Lecture Notes on Jordan's Theorem on sharp k-transitivity for permutation groups, the five Mathieu simple sporadic groups, Steiner systems and related notions (In preparation; Not for distribution)

Jeff Steif

November 6, 2021

Contents

1 Introduction and overview

The goal of these notes is to present some interesting results in a somewhat direct way and with only requiring what might be called a "modest" background (modest meaning the group theory which would be contained in a masters or perhaps Ph.D. level course in algebra). In this section, we try to give the big picture in terms of what we want to do. It is long but the attempt is to give a lot of the coming story. Obviously, this introduction is vague in places and the details will be in the relevant sections.

While relevant definitions will be given in the relevant sections, we do need to give just one definition right away. S_n will always denote the symmetric group on n elements.

Definition 1.1. *A subgroup G of S_n is called (sharply) k-transitive if for each pair of ordered k-tuples of elements from $\{1,\ldots,n\}$, $(a_1,\ldots,a_k)$ and $(b_1,\ldots,b_k)$ with the a_i's distinct and the b_i's distinct, there exists (a unique) $g \in G$ so that for each i, $g(a_i) = b_i$. (Note that 1-transitivity corresponds to the usual notion of a transitive subgroup.)*

The following three examples of sharp transitivity are worth noting but considered to be trivial. The full permutation group S_n is both sharply n-transitive and sharply $(n-1)$-transitive. The alternating group $A_n \subseteq S_n$ is sharply $(n-2)$-transitive. The term nontrivial when used with a permutation group will mean that it is not one of the three above trivial examples.

We can already now state Jordan's Theorem from 1872.

Theorem 1.2. *Assume $G \subseteq S_n$ is sharply k-transitive and nontrivial.*
1. If $k = 4$, then we must have $n = 11$ (and it would follow that $|G| = 11 \times 10 \times 9 \times 8 = 7920$ should such a G exist).
2. If $k = 5$, then we must have $n = 12$ (and it would follow that $|G| = 12 \times 11 \times 10 \times 9 \times 8 = 95,040$ should such a G exist).
3. If $k \geq 6$, there is no such example.

The proof of this will be given in Section 3 and is supposed to be the start of the story.

Once one has Jordan's Theorem, a number of questions immediately come to mind. The first obvious one is whether there are in fact sharp 4-transitive subgroups of S_{11} and sharp 5-transitive subgroups of S_{12}. The answer is (of course since otherwise we wouldn't be here) yes and the corresponding two groups will be our first two so-called *sporadic* simple groups.

Theorem 1.3. *1. There is a group $M_{11} \subseteq S_{11}$ which is sharply 4-transitive and which in addition is simple.*
2. There is a group $M_{12} \subseteq S_{12}$ which is sharply 5-transitive and which in addition is simple.

This theorem other than the simplicity will be proved in Sections 5 and 6. The simplicity will be proved in Section 8.

Background on simple groups: Simple groups are those groups which have no nontrivial normal subgroups. These are viewed as important since all finite groups can be "built up" from simple groups using what is called a composition series. This is not perhaps as satisfying as it might sound. It *is* true that to any group we can assign a finite number of simple groups (those appearing in a so-called composition series) and it *is* true that, while one can have different composition series, the simple groups appearing in any one of them are the same even with repetitions (this is the so-called Jordan–Hölder theorem). In this way, we have a natural way to assign to any group a unique finite set of simple groups from which it is "built". *However,* unlike a finite number of prime numbers which combine to give you a unique integer, there is most of the time more than one group corresponding to a collection of simple groups, i.e. simple groups can be "put together" in different ways to construct different groups. So, understanding all groups amounts to (1) understanding/classifying all simple groups and (2) describing all ways a finite number of simple groups can be "put together". The famous so-called classification of finite simple groups solves (1). It is considered to be one of mathematics' largest achievements. There are various infinite families of finite simple groups, like the alternating groups and the so-called projective special linear groups as we will see, among others. However, there are 26 simple groups which do not fall into one of these infinite families; these 26 groups are called the simple sporadic groups. The so-called monster group which has size

$$808,017,424,794,512,875,886,459,904,961,710,757,005,754,368,000,000,000$$

is the largest of them. The main point of these notes is to introduce the first five of these sporadic simple groups which were found by Mathieu in 1861 and 1873. It seems to have taken about a hundred years before the 6th simple sporadic group was found. □

What about the other three Mathieu simple sporadic groups? Note the word **sharp** in Jordan's Theorem. It does not rule out non-sharp k-transitive subgroups and perhaps there are lots of k's and n's for which they exist. In fact, there are not and the following is another big theorem but its proof requires the classification of finite simple groups (and hence of course we don't discuss this here).

Theorem 1.4. *Assume $G \subseteq S_n$ is k-transitive and nontrivial.*
1. If $k = 4$, G must be M_{11} or M_{12} or one of two further groups $M_{23} \subseteq S_{23}$ and $M_{24} \subseteq S_{24}$, both of which are also simple.
2. If $k = 5$, G must be M_{12} or M_{24}.
3. If $k \geq 6$, there is no such example.

In Section 7, we will prove the existence of M_{23} and M_{24} as well as that of a group M_{22} giving us in total our five simple sporadic groups. The constructions will be based on a technique for extending group actions, so-called "transitive extensions". To motivate it, we first mention that it turns out to be the case that for $i = 11, 22, 23$, a point stabilizer of M_{i+1} is M_i.

Now, if we have $G \subseteq S_{n+1}$, we trivially can get $H \subseteq S_n$, simply by letting $H = G_{n+1}$, the stabilizer of $n+1$. And of course we can iterate this procedure. If G is k-transitive, we can iterate this $k-1$ times and still get a transitive subgroup (we'll see later why this $k-1$ iterated action is still transitive). The more difficult and subtle question is whether one can reverse this procedure and go "upwards" rather than "downwards". This can be formulated as follows.

Question: If $H \subseteq S_n$ is transitive, does there exist $G \subseteq S_{n+1}$ *transitive* such that $H = G_{n+1}$, the G-stabilizer of $n+1$?

The answer is not always. A simple example where such a G does not exist will be given in Section 5 where, more importantly, a nontrivial sufficient condition for "extendability" will also be given and later exploited to construct our five simple sporadic groups. This extension theorem will allow us to construct M_{12} from M_{11} and M_{24} from M_{23} and the latter from M_{22}. But how do we start? Where do we get M_{11} and M_{22} from in order to start this iterative constructive procedure? For M_{22}, it will itself be constructed as an extension of a natural action of the so-called projective special linear group $PSL(3,4)$ on a 21 element set, the latter set being the set of 1-dimensional subspaces of a 3-dimensional vector space over F_4, the field with 4 elements: the latter is also called a projective plane.

On the other hand, the construction of M_{11} via extension will first lead us to Section 4 which answers another question which is naturally posed when seeing Jordan's Theorem. One could ask what happens if $k = 3$ in Jordan's Theorem; i.e. do there exist lots of sharply 3-transitive permutation groups or just a finite number. The following theorem answers this question, providing us with two infinite such families. Also the group $M_{10} := M(9)$ described in the next theorem is the group which will extend up to M_{11} to construct the latter. The group PGL (called the projective general linear group) appearing below as well as the PSL group mentioned above will be precisely defined in Section 7.1. Also, the different concepts in this theorem will be explained in due course.

Theorem 1.5. *1. For each prime p and integer k, letting $q = p^k$, there is a sharply 3-transitive subgroup $L(q)$ of S_{q+1}. Moreover, this will correspond to a faithful action of $PGL(2, q)$ on the "projective line" given by the 1-dimensional subspaces of a 2-dimensional vector space over the field F_q or equivalently will be the set of linear fractional transformations on this "projective line".*
2. For each prime $p \geq 3$ and even integer k, letting $q = p^k$, there is a (different) sharply 3-transitive subgroup $M(q)$ of S_{q+1}.
3. The two groups $L(q)$ and $M(q)$ (when both are defined) which are of order $(q + 1)q(q - 1)$ are not isomorphic.
4. Every sharply 3-transitive action is equivalent to one of these two.

In Section 4, we will prove the first two parts and the third part in the special case $q = 9 = 3^2$ (which is the relevant case for developing the Mathieu groups) by distinguishing their respective Sylow 2-subgroups. The action of $M(q)$ is very similar to the action of $L(q)$ but comes with a twist; it is sometimes called a semi-linear mapping. Part 4 is due to Zassenhaus and appears very nontrivial and will not be discussed.

Not only will this "extension procedure" allow us to construct our groups but it will also assist in proving simplicity since as we will see in Section 8.1, this extension procedure will often (although not always) preserve simplicity. So, when the smoke clears, simplicity of M_{11} will yield simplicity of M_{12} and simplicity of M_{22} will yield simplicity of M_{23} which in turn will give simplicity of M_{24}. But, even given all that, how do we start the process? We will prove simplicity of M_{11} directly from first principles using *only* the size of the group and that it is a 2-transitive subgroup of S_{11} while the fairly well known simplicity of $PSL(3, 4)$ (whose proof will (almost) be reviewed in Section 7.1) allows us to start moving up starting from M_{22} to prove the simplicity of the latter 3 groups. However, we mention that we cannot prove M_{11} is simple by trying to conclude its simplicity from M_{10} since the latter is in fact not simple having A_6 as an index 2 subgroup (but is nonetheless not isomorphic to S_6).

Note that we could easily read off the orders of the first two sporadic groups M_{11} and M_{12} simply by knowing that they are sharp 4 or 5-transitive. Since

M_{22}, M_{23} and M_{24} are not *sharply* transitive, their orders are not read off in this simple way. However, from what is written above, it is not too hard to calculate their orders inductively. By construction, M_{24} will have M_{23} as its 1 point stabilizer, M_{23} will have M_{22} as its 1 point stabilizer and finally M_{22} will have $PSL(3,4)$ as its 1 point stabilizer. It follows (since in general the size of an orbit of an action is the index of its stabilizer), we have $|M_{24}| = 24|M_{23}|$, $|M_{23}| = 23|M_{22}|$ and $|M_{22}| = 22|PSL(3,4)|$. In Section 7.1, we will see that $|PSL(3,4)| = 20,160$ leading, using the above, to $|M_{22}| = 443,520$, $|M_{23}| = 10,200,960$ and $|M_{24}| = 244,823,040$. We lastly mention that we can go downwards from $PSL(3,4)$ (which is sometimes called M_{21} for obvious reasons) meaning we can take a point stabilizer of $PSL(3,4)$ obtaining a subgroup called M_{20} whose order would then be $20,160/21 = 960$. However, M_{20} will not be a simple group.

