MATHIEU GROUPS

SIMON RUBINSTEIN-SALZEDO

Abstract. In this essay, we will discuss some interesting properties of the Mathieu groups, which are a family of five sporadic finite simple groups, as well as some closely related topics.

One construction of the Mathieu groups involves Steiner systems, so we will begin by discussing Steiner systems.

1. Steiner systems

Definition 1. Let $\ell < m < n$ be positive integers. We say that a collection $S_1, S_2, \ldots, S_N$ of distinct subsets of $\{1, 2, \ldots, n\}$ is an (ℓ, m, n)-Steiner system (denoted $S(\ell, m, n)$) if it satisfies the following two properties:

- For each i, $|S_i| = m$.
- For every subset $T \subset \{1, 2, \ldots, n\}$ with $|T| = \ell$, there is exactly one i so that $S_i \supset T$.

Of course, Steiner systems do not exist for most triples (ℓ, m, n), but occasionally they do exist. It can be shown, for instance, that $S(2, 3, n)$ exists if and only if $n \equiv 1 \pmod 6$ or $3 \pmod 6$.

In our discussion of Mathieu groups, we will be most interested in the Steiner systems $S(4, 5, 11)$, $S(5, 6, 12)$, $S(3, 6, 22)$, $S(4, 7, 23)$, and $S(5, 8, 24)$. Let us now show that these Steiner systems do in fact exist.

In fact, we really only need to construct the Steiner systems $S(5, 6, 12)$ and $S(5, 8, 24)$; the others will follow easily from the constructions of these.

In order to construct the Steiner system $S(5, 6, 12)$, we will consider the projective line $\mathbb{P}^1(\mathbb{F}_{11})$ over the field with eleven elements. This consists of the elements $0, 1, 2, \ldots, 10$, together with an additional element, which we will call ∞. Now consider the set of squares in $\mathbb{F}_{11}$, which are $S_1 = \{0, 1, 3, 4, 5, 9\}$. Now suppose $\gamma \in \mathrm{PSL}_2(\mathbb{F}_{11})$. If γ is represented by $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then define

$$\gamma(S_1) = \left\{ \frac{az+b}{cz+d} \,\middle|\, z \in S_1 \right\},$$

Date: October 18, 2011.

where we interpret $\frac{az+b}{cz+d}$ to be ∞ if $cz+d=0$. Then $\{\gamma(S_1) : \gamma \in \mathrm{PSL}_2(\mathbb{F}_{11})\}$ is a Steiner system $S(5,6,12)$.

To construct $S(4,5,11)$ from $S(5,6,12)$, we simply take those elements of $S(5,6,12)$ that do not contain ∞. Thus we have constructed the first two of our five Steiner systems.

We now construct the Steiner system $S(5,8,24)$. Perhaps the simplest way of describing it is by using binary lexicodes, as follows: Let b_1 be the 24-bit binary string consisting of all 0 digits. Now, for $2 \le n \le 4096$, define b_n to be the first string lexicographically that differs from each of the strings $b_1, b_2, \ldots, b_{n-1}$ in at least 8 positions. Of these 4096 strings we have just written down, 759 of them contain exactly eight 1 digits. For each of these 759 strings, take a subset S_i of $\{1,2,\ldots,24\}$ which includes digit d if and only if the d^{th} digit of the corresponding binary string is a 1. The subsets we have chosen form a Steiner system $S(5,8,24)$.

To construct $S(4,7,23)$ and $S(3,6,22)$ from $S(5,8,24)$, we consider those $S_i \in S(5,8,24)$ that do not contain 24 in the first case, and 23 or 24 in the second.

It should be pointed out here that Steiner systems are very rare. In particular, there are no known Steiner systems $S(\ell,m,n)$ with $\ell > 5$, and the only ones known with $\ell = 5$ are $S(5,6,12)$ and $S(5,8,24)$, described here.

2. Mathieu groups

Now that we have defined all the Steiner systems we need, we are able to define the Mathieu groups. These will be the automorphism groups of the Steiner systems.

