Math 2^8: The Theory of Error-Correcting Codes

The ternary Golay codes and related structures

We have seen already that if C is an extremal $[12, 6, 6]$ Type III code then C attains equality in the LP bounds (and conversely the LP bounds show that any ternary $(12, 3^6, 6)$ code is a translate of such a C); earlier we observed that removing any coordinate from C yields a perfect 2-error-correcting $[11, 6, 5]$ ternary code. Since C is extremal, we also know that its Hamming weight enumerator is uniquely determined, and compute

$$W_C(X, Y) = (X(X^3 + 8Y^3))^3 - 24(Y(X^3 - Y^3))^3 = X^{12} + 264X^6Y^6 + 440X^3Y^9 + 24Y^{12}.$$

Here are some further remarkable properties that quickly follow:

A (5,6,12) Steiner system. The minimal words form 132 pairs $\{c, -c\}$ with the same support. Let $\mathcal{S}$ be the family of these 132 supports. We claim $\mathcal{S}$ is a $(5, 6, 12)$ Steiner system; that is, that any pentad (5-element set of coordinates) is a subset of a unique hexad in $\mathcal{S}$. Because $\mathcal{S}$ has the right size $\binom{12}{5} / \binom{6}{5}$, it is enough to show that no two hexads intersect in exactly 5 points. If they did, then we would have some $c, c' \in C$, both of weight 6, whose supports have exactly 5 points in common. But then c' would agree with either c or $-c$ on at least 3 of these points (and thus on exactly 4, because $(c, c') = 0$), making $c' \mp c$ a nonzero codeword of weight less than 5 (and thus exactly 3), which is a contradiction. Hence $\mathcal{S}$ is a $(5, 6, 12)$ Steiner system, as claimed. $\diamondsuit$

A Hadamard matrix of order 12. The *maximal* words form 12 pairwise orthogonal pairs $\{c, -c\}$ of ± 1 vectors. We claim that their lifts to ± 1 vectors in $\mathbf{R}^{12}$ remain pairwise orthogonal, and thus that choosing either representative from each pair yields the rows of a Hadamard matrix H. Again the proof uses the condition that C has minimal weight 6: if c, c' are maximal codewords and $c \neq c'$, then each of $c \pm c'$ has weight at least 6, which means that c and c' agree in six of the 12 coordinates and disagree on the other six.[1] $\diamondsuit$

Both the Steiner system $\mathcal{S}$ and the order-12 Hadamard matrix H turn out to be unique up to automorphism. The automorphism group of $\mathcal{S}$ is Mathieu's sporadic simple group M_{12}, which like M_{24} is 5-transitive, and unlike M_{24} is *sharply* 5-transitive, i.e., acts *simply* transitively on ordered quintuples of distinct coordinates; in particular $|M_{12}| = 12 \cdot 11 \cdot 10 \cdot 9 \cdot 8$ (which is numerically $2^6 3^3 5 \cdot 11 = 95040$). The point stabilizer M_{11}, which is thus sharply 4-transitive and of size $11 \cdot 10 \cdot 9 \cdot 8 = 7920$, is also sporadic, indeed the smallest sporadic simple group.[2] Unlike the case for M_{24}, the two-point stabilizer M_{10} is no longer sporadic, or even simple: it has an index-2 normal subgroup isomorphic with $A_6 \cong \mathrm{PSL}_2(\mathbf{F}_9)$, with a sharply 3-transitive action on the projective line over $\mathbf{F}_9$.

As with M_{24}, the fact that M_{12} is highly transitive means that there are many ways to prove that $\mathcal{S}$ is unique up to automorphism and to count the number of automorphisms; but the fact that M_{12} is sporadic means that each approach suggests different generalizations and variations, and we can't expect any single approach to reveal the full structure. One approach is to mimic our analysis of the $(5, 8, 24)$ Steiner system: the hexads containing any 3 of the points yield a $(2, 3, 9)$ Steiner system, which is an *affine* plane of order 3; by a standard result in combinatorics[3] this is the complement of a line in a projective plane of order 3, which

[1] Alas this is *not* the source of the expression "six of one, half a dozen of the other"...