After doing all of the above, we will discuss so-called Steiner systems from combinatorics. The connection to the Mathieu groups is that the latter are the automorphism groups of so some of these combinatorial systems.

The notes are divided in two parts. The first nine sections have to do with Jordan's Theorem, the construction and simplicity of the five Mathieu simple sporadic groups and other things about permutation groups. This part is all group theory. The second shorter part of the notes introduces Steiner systems and connects up the five Mathieu simple sporadic groups with combinatorics.

We end this introduction with a partial list of all nonabelian simple groups whose order is at most 100,000 and also list the first nine simple sporadic groups. The first list includes only two sporadic groups which are the first two of the five Mathieu groups. Exclamation marks indicate cases which are worthy of extra notice.

Simple groups up to order 100,000		First 9 of 26 sporadic groups	
1. ! $A_5 \cong PSL(2,4) \cong PSL(2,5)$	60	1. M_{11}	7920
2. ! $PSL(2,7) \cong PSL(3,2)$	168	2. M_{12}	95,040
3. ! $A_6 \cong PSL(2,9)$	360	3. Janko J_1	175,560
4. $PSL(2,8)$	504	4. M_{22}	443,520
5. $PSL(2,11)$	660	5. Janko J_2	604,800
6. $PSL(2,13)$	1092	6. M_{23}	10,200,960
7. $PSL(2,17)$	2448	7. Higman-Sims	44,352,000
8. ! A_7	2520	8. Janko J_3	50,232,960
9. $PSL(2,19)$	3420	9. M_{24}	244,823,040
10. $PSL(3,3)$	5616		
11. ! Projective special unitary	6048		
12. $PSL(2,23)$	6072		
13. $PSL(2,25)$	7800		
14. !! The first Mathieu group M_{11}	7920		
15. $PSL(2,27)$	9828		
16. $PSL(2,29)$	12,180		
17. $PSL(2,31)$	14,880		
18. ! $A_8 \cong PSL(4,2)$	20,160		
19. ! $PSL(3,4)$	20,160		
20-29. Various groups			
30. !! The second Mathieu group M_{12}	95,040		

(1899) A_8 and $PSL(3,4)$ are distinguished since A_8 has elements of order 6 and 15

$$(123)(45)(67), \quad (12345)(678)$$

while one can show that $PSL(3,4)$ does not.

There are infinitely many n which have two nonisomorphic simple groups of that order. But none which have 3.

2 Background

2.1 Background: Part 1: (Many readers will know most of what is written here and can skip most or all of it)

It is clear that one will need to have some background in group theory to follow these notes, but I hope the standard algebra courses will cover everything you need. The following things for example should be known or reviewed if you don't recall them.

Subgroups, normal subgroups, homomorphisms, cosets, quotient groups, centralizers, centers, normalizers, conjugate elements, conjugate subgroups, the isomorphism theorems, the Sylow theorems, symmetric groups and their standard properties.

Another important topic which I assume you are familiar with is group actions and many of their various properties. Perhaps this section can be used to recall/tell you about some of these. A great place to quickly read about them (and many other things) is on Keith Conrad's homepage where he has many interesting short lecture notes.

We do an extremely fast repetition of group actions and state without proof a number of standard results.

Our groups and sets on which they act will always be finite; many of these statements hold in a more general context.

Definition 2.1. *If G is a group and X is a set, a* **group action** *of G on X is a mapping*

$$f : G \times X \to X$$

satisfying

$$(1)\ f(g_1 g_2, x) = f(g_1 f(g_2, x))\ \forall g_1, g_2 \in G, x \in X$$

and

$$(2)\ f(1, x) = x\ \forall x \in X$$

where 1 will always denote the identity element in G.

Remarks.
(i). It is simplest to abbreviate $f(g, x)$ by (g, x) which we now do.
(ii). You should think of (g, x) as "the result obtained when g acts on x".

So “g takes x to another element of X” which we call (g, x).
(iii). The key equality (1) says that if you first let g_2 act on x and then act on the result by g_1, you end up with the same thing as you would if you had first multiplied g_1 and g_2 in the group and then let the product act on x. (It has a very similar flavor, but not really the same thing, as the definition of a homomorphism of groups.)
(iv). We have chosen to do ”left actions” rather than ”right actions”. Of course, there is no essential difference but making this convention ties your hands later on (if you don’t want to define things by inverses). For example, composition of permutations should now be from right to left and the conjugate of a subgroup H by g should be gHg^{-1} rather than $g^{-1}Hg$.
(v). Group actions are probably the cleanest way to prove all the Sylow Theorems.

Exercise: Show that if you fix g, then the mapping from X to X given by x goes to (g, x) is a bijection. Is condition (2) in the definition of a group action superfluous?

Exercise. Show that a group action of G on X is equivalent to a group homomorphism from G to S_X where the latter is the symmetric group on X. (Hint. Use the previous exercise).

We now discuss the various key players that arise when studying group actions.

Definition 2.2. *We write $x \sim y$ if there exists g such that $(g, x) = y$. This easily gives an equivalence relation whose equivalence classes are called* **orbits**. *The group action is called* **transitive** *if there is only one orbit.*

Definition 2.3. *Given a group action of G on X and given $x \in X$, the* **stabilizer** *of x, denoted G_x, is $\{g : (g, x) = x\}$. (This is trivially a subgroup)*

Often in books, this is denoted by S_x (”S” for stabilizer) but we use G_x here since we might have more than one group acting on a set at the same time.

Exercises.
1. If x and y are in the same orbit, then G_x and G_y are conjugate subgroups; i.e. there exists $g \in G$ so that $gG_xg^{-1} = G_y$.
2. Show that $\cap_x G_x$ is the kernel of the homomorphism from G to S_X which represents the group action. (The action is called **faithful** if this kernel is trivial.)
3. Denoting the orbit of x by O_x, one has

$$|O_x| = [G : G_x]$$

where the latter denotes the index of the subgroup G_x in G. (This implies of course that $|O_x|$ divides $|G|$.) (Hint: Map the left coset gS_x to (g, x) and check it is well defined and bijective.)
4. There is a slightly more combinatorial way to do the last exercise which from group theory only uses the fact that the set of group elements taking x to y is a coset of G_x and hence has the same size. Namely, look at the set $\{(g, y) : gx = y\}$ and double count.
5. Verify that G acts on the collection of left cosets of any subgroup H by defining

$$(g_1, g_2H) = g_1g_2H.$$

6. Define a natural notion of two G-actions being equivalent and prove that if an action is transitive, then the action is equivalent to the above action on the set of left cosets of G_x for any of the stabilizers.
7. If $G \subseteq S_n$ is transitive and abelian, show that all of the stabilizers of G are trivial.

2.2 Background: Part 2: (Many readers will probably want to look at this)

This section will list a number of somewhat standard facts about groups which we will use later. It is convenient to collect them here.

Throughout, G is a permutation group, ie., a subgroup of S_X for some given finite set X. While we are mostly interested in permutation groups meaning subgroups of the symmetric group (which are essentially faithful group actions), we will need to deal with more general group actions for various proofs. One can define a k-transitive action in the same way as for a subgroup. However, note that if the action is not faithful (i.e., has a kernel), then it is impossible for the action to be sharp k-transitive (using the obvious definition of sharp k-transitive for an action). For $k \leq |X|$, let $O_k(X)$ denote the collection of (ordered) distinct k-tuples from X. Clearly an action of G on X yields a natural action of G on $O_k(X)$ for each k. The following lemma is left to the reader and basically is just reformulations of definitions.

Lemma 2.4. *1. An action of G on X is k-transitive if and only if the induced action of G on $O_k(X)$ is transitive.*
2. $G \subseteq S_X$ is sharp k-transitive if and only if G, viewed as a subgroup of $S_{O_k(X)}$, is sharp 1-transitive.

Another lemma which we will use constantly and which is also left to the reader is the following.

Lemma 2.5. *1. For $k \geq 2$, an action of a group G on $\{1, \ldots, n\}$ is k-transitive if and only if for any (or some) $i \in \{1, \ldots, n\}$, the action of G_i on $\{1, \ldots, n\} \backslash \{i\}$ is $(k-1)$-transitive.*
2. For $k \geq 2$, $G \subseteq S_{\{1,\ldots,n\}}$ is sharp k-transitive if and only if for any (or some) $i \in \{1, \ldots, n\}$, $G_i \subseteq S_{\{1,\ldots,n\} \backslash \{i\}}$ is sharp $(k-1)$-transitive.

Here are other somewhat standard theorems in group theory which we will appeal to without comment.

The commutator subgroup G' of a group G is the subgroup generated by the collection of elements $\{aba^{-1}b^{-1}\}$. This turns out to be a normal subgroup and it gives a measure of the nonabelianness of a group. Modding out by this subgroup is the minimal way to abelianize a group as stated next.

Theorem 2.6. *If $H \lhd G$, then G/H is abelian if and only if $G' \subseteq H$.*

Theorem 2.7. *Any group of order p^k for some prime p (such groups are called p-groups) has a nontrivial center.*

Long remark on primitivity (which is not really crucial to read to understand the notes). Primitivity is an important property that a transitive action might or might not have. The property sits between transitivity and 2-transitivity. The precise definition is that if we have a group G acting transitively on a set X, then there should not be any nontrivial *block* where a block is a subset B of X such that for every $g \in G$, B and gB are either disjoint or identical. Nontrivial means that B is neither X nor a singleton since these are always blocks. Another way to think about a primitive group action is that it is the analogue of a simple group in the context of group actions. Namely, primitivity means that the action has no nontrivial homomorphic images. (Exercise: Make this precise and show that the two descriptions are the same.)

I am still not sure if we will need this concept in these notes but we will see. The reason is that the various times that we might want to use primitivity, we will actually be in a 2-transitive situation. Hence, we might just state and prove the simpler versions and mention when they are true also in the primitive case. Usually, the proofs are a little simpler assuming the stronger 2-transitivity. One example where the proofs differ a lot is in another one of Jordan's theorems. It is immediate that a 2-transitive subgroup of S_n which contains a transposition is all of S_n (since you can get via conjugation to all of the transpositions which of course generate.) However, more interestingly, the same conclusion holds if 2-transitivity is replaced by primitivity. However, the proof is a little harder, although admittedly not *a lot* harder.

Examples.
1. Z/n acting on $\{1,\ldots,n\}$ cyclically is always transitive and never 2-transitive. It is primitive if and only if n is prime.
2. A slightly more interesting example perhaps of a primitive action which is not 2-transitive is the dihedral group D_{2p} acting on a p-gon where $p > 3$ is prime. If p is not prime, we will not have primitivity but still have transitivity. □

2.3 Some Notation (which will be updated as the notes are written)

$C_G(x)$ is the centralizer of x, which is $\{y \in G : xy = yx\}$.