Definition 2. The Mathieu groups M_{11}, M_{12}, M_{22}, M_{23}, and M_{24} are defined as follows:

- $M_{11} = \{\sigma \in S_{11} : \sigma(S) \in S(4,5,11) \text{ for all } S \in S(4,5,11)\}$.
- $M_{12} = \{\sigma \in S_{12} : \sigma(S) \in S(5,6,12) \text{ for all } S \in S(5,6,12)\}$.
- $M_{22} = \{\sigma \in S_{22} : \sigma(S) \in S(3,6,22) \text{ for all } S \in S(3,6,22)\}$.
- $M_{23} = \{\sigma \in S_{23} : \sigma(S) \in S(4,7,23) \text{ for all } S \in S(4,7,23)\}$.
- $M_{24} = \{\sigma \in S_{24} : \sigma(S) \in S(5,8,24) \text{ for all } S \in S(5,8,24)\}$.

The Mathieu groups are the five smallest of the 26 sporadic finite simple groups.

Definition 3. Let G be a group which acts on a set X. We say that G is k-transitive if for any two k-tuples $(x_1,\ldots,x_k)$ and $(y_1,\ldots,y_k)$ of elements of X, where all the x_i's are distinct, and all the y_i's are distinct, there is some $g \in G$ so that $g(x_i) = y_i$ for $1 \le i \le k$.

It follows fairly easily from the definitions that

- M_{22} is 3-transitive,
- M_{11} and M_{23} are 4-transitive,
- M_{12} and M_{24} are 5-transitive.

Theorem 4. *The five Mathieu groups are simple.*

We'll prove this theorem in bits, beginning with M_{11} and M_{23}, which Robin Chapman proves elegantly in [3]. We follow his treatment for the simplicity of M_{11} and M_{23}, and we follow Alder in [1] for the proofs of the simplicity of M_{12}, M_{22}, and M_{24}. We begin with some basic facts about permutation groups.

Let p be a prime and $X = \{1, \ldots, p\}$. If $G \leq S_p$ with $|G| = n$, then G acts on X transitively if and only if $p \mid n$. Suppose now that G is transitive, so that $p \mid n$, and let P be a Sylow p-subgroup of G. Let m_G denote the number of Sylow p-subgroups of G so that $m_G = |G : N_G(P)|$, and let $r_G = |N_G(P) : P|$. Thus $n = pr_G m_G$. Sylow's Theorem tells us that $m_G \equiv 1 \pmod{p}$. Now, since $P \leq N_G(P) \leq N_{S_p}(P)$ and $N_{S_p}(P) \cong \mathbb{A}\,\mathrm{GL}_1(\mathbb{F}_p)$, the affine general linear group of invertible affine maps $x \mapsto ax + b \pmod{p}$, we have $p = |P| \leq |N_G(P)| \leq |N_{S_p}(P)| = p(p-1)$, so $r_G \mid (p-1)$.

Lemma 5. *Suppose G is a transitive subgroup of S_p, and $r_G = 1$. Then $G \cong \mathbb{Z}/p\mathbb{Z}$.*

Proof. Suppose $r_G = 1$. Then G contains $m_G(p-1) = n - m_G$ elements of order p. Since elements of order p do not fix any points of X, G contains at most m_G elements which fix at least one point of X. Now, stabilizers G_j of $j \in X$ each contain $|G|/p = m_G$ elements, so all elements of G other than those of order p fix all points of X. But since G acts faithfully on X, it follows that $m_G = 1$, so $n = p$, or $G \cong \mathbb{Z}/p\mathbb{Z}$. ∎

Lemma 5 already allows us to show that certain subgroups of S_p are simple.