[2] It is a known consequence of the classification of finite simple groups that M_{11}, M_{12}, M_{23}, and M_{24} are the only finite 4-transitive permutation groups other than the symmetric and alternating groups (and thus that the only sharply 4-transitive groups are S_4, S_5, A_6, and M_{11}, and likewise the only sharply 5-transitive groups are S_5, S_6, A_7, and M_{12}). There are infinitely many 3-transitive permutation groups, even sharply 3-transitive (for each finite field k the group $\mathrm{PGL}_2(k)$ acts sharply 3-transitively on the projective line over k).

[3] Introduce a relation $\sim$ on the lines by defining $l \sim l'$ to mean that either $l = l'$ or $l \cap l' = \emptyset$. Then prove "Playfair's axiom" that $\sim$ is an equivalence relation. Extend each line by a "point at infinity" depending on its equivalence class, and add a "line at infinity"

we've already proven unique up to automorphism via ovals; then use affine ovals to reconstruct the full Steiner system.[4] Another approach, which we hinted at in class earlier, starts by using repeated inclusion-exclusion to show that $\mathcal{S}$ is closed under complement, and then fixing a complementary pair h, h' of hexads and identifying h' with the totals of h. (Given a pair $p \subset h$ there are three hexads of the form $h - p + p'$ for some pair $p' \subset h'$; we show that this gives a bijection between pairs in h and synthemes in h', "etc.".) This lets us prove that $\mathrm{Aut}(\mathcal{S})$ acts transitively on hexads, and that the hexad stabilizer is all of S_6, acting on the complementary hexad via the outer automorphism. Either approach also shows that M_{12} is contained in M_{24} as the stabilizer of a weight-12 word c_0 in the binary Golay code $\mathcal{G}_{24}$, because we can construct a $(5, 6, 12)$ Steiner system from the intersections with c_0 of all the octads c with $\mathrm{wt}(cc_0) = 6$.

We do not pursue either of these approaches here, because they don't let us easily recover the code from the Steiner system: here we must also find consistent $\pm$ signs, whereas $\mathcal{G}_{24}$ was obtained immediately from the $(5, 8, 24)$ Steiner system as the span of characteristic functions of octads. Instead we use a Hadamard matrix formed form the maximal words. From these, we can recover the minimal nonzero words (and thus the Steiner system $\mathcal{S}$) as the pair differences $\pm c \pm c'$ with c, c' maximal and $c' \neq \pm c$. Indeed there are $2^2\binom{12}{2} = 264$ such expressions, so we need only show that they are all different. If not then we would have four maximal words c_i, with no $c_j = \pm c_i$ unless $i = j$, such that $c_1 + c_2 + c_3 + c_4 = 0$. Lifting to $\mathbf{R}^{12}$ we would find that $c_1 + c_2 + c_3 + c_4 = 3v$ for some $v \in \mathbf{Z}^{12}$; and this is impossible because $c_1 + c_2 + c_3 + c_4$ has norm $4 \cdot 12 = 48$ which is not a multiple of 3^2. The Hadamard route will also show two features of M_{12} not shared by M_{24}: like S_6, the group M_{12} has both a nontrivial double cover $2.M_{12}$ and an outer automorphism.