$N_G(H)$ is the normalizer in G of the subgroup H, which is

$$\{y \in G : yHy^{-1} = H\}.$$

3 Proof of Jordan's Theorem on sharp k-transitivity (Theorem 1.2)

Strange as it might seem, it turns out cleanest to first prove two special cases of the theorem which we do in the following proposition.

3.1 Two special cases of Jordan's Theorem on sharp k-transitivity

Proposition 3.1. *There is no sharply* 4*-transitive subgroup of* S_{10} *and no sharply* 6*-transitive subgroup of* S_{13}*.*

To do this we first need the following lemma.

Lemma 3.2. *Let* G *be a* k*-transitive subgroup of* S_X *and let* $Y := \{a_1, \dots, a_k\} \subseteq X$*. Let* H *be the pointwise stabilizer of* Y*; i.e.,*

$$H := \{g \in G : g(a_i) = a_i \text{ for each } i\}.$$

Fix p*, let* P *be a Sylow* p*-subgroup of* H *and* $N_G(P)$ *be the normalizer of* P *in* G*. Then* $N_G(P)$ *is* k*-transitive on the set* F_P *of fixed points of* P*; i.e., on*

$$F_P := \{x : gx = x \text{ for each } g \in P\}.$$

Remarks:
1. This is trivial if G is sharp k-transitive since then H would be trivial which would imply that $N_G(P) = G$ and F_P would be X.
2. Similarly it is trivial if p does not divide $|H|$ since again P would be trivial which again implies that $N_G(P) = G$ and F_P would be X.
3. Clearly $Y \subseteq F_P$.
4. If P is the unique Sylow p-subgroup of H, then it would be normal in H implying that $H \subseteq N_G(P)$. Otherwise, there is no general containment relationship between H and $N_G(P)$.

Proof: We first have to show that $N_G(P)$ in fact preserves F_P. However, if $x \in F_P$, $\pi \in N_G(P)$ and $g \in P$, we have, for some $g' \in P$,

$$g\pi x = \pi g' x = \pi x$$

Since $\{a_1, \dots, a_k\} \subseteq F_P$, it suffices to show that if $\{b_1, \dots, b_k\} \subseteq F_P$, then there is $\pi \in N_G(P)$ which takes each a_i to b_i.

By assumption, there is $\sigma \in G$ which takes each a_i to b_i and it trivially follows that $\sigma^{-1}P\sigma$ pointwise fixes $\{a_1, \ldots, a_k\}$ since P pointwise fixes the b_i's. Hence $\sigma^{-1}P\sigma$ is a subgroup of H. Since all Sylow p-subgroups of H are conjugate (in H!!), we conclude there is a $\tau \in H$ with

$$\sigma^{-1}P\sigma = \tau^{-1}P\tau.$$

But this gives that $\tau\sigma^{-1} \in N_G(P)$ and clearly $\tau\sigma^{-1}$ takes each b_i to a_i. QED

Remark: The above argument is often called a "Frattini-type" argument.

Proof of Proposition 3.1

Part 1. If G is a sharply 4-transitive subgroup of S_{10}, then $|G| = 10 \cdot 9 \cdot 8 \cdot 7$. Hence a Sylow 7-subgroup P of G is cyclic of order 7. Since the only order 7 permutations in S_{10} are 7-cycles, WLOG, $P = \langle \pi \rangle$ with

$$\pi = (1234567).$$

We apply Lemma 3.2 with $k = 3$ and $Y = \{8, 9, 10\}$. This then defines our subgroup H. Clearly P is a Sylow 7-subgroup in H as well and we also observe that $F_P = Y$ trivially. The lemma then allows us to conclude that $N_G(P)$ is 3-transitive on Y. This means that we have a surjective homomorphism ϕ from $N_G(P)$ to S_Y.

In addition, $N_G(P)$ acts on P by conjugation with kernel $C_G(P)$. Since conjugations are automorphisms of P, we get $N_G(P)/C_G(P)$ imbeds into $Aut(P)$ which is an abelian group (and is in fact just $Z/6$). It follows from Theorem 2.6 that $C_G(P) \supseteq N_G(P)'$. It is easy to check that taking commutator subgroups commutes with homomorphisms which leads, using also the previous paragaraph, to

$$\phi(C_G(P)) \supseteq \phi(N_G(P)') = (\phi(N_G(P)))' = (S_Y)' = Z/3.$$

This implies (say by the correspondence theorem and Cayley's Theorem) that $C_G(P)$ contains an element σ of order 3, which must consist of 1,2 or 3 3-cycles. (Note that the 3-cycle $(8, 9, 10)$ is trivially in $C_{S_{10}}(P)$ but we need the order 3 element to be in G.)

Clearly $\sigma\pi$ has order 21 (since σ and π commute and have relatively prime orders) and hence $\sigma\pi$ must consist of a 7-cycle and a 3-cycle. Finally now, the permutation $(\sigma\pi)^7$ is clearly nontrivial (due to the 3-cycle) but fixes the whole 7-cycle. This contradicts sharp 4-transitivity.

(In fact, it is not hard to argue that the order 3 element σ can in fact only involve $\{8, 9, 10\}$ but the above argument does not require us to do that; this follows from the fairly easy fact that $\sigma \in C_G(P) \subseteq C_{S_{10}}(P) = P \times S_Y$.)

Part 2. This second part is proved in almost the exact same way. If G is a sharply 6-transitive subgroup of S_{13}, then $|G| = 13 \cdot 12 \cdot 11 \cdot 10 \cdot 9 \cdot 8$. Then a Sylow 5-subgroup of G has order 5 and its generator π must be either a single 5-cycle or two disjoint 5-cycles. If π had only one 5-cycle, it would fix 8 elements, contradicting sharp 6-transitivity. So WLOG

$$\pi = (12345)(678910).$$

We again apply Lemma 3.2 with $k = 3$ and $Y = \{11, 12, 13\}$. Exactly as before, P is a Sylow 5-subgroup in H and $F_P = Y$ trivially. The lemma then allows us to conclude that $N_G(P)$ is 3-transitive on Y. We then can apply verbatum the argument of Part 1 to conclude that $C_G(P)$ contains a σ of order 3. (The only difference is that $Aut(P)$ is now $Z/4$ instead of $Z/6$.) Then $\pi\sigma$ necessarily has order 15 and hence consists of a 5- and a 3-cycle, two 5 cycles and one 3-cycle or one 5 cycle and two 3-cycles. It follows that $(\pi\sigma)^5$ is nontrivial but fixes at least 7 points. This contradicts sharp 6-transitivity. QED

3.2 General proof of Jordan's Theorem on sharp k-transitivity

We can now proceed with the

Proof of Theorem 1.2.

Proof of Statement 1

Main Step: for $k = 4$, show that n must be 10 or 11.

Step 1: Ruling out $n \leq 7$.

Being 4-transitive, $|G| = n \cdot (n-1) \cdot (n-2) \cdot (n-3)$. If $n = 4$, then $|G| = 4!$ and hence $G = S_4$, which is trivial and ruled out by assumption. If $n = 5$, then $|G| = 5!$ and hence $G = S_5$, again trivial. If $n = 6$, then $|G| = 6!/2$ and hence $G = A_6$ (since S_n only has one index two subgroup), again trivial.

$n = 7$. Then $|G| = 7!/6$ and hence G is an index 6 subgroup of S_7. S_7 acts on the left cosets of G, S_7/G (a set of size 6), by translations and the kernel of this action is the so-called core of G, denoted by $core(G)$. This is the largest normal subgroup contained in G or alternatively $\cap_{g \in S_7} gGg^{-1}$. This gives

$$|S_7/core(G)||6!$$

We have $core(G) \lhd S_7$ but S_7 has only three normal subgroups, A_7 and the two trivial ones. $core(G)$ cannot be just 1, since 7! does not divide 6!. It cannot be all of S_7 since it is contained in G which is smaller than S_7. Hence $core(G)$ must be A_7 but this contradicts the fact that $|core(G)| \leq |G| = 7!/6$.

We now assume $n \geq 8$.

Step 2: Any 2 involutions in G are conjugate (in G).
Proof: Any involution in G must have at least two (and in fact at least three) 2-cycles since otherwise it would have four fixed points contradicting sharp 4-transitivity. Let π and σ be two involutions in G which WLOG look like

$$\pi = (12)(34)\cdots, \quad \sigma = (ab)(cd)\cdots$$

where $\cdots$ means we know nothing further about the permutation. By 4-transitivity, there is $\tau \in G$ taking 1 to a, 2 to b, 3 to c and 4 to d. Hence

$$\tau\pi\tau^{-1} = (ab)(cd)\cdots$$

and hence by sharp 4-transitivity,

$$\tau\pi\tau^{-1} = \sigma.$$

Step 3: A 4-Klein group H with very simple structure is contained in G.
By 4-transitivity, we have two elements in G as follows.

$$\pi = (1)(2)(34)\cdots, \quad \sigma = (12)(3)(4)\cdots.$$

(This π and σ have nothing to do with the π and σ in step 2 which was just used locally in that proof.) Since both π^2 and σ^2 have at least four fixed points, these are each 1 by sharp 4-transitivity; i.e. π and σ are involutions. Next since

$$\pi\sigma = (12)(34)\cdots = \sigma\pi,$$

by sharp 4-transitivity we have $\pi\sigma = \sigma\pi$. Hence $H := \langle\{\pi, \sigma\}\rangle$ is the 4-Klein group $Z/2 \times Z/2$. Again, by sharp 4-transitivity, π has at most 3 fixed points and hence either no fixed points other than 1 and 2 or one other fixed point which we can take to be 7. The argument will now continue with the assumption that this third fixed point exists but the argument is easily modified if this third fixed point does not exist.

Since π and σ commute, σ permutes the fixed points of π, namely $\{1, 2, 7\}$. (Check this.) Since σ has the transposition $(1, 2)$, σ must also fix 7. Let now $\tau := \pi\sigma$ be the fourth element of H. By step 2, all involutions are conjugate and hence τ has the same number of fixed points as π. So $\tau(7) = 7$ and it has two other fixed points, which can't be any of 1,2,3,4 and hence we can call them 5 and 6. So

$$\tau = (12)(34)(5)(6)(7)\cdots .$$

Since π and τ commute, π permutes the fixed points of τ, namely 5,6,7. Since π cannot have 4 fixed points, we must therefore have

$$\pi = (1)(2)(34)(56)(7)\cdots .$$

and the exact same argument gives

$$\sigma = (12)(3)(4)(56)(7)\cdots .$$

Observe the exact forms of these three permutations.

