cap B = \emptyset$.

Uncovering by bases

A set $\mathcal{B}$ of bases for G is said to be an uncovering by bases (or UBB) if, for every set E of points of cardinality $e = \lfloor (m-1)/2 \rfloor$, there is a base $B \in \mathcal{B}$ such that $E \cap B = \emptyset$.

Thus, if we have an uncovering by bases, then we can decode: check bases in turn until we find one yielding a transmitted word distant at most e from the received word.

Uncovering by bases

A set $\mathcal{B}$ of bases for G is said to be an uncovering by bases (or UBB) if, for every set E of points of cardinality $e = \lfloor (m-1)/2 \rfloor$, there is a base $B \in \mathcal{B}$ such that $E \cap B = \emptyset$.

Thus, if we have an uncovering by bases, then we can decode: check bases in turn until we find one yielding a transmitted word distant at most e from the received word.

A UBB resembles a covering design, with two differences. First, we uncover rather than cover; so we have to take the complements of the blocks of a covering design. Second, we insist that all these uncovering sets should be bases.

Two conjectures

An easy argument shows that, for any permutation group G, there is a UBB for G. Two features which would make the decoding algorithm more efficient are: a small UBB; and a UBB whose bases belong to a single G-orbit.

Problem

Let G be a permutation group of degree n. Show that there is a UBB for G such that

- *its size is bounded by a low-degree polynomial in n;*
- *it is contained in a single orbit of G on bases.*

Two conjectures

An easy argument shows that, for any permutation group G, there is a UBB for G. Two features which would make the decoding algorithm more efficient are: a small UBB; and a UBB whose bases belong to a single G-orbit.

Problem

Let G be a permutation group of degree n. Show that there is a UBB for G such that

- *its size is bounded by a low-degree polynomial in n;*
- *it is contained in a single orbit of G on bases.*

This has been proved for a variety of permutation groups, by a variety of group-theoretic and combinatorial techniques.

Ordering a UBB

Usually, error patterns with a small number of errors are most likely. So we can improve the average run-time of the decoding algorithm if we can find a UBB $\mathcal{B} = \mathcal{B}_e$ containing a chain of subsets

$$\mathcal{B}_1 \subseteq \mathcal{B}_2 \subseteq \ldots \subseteq \mathcal{B}_e$$

such that

- $\mathcal{B}_i$ is a UBB for sets of size i;
- $|\mathcal{B}_i|$ is (close to) optimal for such a design.

Ordering a UBB

Usually, error patterns with a small number of errors are most likely. So we can improve the average run-time of the decoding algorithm if we can find a UBB $\mathcal{B} = \mathcal{B}_e$ containing a chain of subsets

$$\mathcal{B}_1 \subseteq \mathcal{B}_2 \subseteq \ldots \subseteq \mathcal{B}_e$$

such that

- $\mathcal{B}_i$ is a UBB for sets of size i;
- $|\mathcal{B}_i|$ is (close to) optimal for such a design.

Problem

Do UBBs with this property exist?

This is an interesting question even with no reference to bases (i.e. for general covering designs).