We review some basic facts about Hadamard matrices. A Hadamard matrix H is said to be *equivalent* to any other Hadamard matrix obtained from H by permuting the rows, permuting the columns, and independently multiplying each row and column by ± 1. In our setting, row operations give different choices of constructing H from our code C, and column operations yield a Hadamard matrix associated to an isomorphic code. In general, we may always use the sign operations (with no need for either row or column permutations) to make all the entries in the first row and and column equal $+1$. Having done this, each other row has as many entries equal $+1$ as equal -1, so $2 \mid n$ once $n > 1$; and indeed any pair of rows after the first form a submatrix with each of the four possible columns $(\pm 1, \pm 1)^{\mathsf{T}}$ equally represented, so $4 \mid n$ once $n > 2$. We cannot continue in this fashion because there is a Hadamard matrix of order 4 (unique up to equivalence):

$$\begin{pmatrix} + & + & + & + \\ + & + & - & - \\ + & - & + & - \\ + & - & - & + \end{pmatrix}$$

and clearly the submatrix of rows $2, 3, 4$ does not show all eight possible columns $(\pm 1, \pm 1, \pm 1)^{\mathsf{T}}$. Indeed it is conjectured (though I'm not clear that there's any good theoretical reason for this[5]) that there is a Hadamard matrix of order n for every $n \equiv 0 \bmod 4$.

Two key techniques for constructing Hadamard matrices are tensor products and the QR (quadratic-residue) construction. If H and H' are Hadamard matrices of orders n and n' then $H \otimes H'$ is a Hadamard matrix of order nn'. Applying this repeatedly to the 2×2 matrix $\left(\begin{smallmatrix} + & + \\ + & - \end{smallmatrix}\right)$ yields a Hadamard matrix of order 2^m that is the matrix of the discrete Fourier transform (DFT) on $\mathbf{F}_2^m$ and is also known as the m-th Walsh matrix (since DFT on $\mathbf{F}_2^m$ is also known as the Walsh transform). The automorphism group is the affine linear

consisting of all the new points. See for instance Cameron and van Lint's text *Designs, Graphs, Codes and their Links*, page 14; the argument is given in my Math 155 notes, http://www.math.harvard.edu/~elkies/M155.09/feb17 . The recovery of the $(5, 6, 12)$ Steiner system $\mathcal{S}$ is Exercise 13 on page 27.

[4] Along the way we encounter a $(3, 4, 10)$ Steiner system that, unlike the $(3, 6, 22)$, is part of an infinite family: the finite "inversive planes", which are $(3, q + 1, q^2 + 1)$ Steiner systems for any prime power q, and consist of copies of $\mathbf{P}^1(\mathbf{F}_q)$ in $\mathbf{P}^1(\mathbf{F}_{q^2})$. This is why M_{10}, unlike M_{22}, is not sporadic.

[5] Such matrices are known for most "small" multiples of 4, and currently the first open case is $n = 668$; but for large n I don't think it's even known that the set of orders of Hadamard matrices has positive density.

group $\mathrm{AGL}_m(\mathbf{F}_2)$. The QR Hadamard matrix of order $p+1$ (with p a prime[6] congruent to $-1 \bmod 4$) has rows and columns indexed by $\infty, 0, 1, \ldots, p-1$; the (i,j) entry is $+1$ iff $i=\infty$, $j=\infty$, or $(\frac{i-j}{p}) = +1$ (i.e. $i-j \bmod p$ is a nonzero square). We already saw that this is a Hadamard matrix in our discussion of QR codes; any $\mathrm{PSL}_2(\mathbf{F}_p)$ permutation of the columns lifts to an automorphism of the matrix.

For $n=8$ and $n=12$ it is known that, as was the case for $n=4$, the Hadamard matrix of order n is unique up to equivalence. This is easy to see for $n=4$, and can be proved for $n=8$ via uniqness of the projective plane of order 2. In particular the QR and Walsh matrices of order 8 are isomorphic; this gives one route to the identification of the 168-element simple group $\mathrm{PSL}_2(\mathbf{F}_7)$ with $\mathrm{GL}_3(\mathbf{F}_2)$. [NB once $m>3$, even if 2^m-1 is a prime the Walsh matrix of size 2^m is not isomorphic with the QR matrix of the same size, and there may be yet further non-isomorphic $2^m \times 2^m$ Hadamard matrices.] We shall outline the proof for $n=12$, and also show that $\mathrm{Aut}(H)$ acts on the columns by a sharply 5-transitive permutation group, which thus contains $\mathrm{PSL}_2(\mathbf{F}_{11})$, much as we saw that M_{24} contains $\mathrm{PSL}_2(\mathbf{F}_{23})$ (it so happens that Mathieu first constructed M_{12} and M_{24} starting from $\mathrm{PSL}_2(\mathbf{F}_{11})$ and $\mathrm{PSL}_2(\mathbf{F}_{23})$ respectively).