Step 4: $C_G(H) = H$.
H abelian gives of course $H \subseteq C_G(H)$. Choose now $\rho \in C_G(H)\backslash\{1\}$. ρ must then permute the fixed point sets of each of the three nontrivial permutations in H, i.e. must permute the sets $\{1,2,7\}$, $\{3,4,7\}$ and $\{5,6,7\}$. ρ must then fix 7 and therefore must look like

$$\rho = (12)^*(34)^*(56)^*(7)\cdots .$$

where $*$ means that the transposition may or may not be present. Since $\rho \neq 1$, ρ has at most three fixed points (again by sharp 4-transitivity) and hence at least two of the potential transpositions must be present. So WLOG

$$\rho = (12)(34)\cdots$$

and hence (by sharp 4-transitivity) $\rho = \tau \in H$, as desired.

Step 5: Structure of the H-orbits and the normalizer of H.
Clearly $\{1,2,3,4,5,6,7\}$ is a union of H orbits and includes all fixed points for any nonidentity element in H. Since $n \geq 8$, there must be a disjoint H-orbit X which must have size four since H has size four and all the stabilizers are trivial.

Letting $S \subseteq G$ be the collection of permutations in G which send X to X, we claim that

$$S = N_G(H).$$

(Of course $H \subseteq S$ trivially since X is an H-orbit.) First, the sharp 4-transitivity gives (why?) that

$$S \cong S_4.$$

Since $H \subseteq S$ and S_4 contains only one copy of the 4 Klein group where no nonidentity element has a fixed point, we must have, letting $X = \{x_1, x_2, x_3, x_4\}$,

$$H = \{1, (x_1x_2)(x_3x_4), (x_1x_3)(x_2x_4), (x_1x_4)(x_2x_3)\}.$$

(We are identifying H with the restriction of H to X, using sharp 4-transitivity.) Hence $H \lhd S$ and so $S \subseteq N_G(H)$. For the other inclusion, we will just check cardinalities. We have, using step 4,

$$[N_G(H) : H] = [N_G(H) : C_G(H)] \leq |AutH| = 6.$$

The last equality is clear since $AutH \cong S_3$ while the inequality follows from letting $N_G(H)$ act on H by conjugation (which yield automorphisms of H) and noting that the kernel of this action is $C_G(H)$. The last display gives that $|N_G(H)| \leq 24 = |S|$, proving that $S = N_G(H)$.

Step 6: At most one H-orbit outside of $\{1, 2, 3, 4, 5, 6, 7\}$:
Assume that there was another H-orbit X' outside of $\{1, 2, 3, 4, 5, 6, 7\}$ besides X. Letting $S' \subseteq G$ be the collection of permutations in G which send X' to X', step 5 tells us that we also have $S' = N_G(H)$ and H has the same form as above (since there was nothing special about the H-orbit X).

Choose $g \in N_G(H)\backslash H$ which induces the permutation $(x_1)(x_2)(x_3x_4)$ on X. What permutation does this give on X'? Since g is not in H, it is not one of the 4 Klein group permutations above and since it has order 2, it must also have two fixed points in X'. This would give four fixed points altogether contradicting sharp 4-transitivity. Hence this second H-orbit cannot exist. (Admittedly, this last argument is a little subtle.)

This allows us to conclude that $n = 10$ or 11 depending on whether the fixed point 7 was present and thereby completes the main step. Since Lemma 3.1 says that there is no sharp 4-transitive subgroup of S_{10}, we conclude that n must be 11, completing the first part of Jordan's Theorem.

Proof of Statement 2

We now move to $k = 5$ which will be short using the $k = 4$ result. If G is a sharp 5-transitive subgroup of S_n, then the stabilizer $G_{\{1\}}$ must be a sharp 4-transitive subgroup of S_{n-1}. From the first part, we conclude that either (1) $n - 1 = 11$ or (2) $G_{\{1\}}$ is trivial meaning it is either S_{n-1} or A_{n-1}. In the second case, $G_{\{1\}}$ would have size $(n-1)!$ or $(n-1)!/2$. Hence either $n = 12$ or G has size $n!$ or $n!/2$ and since we assume nontrivialness, we get n must be 12, proving the second statement.

Proof of Statement 3

If G is a sharp 6-transitive subgroup of S_n, then the stabilizer $G_{\{1\}}$ must be a sharp 5-transitive subgroup of S_{n-1}. From the second part, we conclude that either (1) $n - 1 = 12$ or (2) $G_{\{1\}}$ is either S_{n-1} or A_{n-1}. In the second case, as above, (2) would lead to G being trivial, which we assume it is not. Hence n=13. However, Lemma 3.1 says that there is no sharp 6-transitive subgroup of S_{13}. We conclude that there is no sharp 6-transitive subgroup of S_n for any n.

One finally proves there is no sharp k-transitive subgroup of S_n for $k \geq 7$ which is not trivial by induction since if one existed, then the 1-point stabilizer of such a subgroup would yield a nontrivial $(k-1)$-transitive subgroup of S_{n-1}.
QED

4 Description of all sharply 3-transitive permutation groups

In this section, we will give two distinct constructions of an infinite family of sharply 3-transitive subgroups, proving most of Theorem 1.5.

4.1 First construction of a family of sharply 3-transitive subgroups

Let F_q be a finite field with $q = p^k$ elements where p is prime and k is an integer. (The case $q = 9$ will be relevant to the construction of the two smallest order Mathieu simple sporadic groups.) Let $X = F_q \cup \{\infty\}$ and so $|X| = p^k + 1$.

Consider the subgroup $L(q)$ of S_X consisting of the permutations

$$\{x \to \frac{ax+b}{cx+d}\}$$

where a, b, c, d belong to F_q and $ad - bc \neq 0$. (This map takes ∞ to $\frac{a}{c}$ and takes $\frac{-d}{c}$ to ∞.) One needs to check that these are bijections and closed under composition but that is easy; one can check that composition corresponds to matrix multiplication. These mappings are called linear fractional transformations, at least in a complex analysis setting.

The following remark connects up two things and explains what the size of $L(q)$ should be and why it might be sharp 3-transitive on X.

Remark: X can be identified with the 1-dimensional subspaces of F_q^2 (F_q^2 is the 2-dimensional vector space over F_q) where we identify $x \in F_q$ with the vector space spanned by $(x, 1)$ and ∞ is matched with the vector subspace spanned by $(1, 0)$. (One can make a mental check that the number of such 1-d subspaces is $q + 1$ as it should be.) Also, the action of $L(q)$ on X can be checked (do it) to correspond to the action of ordinary matrix multiplication on the 1-d subspaces. The latter action has a kernel, the scalar matrices, but when we mod out by them, we get a faithful action of the projective general linear group $PGL(2, q)$ (this group will be reviewed in Section 7.1) on the 1-d subspaces. This basically identifies $L(q)$ with $PGL(2, q)$, both subgroups of S_{q+1}. $|PGL(2, q)| = (q + 1)q(q - 1)$ (see again Section 7.1) which tells us that $L(q)$ has size $(q + 1)q(q - 1)$ which while it does not prove it, it is consistent with $L(q)$ being a sharp 3-transitive subgroup of S_X.

We let $H(g) := L(q)_\infty$, the stabilizer of ∞. Clearly this consists of the mappings

$$\{x \to ax + b\}$$

where a, b belong to F_q and $a \neq 0$.

The next theorem gives us our first infinite family of sharply 3-transitive permutation groups.

Theorem 4.1. *1. $H(q)$ is sharply 2-transitive on F_q (and hence has order $q(q-1)$).*
2. $L(q)$ is sharply 3-transitive on X (and hence has order $(q+1)q(q-1)$).

Proof.
1. To show sharp 2-transitivity of $H(q)$, we need to show that given (x, y) and (x', y') in $F_q \times F_q$ with $x \neq y$ and $x' \neq y'$, there exists unique a and b with $a \neq 0$ satisfying

$$ax + b = x', \quad ay + b = y'.$$

A unique solution to this system follows immediately from linear algebra if $x \neq y$ and a will be nonzero if and only if $x' \neq y'$.
2. The second part follows from an application of the second part of Lemma 2.5.
QED

Remark: (Analogous things will come up elsewhere). In view of the remark above, we have essentially shown that $GL(2, F)$ is 3-transitive on the collection of 1-dimensional spaces of F^2 for any field F. It is natural to ask if for $n \geq 3$, $GL(n, F)$ is 3-transitive on the collection of 1-dimensional spaces of F^n for a given field F. The answer is no. Unlike in 2-dimensions where any two sets of three 1-dimensional spaces are essentially "the same" (this is basically what 3-transitivity means) this is not true for $n \geq 3$. For example, for such n, three 1-dimensional spaces might either (1) span a 2-dimensional space or (2) span a 3-dimensional space. These are different enough so that no matrix can map the first collection to the second and hence one is not 3-transitive.

4.2 Second construction of a family of sharply 3-transitive subgroups

This construction will be obtained by taking the previous construction and "adding a twist". It is not perhaps as natural as the previous one but it is precisely what is needed to lead to the construction of the first two Mathieu groups.

We consider again finite fields with $q = p^{2m}$ elements with p prime but we now restrict ourselves to $p \geq 3$ and m an integer (so that the exponent is even). The reason for the restriction on p is that this will yield that half of the nonzero elements of F_q will be squares. (For fields which have characteristic 2, all the nonzero elements of F_q are squares.) The restriction on the exponent is that it allows us to have a field automorphism which has order 2.

The (field) automorphism group of F_q is isomorphic to $Z/2m$ with generator given by $x \rightarrow x^p$ which is called the Frobenius map. We let σ be the mth power of the generator so that $\sigma(x) = x^{p^m}$; this is our involution. Again we take $X = F_q \cup \{\infty\}$ so $|X| = q + 1$. We now consider the subgroup $M(q)$ of S_X consisting of (1) the permutations

$$\{x \rightarrow \frac{ax+b}{cx+d}\}$$

where a, b, c, d belong to F_q and $ad - bc \neq 0$ is a square in F_q^* and (2) the permutations

$$\{x \rightarrow \frac{a\sigma(x)+b}{c\sigma(x)+d}\}$$

where a, b, c, d belong to F_q and $ad - bc \neq 0$ is a nonsquare in F_q^*. (As before, these maps all take ∞ to $\frac{a}{c}$ and takes $\frac{-d}{c}$ to ∞ for the first class and $\sigma(\frac{-d}{c})$ to ∞ for the second class.)