Theorem 6. *Let $G \leq S_p$ act transitively on X, where $|G| = pmr$ with $m > 1$, $m \equiv 1 \pmod{p}$ and $r < p$ prime. Then G is a simple group.*

Proof. Clearly, in this case, $r_G = r$ and $m_G = m$. Suppose H is a nontrivial normal subgroup of G. Then the H-orbits of X are permuted transitively by G, so all the orbits have the same size. Thus X consists of a unique H-orbit, so H acts transitively on X. Thus $p \mid |H|$, so H has some Sylow p-subgroup Q. Since $H \lhd G$, H contains all Sylow p-subgroups of G (since they are all conjugate), so $m_H = m_G = m$. Thus $|H| = pmt$ for some $t \mid r$. But $t > 1$ by Lemma 5, and r is prime by hypothesis, so $H = G$. Thus G is simple. ∎

In order to show that the Mathieu groups M_{11} and M_{23} are simple, we need to know something about their orders.

Lemma 7.
- $|M_{11}| = 2^4 \times 3^2 \times 5 \times 11$.
- $|M_{23}| = 2^7 \times 3^2 \times 5 \times 7 \times 11 \times 23$.

We can now put all this information together to show that M_{11} and M_{23} are simple.

Theorem 8. *M_{11} and M_{23} are simple.*

Proof. We first check that M_{11} is simple. Since $11 \mid |M_{11}|$, M_{11} acts transitively on $\{1, \ldots, 11\}$. Since $|M_{11}|/11 = 720 \equiv 5 \pmod{11}$, we have $r_G = 5$ and $m_G = 144$. Thus Theorem 6 tells us that M_{11} is simple.

We now check that M_{23} is simple. Now, $23 \mid |M_{23}|$, so M_{23} acts transitively on $\{1, \ldots, 23\}$. Since $|M_{23}|/23 = 443520 \equiv 11 \pmod{23}$ so $r_G = 11$ and $m_G = 40320$. Again Theorem 6 tells us that M_{23} is simple. ■

In order to prove the simplicity of M_{12} and M_{24} we cite (without proof) a lemma from Rotman [6], p. 193.

Lemma 9. *Suppose G acts faithfully and k-transitively on a set X. If there is some $x \in X$ so that the stabilizer G_x is simple, then*

- *If $k \geq 4$, G is simple.*
- *If $k \geq 3$ and $|X| \neq 2^r$ for any r, then either $G \cong S_3$ or G is simple.*

Together with Theorem 8, Lemma 9 immediately implies the following result:

Theorem 10. *The Mathieu groups M_{12} and M_{24} are simple.*

The one remaining Mathieu group is M_{22}. This again follows from Lemma 9 once we note that a point stabilizer of M_{22} is $\mathrm{PSL}_3(\mathbb{F}_4)$, which is simple. Hence

Theorem 11. *M_{22} is simple.*

3. Binary Golay Codes

Let us return to the code we used to construct the Steiner system $S(5, 8, 24)$. Recall that this code consists of those 24-digit binary strings which are lexicographically first with respect to differing from all prior strings in at least eight positions. We call this code the extended Golay code. If we delete the rightmost digit in each codeword of the extended Golay code, we obtain the perfect Golay code.

The perfect Golay code is very important in coding theory, for it allows us to send data which may be corrupted as efficiently as possible. More precisely, suppose we wish to send a string of twelve binary digits, but we know that up to three of the digits in any string we send may be changed. Then, by sending a certain 23-digit string with the given twelve initial digits, we can correct any three (or fewer) errors in order to recover the original string.

Furthermore, with the Golay code, this recovery is optimal, in the following sense: let $f : \mathbb{F}_2^{12} \to \mathbb{F}_2^{23}$ be the encoding of data (so that the first twelve digits of $f(S)$ are equal to those of S). Then any element of $\mathbb{F}_2^{23}$ differs from some element of $f(\mathbb{F}_2^{12})$ in at most three digits, so any string in $\mathbb{F}_2^{23}$ encodes a unique string from $\mathbb{F}_{12}$.

We mentioned earlier that M_{24} is the automorphism group of the set of codewords with exactly eight 1's. However, it is also true that M_{24} is the automorphism group of the entire extended Golay code. Similarly, M_{23} is the automorphism group of the perfect Golay code.