Let H, then, be a 12×12 Hadamard matrix. Choose any three rows, and permute rows to make them first, second, and third. Flip columns signs as necessary to make the first row all $+1$, and permute the columns to bring the first three rows to the form

$$\left(\begin{array}{ccc|ccc|ccc|ccc} + & + & + & + & + & + & + & + & + & + & + & + \\ + & + & + & + & + & + & - & - & - & - & - & - \\ + & + & + & - & - & - & + & + & + & - & - & - \end{array}\right).$$

Now suppose a further row has a signs $+1$ in columns 1–3, b signs $+1$ in columns 4–6, c in 7–9, and d in 10–12. Then $a+b+c+d=6$, $a+b=c+d$, and $a+c=b+d$. Hence $a=d$ and $b=c$, so $(a,b,c,d)=(a,3-a,3-a,a)$. But we cannot have $a=3$ or $a=0$, because then the matrix has a row $+++|---|---|+++$ or $---|+++|+++|---$ and no further row can be orthogonal to all four. Thus $(a,b,c,d)=(1,2,2,1)$ or $(2,1,1,2)$, and by flipping row signs as necessary we can put each of the remaining nine rows in $(1,2,2,1)$ form. We simplify notation for these rows by marking only the odd-sign-out in each column triplet; for instance if the next row is $+--|+-+|++-|--+$ then we abbreviate

$$\left(\begin{array}{ccc|ccc|ccc|ccc} + & + & + & + & + & + & + & + & + & + & + & + \\ + & + & + & + & + & + & - & - & - & - & - & - \\ + & + & + & - & - & - & + & + & + & - & - & - \\ \hline + & - & - & + & - & + & + & + & - & - & - & + \end{array}\right)$$

by writing

$$\left(\begin{array}{ccc|ccc|ccc|ccc} + & + & + & + & + & + & + & + & + & + & + & + \\ + & + & + & + & + & + & - & - & - & - & - & - \\ + & + & + & - & - & - & + & + & + & - & - & - \\ \hline + & & & & - & & & & - & & & + \end{array}\right).$$

[NB there's no essential distinction between the outer and inner column-triples: we could even flip the inner six columns' signs to make all the triples visibly equivalent, at the cost of the all-plus normalization of the first row.]

Now the condition that the remaining nine rows be pairwise orthogonal is equivalent to the requirement that in each two exactly one of the odd-signs-out should match. In particular, no two rows can match in two of the column-triples; but in each pair of column-triples there are only $3^2=9$ possibilities, so all must arise!

[6] More generally p may be a prime power congruent to $-1 \bmod 4$, and then the rows and columns are indexed by ∞ and the elements of a finite field of p elements. Wikipedia credits to Paley (1933) this construction and a related "double circulant" one that works for prime powers $p \equiv 1 \bmod 4$ and yields matrices of order $2(p+1)$.