Exercise: Verify that $M(q)$ is a subgroup S_X.

We let $S(q) := M(q)_\infty$, the stabilizer of ∞. Clearly this consists of (1) the set of mappings

$$\{x \rightarrow ax + b\}$$

where a, b belong to F_q, $a \neq 0$ and a is a square and (2) the set of mappings

$$\{x \rightarrow a\sigma(x) + b\}$$

where a, b belong to F_q, $a \neq 0$ and a is a nonsquare.

The next theorem gives us our second infinite family of sharply 3-transitive permutation groups.

Theorem 4.2. *1. $S(q)$ is sharply 2-transitive on F_q (and hence has order $q(q-1)$).*
2. $M(q)$ is sharply 3-transitive on X (and hence has order $(q+1)q(q-1)$).

Proof.
1. To show sharp 2-transitivity of $S(q)$, it is enough to show that for each a, b in F_q with $a \neq 0$, there is a unique element of $S(q)$ taking 0 to b and 1 to $a + b$. The mappings doing this are $x \to ax + b$ and $x \to a\sigma(x) + b$ and exactly one of these will be in $S(q)$.
2. The second part follows, as in the first construction, from an application of the second part of Lemma 2.5.
QED

4.3 $L(q) \ncong M(q)$

These two groups are not isomorphic when these two groups are both defined, i.e. when q is an even power of an odd prime. We will prove this only for the smallest possible q, namely for $q = 9$ which is the case which we use for the construction of two of our Mathieu groups.

Proposition 4.3. $L(9) \ncong M(9)$

Proof. We have that each of these groups has size $10 \cdot 9 \cdot 8 = 720 = 5 \cdot 3^2 \cdot 2^4$. It suffices to show that their respective Sylow 2-subgroups are not isomorphic. We will show (1) the Sylow 2-subgroup of $L(9)$ is the dihedral group D_{16} with 16 elements and (2) $M(9)$ contains a copy of Q_8, the quaternions. Since (i) D_{16} does not contain Q_8 as a subgroup, (ii) all p-groups are contained in a Sylow p-subgroup and (iii) all Sylow p-subgroups are conjugate and hence isomorphic, the Sylow 2-subgroup of $M(9)$ cannot be D_{16} concluding the proof.

(1). Consider the 2 point stabilizer $L(9)_{0,\infty}$ of $L(9)$ which corresponds to maps $x \to ax$ with $a \in F_9^*$. This subgroup is clearly isomorphic to F_9^* and hence is cyclic and $\cong Z/8$. Let $t \in L(9)$ be the map $x \to 1/x$. Clearly t is an involution and a trivial computation shows that for $a \in F_9^*$,

$$tat^{-1} = a^{-1}$$

where we identified $L(9)_{0,\infty}$ with F_9^*. Hence $\langle L(9)_{0,\infty}, t\rangle = D_{16}$.

(2). Consider the 2 point stabilizer $M(9)_{0,\infty}$ of $M(9)$ which corresponds to maps $x \to ax$ with $a \in F_9^*$ a square and maps $x \to ax^3$ with $a \in F_9^*$ a nonsquare. Clearly this group has 8 elements. To show $M(9)_{0,\infty}$ is Q_8, we first check that it is nonabelian by taking a square a_1 with $a_1^2 \neq 1$, a nonsquare a_2 and checking with an easy computation that $x \to a_1x$ and $x \to a_2x^3$ do not commute. Hence we know $M(9)_{0,\infty}$ is either D_8 or Q_8 since these are the only nonabelian groups on 8 elements. We count the number of involutions. One can check that D_8 has five involutions while Q_8 has only one. Looking now at $M(9)_{0,\infty}$, if $x \to ax$ is an involution, we must have $a^2 = 1$ which gives us an involution in $M(9)_{0,\infty}$ since -1 can be checked to be a square in F_9^* since the latter is cyclic and 4 divides $|F_9^*|$. On the other hand, one checks that $x \to ax^3$ is an involution if and only if $a^4 = 1$ but, as can be checked, there is no nonsquare a satisfying $a^4 = 1$ (since the nonsquares in $Z/8$ all generate). So $M(9)_{0,\infty}$ has one involution and hence is Q_8.

In case that one wants to make sure that (i) D_{16} does not contain Q_8, one can note that an index 2 subgroup of D_{16} is either $Z/8$ or contains at least half involutions and so cannot be Q_8.
QED

5 Transitive Extensions: A key method in the construction of the 5 Mathieu simple sporadic groups

There is a natural and easy way to go from subgroups of S_{n+1} to subgroups of S_n; namely, if $G \subseteq S_{n+1}$, then $G_{n+1} \subseteq S_n$. (Of course, as usual, G_{n+1} is the stabilizer of $n+1$.) In addition, if G is k-transitive, then G_{n+1} will be $(k-1)$-transitive. It is also trivial to go the other way if one does not make any further requirements. (Why?) However the following is an interesting question.

Question: If $H \subseteq S_n$ is a transitive subgroup, does there exist $G \subseteq S_{n+1}$ which is transitive and satisfies $G_{n+1} = H$? Such a G, if it exists, is called a *transitive extension of H*.

Exercise:
Convince yourself that H *not* having a transitive extension would mean that if we view H as sitting in S_{n+1} (doing nothing to $n+1$) and take any permutation $\sigma \in S_{n+1}$ which does not fix $n+1$, then the stabilizer of $n+1$ for the group $\langle H, \{\sigma\}\rangle$ is strictly larger than H.

Example where no transitive extension exists:
Let $n = 4$ and let $H \subseteq S_4$ be the unique Klein 4-group $(Z/2 \times Z/2)$ in S_4 which has no fixed points; this is 1 and the three permutations consisting of two transpositions. Note H is transitive. We claim that there is no transitive $G \subseteq S_5$ with $G_5 = H$. The proof is as follows. If such a transitive G existed, it would necessarily have $5 \cdot 4 = 20$ elements. The Sylow theorems immediately imply that there is a unique Sylow 5-subgroup P which is normal in G. P is generated by a 5-cycle σ which WLOG can be taken to be (12345). Since $P \lhd G$, we have

$$(12)(34)\sigma(12)(34) = (21435)$$

must be in P. But one quickly checks that (21435) is not σ^i for any i giving a contradiction. QED

Question: Let $n = 4$ and let $H \subseteq S_4$ be a 4-cycle. Does there exist a transitive extension?

The following theorem due to Witt gives us a condition under which we can extend. It looks a little strange but it does the trick.

Theorem 5.1. *Let $G \subseteq S_n$ be 2-transitive. Assume there exists $x \in \{1, 2, \ldots, n\}$, $g \in G$ and $h \in S_{n+1}$ so that*

1. $g(x) \neq x$
2. $h(n+1) \in \{1, 2, \ldots, n\}$
3. $h^2 \in G$
4. $(gh)^3 \in G$ *and*
5. $hG_x h = G_x$.
Then $\tilde{G} := \langle G, \{h\} \rangle \subseteq S_{n+1}$ *is a transitive extension of* G*; i.e., the stabilizer of* $\tilde{G}$ *at* $n+1$ *is* G.

Proof. First, note that 2. guarantees that $\tilde{G}$ is transitive. We will show that

$$\tilde{G} = G \cup GhG;$$

this will suffice since then

$$\tilde{G}_{n+1} = G_{n+1} \cup (GhG)_{n+1} = G$$

since no element of GhG can fix $n+1$ so that the second set is empty. To verify this, it is enough to show that $G \cup (GhG)$ is a subgroup of S_{n+1}; i.e. closed under products since we are in a finite group.

We first need the following preliminary claim that

$$G = G_x \cup G_x g G_x \tag{1}$$

which we demonstrate as follows. Assume $g_1 \in G \backslash G_x$ takes x to $z \neq x$. Let $y := g(x) \neq x$ by assumption 1. By 2-transitivity of G which implies transitivity of G_x, choose $h_1 \in G_x$ taking y to z (y and z can be the same) and note then that $h_1 g$ takes x to z. This implies that $(h_1 g)^{-1} g_1$ takes x to x and hence

$$g_1 = h_1 g h_2$$

for some $h_2 \in G_x$, proving the claim.

Now note that

$$(G \cup GhG)(G \cup GhG) \subseteq G \cup GhG \cup GhGhG.$$

We need only worry now about $GhGhG$. To show this is contained in $G \cup GhG$, it suffices to show that $hGh \subseteq G \cup GhG$.

Letting $\gamma := h^2$ and $\delta := (gh)^3$ which are in G by assumption, we get, using (1)

$$hGh = h(G_x \cup G_x g G_x)h = hG_x h \cup hG_x g G_x h = hG_x h \cup hG_x h h^{-1} g h^{-1} h G_x h.$$

By assumption 5., this becomes

$$G_x \cup G_x h^{-1} g h^{-1} G_x. \tag{2}$$

Noting that $h\gamma^{-1} = h^{-1} = \gamma^{-1}h$ and $hgh = g^{-1}h^{-1}g^{-1}\delta$, we have

$$h^{-1}gh^{-1} = \gamma^{-1}hgh\gamma^{-1} = \gamma^{-1}g^{-1}h^{-1}g^{-1}\delta\gamma^{-1} = \gamma^{-1}g^{-1}h\gamma^{-1}g^{-1}\delta\gamma^{-1}.$$

It follows that (2) is contained in $G \cup GhG$, completing the proof.
QED

6 Construction of the first two Mathieu groups using transitive extensions starting from the sharp 3-transitive subgroup $M(9)$ of S_{10}

In this section, we will construct the two smallest Mathieu groups by applying Theorem 5.1 twice starting from the sharp 3-transitive subgroup $M(9)$ of S_{10} given in Theorem 4.2. For obvious reasons, one often denotes $M(9)$ by M_{10} which we do here.

6.1 Construction of M_{11}

Theorem 6.1. *There exists a sharp 4-transitive subgroup M_{11} of S_{11} which is a transitive extension of M_{10}. (The sharp 4-transitivity implies that $|M_{11}| = 11 \cdot 10 \cdot 9 \cdot 8 = 7920$.)*

Proof.
We will apply Theorem 5.1 with $G = M_{10}$ which is 2-transitive (being sharp 3-transitive) on the set $F_9 \cup \{\infty\}$. We will call our added 11th element ω. F_9 can be taken to be $\frac{Z}{3}[x]/(x^2 + x - 1)$ where each element can be written as $a + b\pi$ where a and b are in F_3 and π is a variable satisfying $\pi^2 + \pi = 1$ and that π is a primitive element.