It turns out that the extended Golay code is linear, in the sense that if $v, w \in \mathbb{F}_2^{24}$ are both in the Golay code, so is $v+w$. Thus the Golay code forms a 12-dimensional vector subspace of $\mathbb{F}_2^{24}$. To prove this, we need some results about impartial combinatorial games. We first introduce the game of nim. Proofs of all statements about general games can be found in either [2] or [4]. More details about the relation between codes and games can be found in [5].

In the game of nim, there are several piles of stones, and two players who will take turns making moves. A move consists of choosing one of the piles with at least one stone, and removing as many stones from that pile as are desired (but at least one). The winner is the player who removes the last stone. (Alternatively, the loser is the player whose turn it is to move when there are no more legal moves remaining.)

We can generalize the game of nim so that we allow various different types of moves, but so that players still alternate moves and so that the unfortunate player who is to move but does not have any legal moves remaining is the loser. We will still have several piles of stones, but now we will have a certain family of so-called turning sets $\{h, i, j, \ldots\}$, where $h > i > j > \cdots$, so that it is legal to replace a pile of size h with piles of size $i, j, \ldots$ A general position therefore is a sum of piles $P_a + P_b + P_c + \cdots$, where P_r denotes a pile of size r. Thus nim consists of turning sets $\{h, i\}$ where $h > i \geq 0$.

We now define precisely what it means for certain positions to be winning for one player.

Definition 12. Let G be an impartial combinatorial game of the type described above. Then we divide the class of positions of games into two outcome classes called $\mathcal{N}$ and $\mathcal{P}$, in the following manner. We say that the zero position, denoted 0, consisting of no piles, is in class $\mathcal{P}$. We then say:

- $G \in \mathcal{N}$ if there is some legal move $G \to G'$ so that $G' \in \mathcal{P}$,
- $G \in \mathcal{P}$ if all legal moves $G \to G'$ satisfy $G' \in \mathcal{N}$.

The $\mathcal{N}$ positions (standing for $\mathcal{N}$ext player) are those in which the player to move can win with optimal play, and the $\mathcal{P}$ positions (standing for $\mathcal{P}$revious player) are those in which the player who just moved wins with optimal play.

Since all are games are decided in a fixed number of moves, any G is either in class $\mathcal{N}$ or $\mathcal{P}$.

We now say what it means to add two games. If G and H are games, a move consists of making a legal move in exactly one of G and H. The game ends when there are no legal moves left in either game.

Definition 13. We say two games G and H are equivalent (and write $G = H$) if whenever X is a game, either $G + X \in \mathcal{P}$ and $H + X \in \mathcal{P}$, or $G + X \in \mathcal{N}$ and $H + X \in \mathcal{N}$. That is, G and H act in the same way with respect to adding other games.

Theorem 14 (Sprague–Grundy)**.** *Every impartial combinatorial game is equivalent to a unique nim game with exactly one pile. All $\mathcal{P}$ positions are equivalent to 0.*

A very important (but easy) corollary of the Sprague–Grundy Theorem is that the sum of two $\mathcal{P}$ positions is again a $\mathcal{P}$ position. Much more generally, however, is the following: Let $*n$ denote the nim position with exactly one pile, consisting of n stones. Then $*m + *n = *(m \oplus n)$, where $m \oplus n$ is the number obtained by writing m and n in binary and adding without carrying (alternatively, addition when considered as being elements of an $\mathbb{F}_2$-vector space). If P is some game position, write $\mathcal{G}(P) = n$ if $P = *n$. $\mathcal{G}(P)$ is called the Grundy value of P.