Choose any of those remaining rows and make them fourth and fifth. By permuting and flipping column-triples, and permuting columns within each of the triples, we can normalize the top five rows of our matrix to

$$\left(\begin{array}{ccc|ccc|ccc|ccc}
+ & + & + & + & + & + & + & + & + & + & + & + \\
+ & + & + & + & + & + & - & - & - & - & - & - \\
+ & + & + & - & - & - & + & + & + & - & - & - \\
\hline
+ & & & - & & & - & & & + & & \\
+ & & & & - & & & - & & & + & \\
\end{array}\right).$$

We then permute the remaining rows so that their first two columns are in order as follows; the sixth row is then completely determined, and we're left with only the bottom right 6×6 matrix to fill:

$$\left(\begin{array}{ccc|ccc|ccc|ccc}
+ & + & + & + & + & + & + & + & + & + & + & + \\
+ & + & + & + & + & + & - & - & - & - & - & - \\
+ & + & + & - & - & - & + & + & + & - & - & - \\
\hline
+ & & & - & & & - & & & + & & \\
+ & & & & - & & & - & & & + & \\
+ & & & & & - & & & - & & & + \\
\hline
 & + & & - & & & & ? & ? & & ? & ? \\
 & + & & & - & & ? & & ? & ? & & ? \\
 & + & & & & - & ? & ? & & ? & ? & \\
\hline
 & & + & - & & & & ? & ? & & ? & ? \\
 & & + & & - & & ? & & ? & ? & & ? \\
 & & + & & & - & ? & ? & & ? & ? & \\
\end{array}\right)$$

(there is no "?" where it would cause two rows to match in more than one odd-sign-out). Thus each of the 3×3 blocks still in question must be either $\left|\begin{smallmatrix} * & * \\ * & \end{smallmatrix}\right|$ or $\left|\begin{smallmatrix} * & * \\ & * \end{smallmatrix}\right|$ and each must be different from its horizontal and vertical neighbors. This brings us down to two choices, which are equivalent under the remaining column permutation that switches columns 2 and 3. Filling the empty squares in rows 4 through 12 with the $\pm$ signs that they represent, we finally find that H is equivalent with

$$\left(\begin{array}{ccc|ccc|ccc|ccc}
+ & + & + & + & + & + & + & + & + & + & + & + \\
+ & + & + & + & + & + & - & - & - & - & - & - \\
+ & + & + & - & - & - & + & + & + & - & - & - \\
\hline
+ & - & - & - & + & + & - & + & + & + & - & - \\
+ & - & - & + & - & + & + & - & + & - & + & - \\
+ & - & - & + & + & - & + & + & - & - & - & + \\
\hline
- & + & - & - & + & + & + & - & + & - & - & + \\
- & + & - & + & - & + & + & + & - & + & - & - \\
- & + & - & + & + & - & - & + & + & - & + & - \\
\hline
- & - & + & - & + & + & + & + & - & - & + & - \\
- & - & + & + & - & + & - & + & + & - & - & + \\
- & - & + & + & + & - & + & - & + & + & - & - \\
\end{array}\right).$$

We can now recover $\mathcal{G}_{12}$ as the row span of $H \bmod 3$. We must show that this row span is in fact an extremal Type III code; the parameters here are small enough that this can be done by direct computation, but this is not even needed: the row span is clearly self-orthogonal, so of dimension at most 6, and we've already constructed $264 > 3^5$ distinct words of weight 6, so the dimension cannot be 5 or less. This shows that the row span has the right dimension, and the minimal weight can then be obtained even using Gleason's theorem (if a Type III code of length 12 has $N_3 > 0$ then $N_6 = 264 - 3N_3 < 264$).

Keeping track of choices made along the way, we see that $\mathrm{Aut}(H)$ acts sharply 5-transitively on the rows of H, as claimed; the group of row permutations is isomorphic with Mathieu's sporadic group M_{12}. Since

H^T is also a Hadamard matrix, it follows that $\mathrm{Aut}(H)$ also acts by M_{12} on the columns. (This makes it easy to check in a more direct fashion that $\mathcal{G}_{12}$ has no words of weight 3: since all coordinate triples are equivalent under M_{12}, it is enough to choose some three columns and verify that they are linearly independent and thus do not support a dual codeword of weight 3.) We next describe the nontrivial double cover of M_{12} and the outer automorphism of M_{12}.