We now let $x = \infty$,

$$g = (0\ \infty)(\pi\ \pi^7)(\pi^2\ \pi^6)(\pi^3\ \pi^5)(1)(\pi^4)$$

and

$$h = (\omega\ \infty)(\pi\ \pi^2)(\pi^3\ \pi^7)(\pi^5\ \pi^6)(1)(\pi^4)(0).$$

We need to check all the five conditions but before that we need to remember to check that $g \in G$. In terms of a mapping, one quickly checks that g corresponds to the map $x \to 1/x$ which gives $a, d = 0$ and $b, c = 1$ which yields a determinant of -1. However -1 is the square π^4 and hence this map belongs to G. Now we check the five conditions; all will be straightforward except the last.

1. and 2. are immediate. 3. is clear since $h^2 = 1$ being a product of transpositions. For 4., first compute gh getting

$$gh = (\omega\ 0\ \infty)(\pi\ \pi^6\ \pi^3)(\pi^2\ \pi^7\ \pi^5)(1)(\pi^4)$$

implying that $(gh)^3 = 1$.

For 5., we first give an analytic expression for h. One checks that

$$h = (\omega\,\infty)(x \to \pi^2 x + \pi x^3).$$

(I don't see how one checks this without taking out a piece of paper and doing the necessary arithmetic in F_9.)

Since h is an involution, we need only show that $h(M_{10})_\infty h \subseteq (M_{10})_\infty$. Now if $f \in (M_{10})_\infty$ (or even in fact just in M_{10}), one easily checks that $hfh(\infty) = \infty$ and so we need to show that $hfh \in M_{10}$.

If $f \in (M_{10})_\infty$, then

$$f(x) = \pi^{2i}x + \alpha \text{ or } \pi^{2i+1}x^3 + \alpha$$

for some $i \in \{0, \dots, 3\}$ and $\alpha \in F_9$. We deal only with the first case only; the second can be handled in an analogous way. We use the second description of h above and note that, since we are in characteristic 3, we always have $(x+y)^3 = x^3 + y^3$. A straightforward computation then gives

$$hfh(x) = (\pi^{2i+4} + \pi^{6i+4})x + (\pi^{2i+3} + \pi^{6i+7})x^3 + \pi^2\alpha + \pi\alpha^3.$$

We need to consider whether i is even or odd.

If $i = 2j$, the coefficient of x^3 becomes (recall π has order 8)

$$\pi^{4j+3} + \pi^{12j+7} = \pi^{4j+3}(1 + \pi^4) = 0$$

while the coefficient of x becomes

$$\pi^{4j+4} + \pi^{12j+4} = 2\pi^{2(2j+2)} = -1\pi^{2(2j+2)}$$

which is a square and hence $hfh \in M_{10}$ in this case.

If $i = 2j + 1$, the coefficient of x becomes

$$\pi^{4j+6} + \pi^{12j+10} = \pi^{4j+6}(1 + \pi^4) = 0$$

while the coefficient of x^3 becomes

$$\pi^{4j+5} + \pi^{12j+13} = (-1)\pi^{4j}\pi^5$$

which is a nonsquare and hence $hfh \in M_{10}$ in this case as well. This completes the proof of the transitive extension.

Finally the sharp 3-transitivity of M_{10} allows us to conclude that $M_{11} := \langle M_{10}, \{h\}\rangle$ is sharp 4-transitive using Lemma 2.5, as desired.
QED

6.2 Construction of M_{12}

We now construct our second smallest sporadic simple group M_{12} using M_{11}.

Theorem 6.2. *There exists a sharp* 5*-transitive subgroup* M_{12} *of* S_{12} *which is a transitive extension of* M_{11}*. (The sharp* 5*-transitivity implies that* $|M_{12}| = 12 \cdot 11 \cdot 10 \cdot 9 \cdot 8 = 95,040$*.)*

Proof.
We will apply Theorem 5.1 with $G = M_{11}$ which is sharp 4-transitive on the set $F_9 \cup \{\infty\} \cup \{\omega\}$ by Theorem 6.1. We will call our added 12th element Ω so that our 12 element set is $F_9 \cup \{\infty\} \cup \{\omega\} \cup \{\Omega\}$.

We let $x = \omega$ and let

$$h := (\omega\ \infty)(\pi\ \pi^2)(\pi^3\ \pi^7)(\pi^5\ \pi^6)(1)(\pi^4)(0).$$

Note that this h is the same h as in the previous proof but it will now play the role of g in the statement of Theorem 5.1. Note that since $h \in M_{11}$ by construction, we will not need to do that prestep as we did in the previous proof where we verified that g actually belonged to M_{10}. We next let

$$k := (\omega\ \Omega)(\pi\ \pi^3)(\pi^2\ \pi^6)(\pi^5\ \pi^7)(1)(\pi^4)(0)(\infty).$$

This k will in turn play the role of h in the statement of Theorem 5.1. We now look at 1-5 in the latter theorem keeping in mind that the g and h there are h and k here. 1-3 are immediate and 4. follows when one checks that

$$hk = (\omega\ \Omega\ \infty)(\pi\ \pi^7\ \pi^6)(\pi^2\ \pi^5\ \pi^3)(1)(\pi^4)(0).$$

For 5., one first notes that k has the following analytic expression which is immediate:

$$k = (\omega\ \Omega)(x \to x^3).$$

Since k is an involution, we need only show that $k(M_{11})_\omega k \subseteq (M_{11})_\omega$. Now if $f \in (M_{11})_\omega$ (or even in fact just in M_{11}), one easily checks that $kfk(\omega) = \omega$ and so we need to show that $kfk \in M_{11}$.

Now $f \in (M_{11})_\omega$ means (essentially) that $f \in M_{10}$ and hence either (1)

$$f(x) = \frac{ax+b}{cx+d}$$

with $ad - bc$ a square in F_9^* or (2)

$$f(x) = \frac{ax^3+b}{cx^3+d}$$

with $ad - bc$ a nonsquare in F_9^*.

Assuming the first case and recalling that on $F_9 \cup \{\infty\}$, $k(x) = x^3$ and that we are in characteristic 3, we get

$$kfk(x) = (\frac{ax^3 + b}{cx^3 + d})^3 = \frac{a^3x^9 + b^3}{c^3x^9 + d^3} = \frac{a^3x + b^3}{c^3x + d^3}.$$

For this to be in M_{10}, we need the determinant of this, $a^3d^3 - b^3c^3$, to be a square. However, this equals $(ad - bc)^3$ which is a square since $ad - bc$ is a square.

For the second case, we obtain, in a completely analogous way,

$$kfk(x) = \frac{a^3x^3 + b^3}{c^3x^3 + d^3}.$$

Now, for this to be in M_{10}, we need the determinant of this $a^3d^3 - b^3c^3$ to be a nonsquare. However, we have that $ad - bc$ is a nonsquare which easily implies that the equivalent $(ad - bc)^3$ is a nonsquare. (Why?) This proves we have a transitive extension.

Finally the sharp 4-transitivity of M_{11} allows us to conclude that $M_{12} := \langle M_{11}, \{k\}\rangle$ is sharp 5-transitive using Lemma 2.5, as desired.
QED

Questions for thought (which I have not thought at all about):
1. It is a consequence of Jordan's Theorem that there is no transitive extension for M_{12}. Can one see that directly?
2. To get M_{11}, we needed to do a transitive extension of $M(9)$ rather than a transitive extension of the more natural group $L(9)$. Can one also do a transitive extension of $L(9)$ and what does one get if one can do that?

7 Construction of the last three Mathieu groups using transitive extensions starting from the 2-transitive subgroup $PSL(3,4)$ of S_{21}

7.1 Linear groups

Subsection 7.1 contains background and motivation that some readers might not need. Those who understand and know the following statements can move on to subsection 7.2.

$PSL(n,q)$ acts faithfully on n-dimensional projective space over the field F_q and hence can be viewed as a subgroup of the symmetric group on n-dimensional projective space, a set containing $\frac{q^n-1}{q-1}$ elements; moreover $PSL(n,q)$ is a simple group except when (n,q) is $(2,2)$ *or* $(2,3)$*. In particular, $PSL(3,4)$ is a subgroup of S_{21}.*

7.1.1 The definition of the Linear groups

We need to quickly review some basic definitions and results about linear groups since $PSL(3,4)$ (defined below) will be the starting point for the inductive construction of the last three Mathieu simple sporadic groups. Fix a field F with $q = p^k$ elements with p prime. Fix an integer $n \geq 2$.

Definition 7.1. *1. $GL(n,F)$ is the group (under multiplication) of $n \times n$ matrices over F*
2. $SL(n,F)$ is the subgroup of $GL(n,F)$ consisting of determinant 1 matrices
3. $PGL(n,F)$ is the quotient group of $GL(n,F)$ when we mod out by its center (which can be shown to be the scalar matrices)
and
4. $PSL(n,F)$ is the quotient group of $SL(n,F)$ when we mod out by its center (which can be shown to be the scalar matrices within $SL(n,F)$).

The names of these groups are respectively (1) the general linear group, (2) the special linear group, (3) the projective general linear group and (4) the projective special linear group.

Remarks: It it not immediately obvious perhaps that the centers of these groups are as claimed but it is not a hard exercise. (Once you know the center of $GL(n,F)$ is what it is claimed to be, there is no general group theory fact which gives you something analogous for $SL(n,F)$. This is because if $H \subseteq G$, while one does have $C(G) \cap H \subseteq C(H)$, one does not have equality in general.)

How many elements are there in these groups?
Exercises:
1. $|GL(n,F)| = (q^n - 1)(q^n - q)(q^n - q^2)\ldots(q^n - q^{n-1})$. (Hint: Think of bijective linear transformations.)
2. $|SL(n,F)| = |GL(n,F)|/(q-1)$. (Hint: $SL(n,F)$ is the kernel of the determinant mapping.)
3. $|PGL(n,F)| = |GL(n,F)|/(q-1)$. (Hint: The number of elements in the center is $q-1$.)
4. $|PSL(n,F)| = |SL(n,F)|/(n,q-1)$ where $(n,q-1)$ is the gcd of n and $q-1$. (Hint: This is probably the most interesting of the above formulas since it depends on both n and q "together". One needs to recall some things about ordinary cyclic groups.)

One of the main (and very old) theorems about the projective special linear groups is the following.