Since a game with two piles of the same size is clearly a $\mathcal{P}$ position, we do not change anything in the evaluation of a game by deleting piles of the same size in pairs. Thus without loss of generality, we may assume that for each n, the number of piles of size n is either 0 or 1. Therefore, we may express a game position in terms of a binary string $\cdots a_3a_2a_1$, where a_i is the number of strings of size i (either 0 or 1). Since we can express any position as $P = a_1P_1 + a_2P_2 + \cdots$ (where P_i is a pile of size i), we have

$$\mathcal{G}(P) = \bigoplus i \geq 1 a_i \mathcal{G}(P_i).$$

Since the Grundy numbers of general positions are defined by a linear equation, we have the following result:

Theorem 15. *The set of binary strings corresponding to the $\mathcal{P}$ positions is $\mathbb{F}_2$-linear.*

We now consider the following game which helps us to prove the linearity of the Golay code. It suffices to describe the turning sets for the game; they are all sets of size one through seven so that the largest element is at most 24. (Since piles of size 25 or higher are irrelevant, we do not list them.) It is easy to check using Definition 12 that the words of the extended Golay code are exactly the $\mathcal{P}$ positions of this game. Hence Theorem 15 implies the following result:

Theorem 16. *The extended Golay code is $\mathbb{F}_2$-linear.*

4. The Leech Lattice and other sporadic simple groups

The Mathieu groups and the Golay code are related to a lattice in $\mathbb{R}^{24}$, called the Leech lattice, which generates a sphere packing in 24 dimensions. The maximum number of spheres that can be tangent to a fixed sphere in $\mathbb{R}^{24}$ so that all spheres have lattice point centers is 196,560. We now describe how this can be done by giving the coordinates of the centers. We take our fixed sphere to be centered at the origin in $\mathbb{R}^{24}$.

Of the 196,560 tangent spheres, we now describe the centers of 97,152 of them. If we fix a code work α with eight 1's in the extended Golay code (alternatively, an element of the Steiner system $S(5, 8, 24)$), we generate 2^7 sphere centers from α by placing ±2 in the positions in which α has a 1 in such a way that the number of -2 terms is even. We place a 0 everywhere else. This gives $2^7 \times 759 = 97,152$ points.

We now describe another 1,104 sphere centers. We place ±4 in any two of the coordinates and 0 everywhere else. This gives $2^2 \times \binom{24}{2} = 1,104$ points.

We now describe the remaining 97,308 sphere centers. One coordinate is ±3, and the rest are ±1. The sign choices are given by the words in the extended Golay code.

We take the radius of each sphere to be 8.

The Leech lattice allows us to construct more sporadic simple groups. For instance, the automorphism group of the Leech lattice contains the Conway group Co_1 as an index 2 subgroup, and the smaller Conway groups Co_2 and Co_3 are the stabilizers of one vector of the second and third type described above, respectively.

The McLaughlin group, the Higman-Sims group, the Suzuki group, and the Hall-Janko group are also stabilizers of various sets of vectors in the Leech lattice. Thus straightfoward considerations of the Leech lattice provides us with seven of the sporadic groups, sometimes called the second generation of the Happy Family. (The five Mathieu groups are the first generation.)

These sporadic groups related to the Leech lattice as well as the others are fascinating objects, but are unfortunately beyond the scope of this essay.

References

[1] Stuart Alder, *Classical papers in group theory: the Mathieu groups*, Master's thesis, University of East Anglia, 2006.

[2] Elwyn R. Berlekamp, John H. Conway, and Richard K. Guy, *Winning ways for your mathematical plays. Vol. 1*, second ed., A K Peters Ltd., Natick, MA, 2001. MR MR1808891 (2001i:91001)

[3] Robin J. Chapman, *An elementary proof of the simplicity of the Mathieu groups M_{11} and M_{23}*, Amer. Math. Monthly **102** (1995), no. 6, 544–545. MR MR1336642

[4] J. H. Conway, *On numbers and games*, second ed., A K Peters Ltd., Natick, MA, 2001. MR MR1803095 (2001j:00008)

[5] John H. Conway and N. J. A. Sloane, *Lexicographic codes: error-correcting codes from game theory*, IEEE Trans. Inform. Theory **32** (1986), no. 3, 337–348. MR MR838197 (87f:94049)

[6] Joseph J. Rotman, *An introduction to the theory of groups*, fourth ed., Graduate Texts in Mathematics, vol. 148, Springer-Verlag, New York, 1995. MR MR1307623 (95m:20001)

Department of Mathematics, Stanford University, Stanford, CA 94305
E-mail address: simonr@math.stanford.edu