The group $\mathrm{Aut}(H) = \mathrm{Aut}(\mathcal{G}_{12})$ acts by a double cover of M_{12}; i.e., the map $\mathrm{Aut}(\mathcal{G}_{12}) \to M_{12}$ has kernel $\{\pm 1\}$. We claim that this is a nontrivial double cover, i.e. that $\mathrm{Aut}(\mathcal{G}_{12}) \ncong \{\pm 1\} \times M_{12}$. It is enough to find an involution in M_{12} that lifts to an element of order 4 in $\mathrm{Aut}(\mathcal{G}_{12})$. We'll use an involution in the copy of $\mathrm{PSL}_2(\mathbf{F}_{11})$ in M_{12}. The involution flips the sign of the coordinates $\infty, 2, 6, 7, 8, 10$ (infinity and the quadratic nonresidues) and then takes each coordinate x to $-1/x$. You should check that this permutes the rows of H up to sign, and iterates to multiplication by -1. This double cover $\mathrm{Aut}(\mathcal{G}_{12})$ of M_{12} is often called $2M_{12}$; it is known (though we don't pretend to prove here) that this fully accounts for the Schur multiplier of M_{12}.

The action of $\mathrm{Aut}(H)$ on the rows and columns of H (equivalently: on the coordinates and maximal codeword-pairs $\pm c$ of $\mathcal{G}_{12}$) are isomorphic, but not under a conjugation in S_{12}, and thus *a fortiori* not by an inner automorphism of M_{12}. This is because Steiner hexads of coordinates correspond to pairs, not hexads, of $\pm c$'s. (We already saw this inside M_{24} when constructing $\mathcal{S}$ from a weight-12 word in $\mathcal{G}_{24}$.) Therefore the two actions of M_{12} are related by an outer automorphism. Again it turns out, though we do not prove it here, that this fully acounts for $\mathrm{Aut}(M_{12})$. (Necessarily this outer automorphism maps the point stabilizer M_{11} to a transitive subgroup of M_{12}; it turns out that the intersection of any two non-conjugate subgroups M_{11} of M_{12} is $\mathrm{PSL}_2(\mathbf{F}_{11})$, acting transitively on 11 letters — Galois already found this action and showed that it is the last case where $\mathrm{PSL}_2(\mathbf{F}_p)$ acts faithfully on fewer than $p+1$ letters.)

Because M_{12} acts transitively on the coordinates of $\mathcal{G}_{12}$, the $[11, 6, 5]$ code obtained by removing one coordinate does not depend on which coordinate is chosen, so we may speak of *the* Golay code $\mathcal{G}_{11}$. We conclude by proving that every perfect 2-error-correcting ternary $[11, 6, 5]$ code is isomorphic with $\mathcal{G}_{11}$.[7] We have seen already while developing Lloyd's theorem that all nonzero words of the dual code $C^\perp$ have weight 6 or 9. It follows that $C^\perp$ is self-orthogonal, and thus that C is the disjoint union of $C^\perp$, $C^\perp + v$, and $C^\perp - v$, where v is any word in the complement, call it C_1, of $C^\perp$ in C. But then all words in C_1 have the same weight mod 3: any such word can be written as $c_0 \pm v$ with $c_0 \in C^\perp$, and then $(c_0, c_0) = (c_0, v) = 0$ implies $\mathrm{wt}(c_0 \pm v) \bmod 3 = (c_0 \pm v, c_0 \pm v) = (v, v)$. In particular $(v, v) = 2$. We can now reconstruct an extremal Type III code by writing any word of C as $c_0 + av$ for some $c_0 \in C^\perp$ and $a \in \mathbf{F}_3$, and extending $c_0 + av$ by a 12th coordinate a. Being extremal, this code is isomorphic with $\mathcal{G}_{12}$, whence $C \cong \mathcal{G}_{11}$ as claimed, **QED**.

[7] Can you prove that in fact every ternary $(11, 3^6, 5)$ code, linear or not, is isomorphic with $\mathcal{G}_{11}$?