Theorem 7.2. *$PSL(n,F_q)$ is a simple group except when (n,q) is $(2,2)$ or $(2,3)$. (For the results later on concerning the simplicity of the last three Mathieu groups, we will just need that $PSL(3,4)$ is simple.)*

The exceptions are easy to check. First one easily checks that $PSL(2,2)$ and $PSL(2,3)$ have cardinalities 6 and 12 respectively and one can check easily that no group of these orders are simple (using say the Sylow Theorems for 12). In fact, these groups are isomorphic to S_3 and A_4 respectively but one does not need that to conclude that they are not simple. In fact, from the next subsubsection 7.1.2, these last two facts easily follow.

Since Theorem 7.2 is not needed for the construction of the Mathieu groups (although it is used in these notes to prove their simplicity), we delay the proof of this theorem until Subsection 8.3

A few interesting factoids to keep in mind: Sometimes these projective special linear groups correspond to alternating groups, which is our other large class of simple groups. First, we have $PSL(2,4) \cong PSL(2,5) \cong A_5$. The first one of this family of groups which is not an alternating group is $PSL(2,7)$ with 168 elements. Also, $PSL(2,7) \cong PSL(3,2)$, $PSL(2,9) \cong A_6$ and $PSL(4,2) \cong A_8$. But these are the only ones (I think). One last interesting thing is that $PSL(3,4) \ncong A_8$ although they have the same order. The order of these two groups is 20,160 which is the smallest number to have two nonisomorphic simple groups. They can be distinguished since A_8 contains an element of order 15, namely (12345)(678), while it can be shown that $PSL(3,4)$ contains no element of order 15. Alternatively, A_8 contains an element of order 6, namely (123)(45)(67), while it can be shown that $PSL(3,4)$ contains no element of order 6.

7.1.2 Projective linear groups as permutation groups

It might at first seem arbitrary to define the projective groups as we did. The definitions might perhaps feel more motivated from a "group actions" point of view as we now explain. We also have an important theorem which allows us to view $PSL(n, q)$ as a 2-transitive permutation group.

We begin however our discussion with $GL(n, q)$ and its corresponding quotient $PGL(n, q)$. Clearly $GL(n, q)$ acts on F_q^n in a natural and obvious way and it is immediate that this action is faithful (essentially by definition). A problem however with this action is that it is not primitive (except when $q = 2$) since any 1-dimensional space with zero removed is a block (as defined earlier in these notes) for this action. In particular, when $q \neq 2$, this action cannot be 2-transitive on the nonzero vectors, something which can be directly checked anyway. It then seems natural in order to remove this problem to identify the nonzero vectors in a block, i.e., identify nonzero vectors which sit inside a 1-d subspace. We therefore consider the collection of 1-dimensional subspaces which is called projective space and which we denote by $P(n, q)$. (Note that we have thrown away zero.) Since we identified blocks, we obtain a well-defined and natural induced action of $GL(n, q)$ on $P(n, q)$. However, this action is often no longer faithful and it then seems natural to mod out by the kernel of this action in order to get a faithful action and thereby actually obtain a subgroup of $S_{P(n,q)}$. It is clear that the center consisting of scalar matrices of $GL(n, q)$ is certainly a subset of the kernel of this action. The key fact is that the kernel is in fact no larger than the center and hence $PGL(n, q)$ acts faithfully on $P(n, q)$ and hence becomes a subgroup of $S_{P(n,q)}$.

Everything stated above is true for $SL(n, q)$ and $PSL(n, q)$ and so we can view $PSL(n, q)$ as a subgroup of $S_{P(n,q)}$. In particular, the following are equivalent: (1) the center of $SL(n, q)$ is trivial, (2) the action of $SL(n, q)$ on $P(n, q)$ is faithful and (3) $SL(n, q) = PSL(n, q)$. The only difference is that for $GL(n, q)$, the analogues of the above hold if and only if $q = 2$, while for $SL(n, q)$, these occur if and only if $(n, q) = 1$.

The next theorem says some of this. We take as a given that the centers of these two matrix groups $GL(n, q)$ and $SL(n, q)$ are the scalar matrices within each one.

Theorem 7.3. *1. The action of $SL(n, q)$ on $P(n, q)$ is 2-transitive.*
2. The kernel of this action is the center of $SL(n, q)$ (which implies $PSL(n, q)$ is a permutation group).

Proof.

1. Let (v_1, v_2) and (w_1, w_2) be two pairs of linearly independent vectors. Extend them to bases $(v_1, v_2, \ldots, v_n)$ and $(w_1, w_2, \ldots, w_n)$ and let M be an invertible map sending v_i to w_i for each i. We would be done except M might not be in $SL(n, q)$. In that case, let a be the determinant of M and let T be the map taking w_1 to $\frac{w_1}{a}$ and w_i to w_i for $i \geq 2$. Then $T \circ M$ is the desired map in $SL(n, q)$ taking $\langle v_i \rangle$ to $\langle w_i \rangle$ for $i = 1, 2$ (and in fact doing this for all i).
2. Clearly the center (or scalar matrices) is contained in the kernel of the action. Now let $M \in SL(n, q)$ be in the kernel of the action and we want to show M is a scalar matrix. Now, for all nonzero vectors v, there is a nonzero a_v so that $Mv = a_v v$. Let $v_1, \ldots, v_n$ be a basis. To show M is a scalar matrix, it suffices to show all the a_{v_i}'s are the same and it then suffices to show that $a_{v_1} = a_{v_2}$. We have

$$a_{v_1+v_2}(v_1 + v_2) = M(v_1 + v_2) = a_{v_1}v_1 + a_{v_2}v_2.$$

Linear independence of v_1 and v_2 now gives $a_{v_1} = a_{v_2}$.
QED

There are a few examples where the 2-transitive subgroup $PSL(n, q)$ of $S_{P(n,q)}$ is sharply 2-transitive such as $PSL(2, 2)$ and $PSL(2, 3)$ (check this!) but this will almost never be the case. Perhaps it is only in these 2 cases; I haven't really checked this.

7.2 Construction of M_{22}

Theorem 7.4. *There exists a 3-transitive subgroup $M_{22} \subseteq S_{22}$ with 443,520 elements which is a transitive extension of $PSL(3, 4) \subseteq S_{21}$.*

Proof.
We will apply Theorem 5.1 with $G := PSL(3, 4)$ which is a 2-transitive subgroup of S_{21} where the 21 element set is represented by $\{[x, y, z]\}$ where $[x, y, z]$ denotes the 1-dimensional vector space spanned by a nonzero (x, y, z) in F_4^3. (The h in that theorem will be called h_1 here.) Denote this latter set by X. We will call our added 22nd element ∞, so our 22 elements are $X \cup \{\infty\}$.

We let $x = [1, 0, 0]$, g be defined by $g([x, y, z]) = [y, x, z]$ and

$$h_1 := ([1, 0, 0]\ \infty) f_1$$

where

$$f_1([u, v, w]) = [u^2 + vw, v^2, w^2].$$

Note first that $g \in PSL(3,4)$ since the determinant of the corresponding matrix is $-1 = 1$. Note also f_1 is well-defined on projective space since all the components are 2-homogeneous. Note next that $[1,0,0]$ is fixed by f_1 and that f_1 is an involution. The latter is a trivial computation left to the reader which uses the fact we are in the field F_4 which has characteristic 2 and the cube of any nonzero element is 1. The first three conditions in Theorem 5.1 are now immediate while the fourth condition is another fairly easy computation (when you keep your tongue straight in your mouth) left to the reader where again all the computations simplify since we are in F_4.

For the last condition, $h_1 G_{[1,0,0]} h_1 = G_{[1,0,0]}$, since h_1 is an involution, we just need that $h_1 G_{[1,0,0]} h_1 \subseteq G_{[1,0,0]}$. Now if $k \in G_{[1,0,0]}$ (or even in fact just in G), one easily checks that $h_1 k h_1([1,0,0]) = [1,0,0]$ and so we need to show that $h_1 k h_1 \in G$. If $k \in G_{[1,0,0]}$, it can be represented by a matrix

$$\begin{bmatrix} 1 & \star & \star \\ 0 & a & b \\ 0 & c & d \end{bmatrix}$$

with $ad - bc = 1$. Another computation (not hard but takes a little more time) *again* left to the reader shows that $h_1 k h_1$ is represented by the matrix

$$\begin{bmatrix} 1 & \star & \star \\ 0 & a^2 & b^2 \\ 0 & c^2 & d^2 \end{bmatrix}$$

The determinant of this is given by $a^2d^2 - b^2c^2 = (ad - bc)^2 = 1$ since we are in characteristic 2. Since this is then in G, we conclude that this yields a transitive extension for us.

The 2-transitivity of $PSL(3,4)$ as a subgroup of S_{21} allows us to conclude that $M_{22} := \langle PSL(3,4), \{h_1\}\rangle$ is a 3-transitive subgroup of S_{22} using Lemma 2.5. Finally, since $(M_{22})_\infty = PSL(3,4)$, we have $|M_{22}| = 22 \cdot |PSL(3,4)| = 22 \cdot 20,160 = 443,520$.
QED

7.3 Construction of M_{23}

Theorem 7.5. *There exists a 4-transitive subgroup $M_{23} \subseteq S_{23}$ with 10,200,960 elements which is a transitive extension of $M_{22} \subseteq S_{22}$.*

Proof.
We will again apply Theorem 5.1 with $G := M_{22}$ which is 3-transitive on

$X \cup \{\infty\}$ by Theorem 7.4 with the definition of X given there. We will now add a 23rd element ω, so our 23 elements are $X \cup \{\infty\} \cup \{\omega\}$. Let $x = \infty$, $g = h_1$ from the previous proof and

$$h_2 := (\omega\ \infty) f_2$$

where

$$f_2([u, v, w]) = [u^2, v^2, \beta w^2]$$

where β is a primitive root in F_4 (i.e., any nonzero element other than 1).

One notes that $[1, 0, 0]$ is fixed by f_2 and that f_2 is an involution. (h_2 will play the role of h in Theorem 5.1.) The first three conditions in Theorem 5.1 are now immediate while the fourth condition is another (not so hard) computation left to the reader where you follow your nose and where again all the computations simplify since we are in F_4.

For 5., since h_2 is an involution, we need only show that $h_2(M_{22})_\infty h_2 \subseteq (M_{22})_\infty$. Now if $f \in (M_{22})_\infty$ (or even in fact just in M_{22}), one easily checks that $h_2 f h_2(\infty) = \infty$ and so we need to show that $h_2 f h_2 \in M_{22}$. Note that $f \in (M_{22})_\infty$ means that $f \in PSL(3, 4)$ and another (fairly straightforward) computation left to the reader shows that $h_2 f h_2$ is a linear mapping even though h_2 is not. Moreover, the resulting matrix has the original matrix elements of f squared and in addition each term is multiplied either by β, β^2 or 1. Moreover these extra factors of $1, \beta$ or β^2 are such that when computing the determinant, each product corresponding to the 6 permutations is β^3 times the original product for f squared. Since $\beta^3 = 1$ and we are in characteristic 2, the new determinant is still one (why?) and hence belongs to $PSL(3, 4) = (M_{22})_\infty$. So 5. holds and we have a transitive extension of (M_{22}).

The 3-transitivity of M_{22} allows us to conclude that $M_{23} := \langle M_{22}, \{h_2\}\rangle$ is 4-transitive using Lemma 2.5. Finally, since $(M_{23})_\omega = M_{22}$, we get $|M_{23}| = 23 \cdot |M_{22}| = 23 \cdot 443,520 = 10,200,960$.
QED

7.4 Construction of M_{24}

Theorem 7.6. *There exists a* 5*-transitive subgroup* $M_{24} \subseteq S_{24}$ *with 244,823,040 elements which is a transitive extension of* $M_{23} \subseteq S_{23}$*.*

Proof.
We will again apply Theorem 5.1 with $G := M_{23}$ which is 4-transitive on $X \cup \{\infty\} \cup \{\omega\}$ by Theorem 7.5 with the definition of X being unchanged.

We will now add a 24th element Ω, so our 24 elements are $X \cup \{\infty\} \cup \{\omega\} \cup \{\Omega\}$. Let $x = \omega$, $g = h_2$ from the previous proof and

$$h_3 := (\omega\ \Omega) f_3$$

where

$$f_3([u, v, w]) = [u^2, v^2, w^2].$$

One notes that $[1, 0, 0]$ is fixed by f_3 and that f_3 is an involution. (h_3 will play the role of h in Theorem 5.1.) The first three conditions in Theorem 5.1 are now immediate while the fourth condition is verified as in the proof of Theorem 7.5.

For 5., since h_3 is an involution, we need only show that $h_3 (M_{23})_\omega h_3 \subseteq (M_{23})_\omega$. Now if $f \in (M_{23})_\omega$ (or even in fact just in M_{23}), one easily checks that $h_3 f h_3(\omega) = \omega$ and so we need to show that $h_3 f h_3 \in M_{23}$.

There is a subtle point here. If $f \in (M_{23})_\omega = M_{22}$ were in fact also in $(M_{22})_\infty = PSL(3, 4)$, then we could argue exactly as in the proof of Theorem 7.5 that $h_3 f h_3 \in PSL(3, 4) \subseteq (M_{23})_\omega$. In fact, in that case, the proof is slightly simpler since the β factors are not present. However, f need not belong to $(M_{22})_\infty$; for example, f might well be h_1 from the proof of Theorem 7.4 which is an element of $M_{22} = (M_{23})_\omega$. One can in that case compute that

$$h_3 h_1 h_3 = h_1 \in M_{22} \subseteq M_{23}.$$

Since $(M_{23})_\omega = M_{22}$ is generated by $PSL(3, 4)$ and h_1, we can conclude that $h_3 f h_3 \in M_{23}$ for all $f \in M_{22} = (M_{23})_\omega$, as desired.

The 4-transitivity of M_{23} allows us to conclude that $M_{24} := \langle M_{23}, \{h_3\} \rangle$ is 5-transitive using Lemma 2.5. Finally, since $(M_{24})_\Omega = M_{23}$, we get $|M_{24}| = 24 \cdot |M_{23}| = 24 \cdot 10,200,960 = 244,823,040$.
QED

8 Simplicity of the 5 Mathieu groups

Some of the proofs of simplicity for the Mathieu groups rely on somewhat advanced topics (such as Burnsides so-called p-complement theorem) which certainly would not be covered in a basic group theory course. In 1995, Robin Chapman found a proof that M_{11} and M_{23} are simple using only "standard" background and moreover the proofs only use the cardinality of the two groups and the fact that they are transitive permutation groups. We give this argument, which is a crucial step in our proof of simplicity of all of the Mathieu groups, in a more general form in Subsection 8.2.

8.1 Some needed lemmas

It will be convenient to collect in this subsection various general results about permutation groups that will be needed.

Lemma 8.1. *Let G be a 2-transitive subgroup of S_n and let N satisfy*

$$1 \lneq N \lhd G.$$

Then N is transitive.

Proof.
Since N is not trivial, there must exist $h \in N$ and $a \in \{1, \ldots, n\}$ so that $h(a) \neq a$. Let $x \neq y$ be arbitrary. We want to find an element in N taking x to y. Since G is 2-transitive, there is $g \in G$ so that $g(a) = x$ and $g(h(a)) = y$. Now ghg^{-1} takes x to y but belongs to N since N is normal.
QED

Remark: The conclusion of this lemma is still true under the weaker assumption that G is primitive. The argument in that case is a little bit longer.

Lemma 8.2. *Let G be a 2-transitive subgroup of S_n. Then each stabilizer G_x is a maximal subgroup.*

Proof.
Assume $G_x \lneq H \lneq G$ and we want to get a contradiction. Choose $h \in H \backslash G_x$ and $g \in G \backslash H$. $h(x) \neq x \neq g(x)$ and so by 2-transitivity, there is $g_0 \in G$ so that g_0 takes x to x and $h(x)$ to $g(x)$. Then $\{g_0, g^{-1}g_0h\} \subseteq G_x \subseteq H$. Since $h \in H$ also, we get $g \in H$, a contradiction. QED

Remark: The conclusion of this lemma is also true under the weaker assumption that G is primitive and is in fact equivalent to primitivity.

Lemma 8.3. *If $H \subseteq G \subseteq S_X$ and H is transitive, then for any $x \in X$,*

$$G = HG_x.$$

Proof.
Fix $g \in G$ and let $y = gx$. Since H is transitive, there is $h \in H$ such that $x = hy$. Then $hg \in G_x$ and hence $g \in HG_x$. QED

8.2 Two "easy" cases of simplicity: M_{11} and M_{23}

Theorem 8.4. *Let G be a 2-transitive subgroup of S_p where p is prime. Assume that*

$$|G| = pmr$$

where $m \equiv 1(p)$, $m > 1$ and $r < p$ is also prime. Then G is simple.

Before giving the proof of this result, let's first apply it to M_{11} and M_{23}.

Corollary 8.5. *M_{11} and M_{23} are simple.*

Proof (of the corollary).
First, $|M_{11}| = 7920 = 11 \cdot 3^2 \cdot 2^4 \cdot 5$. Letting $p = 11$ (as it has to be), $m = 144$ and $r = 5$, one sees the conditions of the theorem are met.

Secondly, $|M_{23}| = 10,200,960 = 23 \cdot 7 \cdot 5 \cdot 3^2 \cdot 2^7 \cdot 11$. Letting $p = 23$ (as it has to be), $m = 40,320$ and $r = 11$, one checks the conditions of the theorem are met.
QED

To prove Theorem 8.4, we begin with a few lemmas. The first lemma is straightforward and left to the reader.

Lemma 8.6. *If G is a transitive subgroup of S_p with p prime, then $p||G|$, $p^2 \not| |G|$ and any Sylow p-subgroup (which then has to have size p) is generated by a p-cycle.*

The next lemma is relatively standard but we include a proof.

Lemma 8.7. *If P is a Sylow p-subgroup of S_p with p prime, then $|N_{S_p}(P)| = p(p-1)$.*
(The normalizer is in fact a semi-direct product $Z/p \times_\phi Z/(p-1)$ but this is not needed. Note that with the exception of $p = 3$, P cannot be normal since it's normalizer is too small.)

Proof.
WLOG P is generated by $(1, 2, \ldots, p)$. The number of conjugates of this element is just the number of p-cycles, which is $(p-1)!$. Since p is prime, the subgroups conjugate to P (which are all the subgroups of this size by Sylow) each contain $p-1$ such p-cycles and these are disjoint, other than 1. We conclude that the number of subgroups conjugate to P is $(p-1)!/(p-1) = (p-2)!$. Since the number of such conjugate subgroups is also $[S_p : N_{S_p}(P)]$ (using the general orbit-stabilizer theorem), we obtain the result. QED

Forgetting for the moment the setup of Theorem 8.4, let us assume now that we have a prime p and a transitive subgroup G of S_p implying that p divides $|G|$. Let P be an arbitrary Sylow p-subgroup (necessarily of size p) of G. The following diagram will be useful to assist in our thinking. We have

$$1 \subseteq P \subseteq N_G(P) \subseteq \begin{matrix} N_{S_p}(P) \\ G \end{matrix} \subseteq S_p$$

where when two subgroups are on top of each other, it means we don't necessarily have a containment either way.

Using this diagram, we are going to *define* $r_G := [N_G(P) : P]$ and $m_G := [G : N_G(P)]$ which is also the number of Sylow p-subgroups in G. Note that $|G| = p \cdot r_G \cdot m_G$.

Corollary 8.8. $r_G|(p-1)$.

Proof. Referring to the diagram and using Lemma 8.7, we clearly have

$$[N_{S_p}(P) : N_G(P)]r_G = [N_{S_p}(P) : N_G(P)][N_G(P) : P] = [N_{S_p}(P) : P] = p-1.$$

QED

Lemma 8.9. *r_G is the smallest residue of $\frac{|G|}{p}$ (mod p).*

Proof.
The Sylow theorems imply that $m_G \equiv 1\ (p)$ which implies

$$\frac{|G|}{p \cdot r_G} \equiv 1(p)$$

and hence $\frac{|G|}{p} \equiv r_G\ (p)$. Since $r_G|(p-1)$ by the previous corollary, r_G is clearly the smallest such residue.
QED

Lemma 8.10. *If $r_G = 1$, then $m_G = 1$*

Proof.
Since p is prime, the number of elements of order p (i.e. p-cycles) in G is $m_G(p-1)$. If $r_G = 1$, then this is $|G| - m_G$. Hence G has at least $|G| - m_G$ elements with no fixed points and hence at most m_G elements (including 1) with a fixed point.

Fix $x \in \{1, \dots, p\}$ and note that since $[G : G_x] = p$, it follows, if $r_G = 1$, that $|G_x| = m_G$. So each stabilizer G_x gives us m_G elements with a fixed point. Since there are only m_G such elements in total and all stabilizers have the same size, we conclude that the stabilizers must be all equal. This implies they are all trivial (since we are dealing with a permutation group, not a general group action) and hence from the above, we can conclude that $m_G = 1$.
QED

We can now proceed with the

Proof of Theorem 8.4
We have both $|G| = pmr$ and $|G| = pm_Gr_G$ where m_G and r_G are defined above. We first want to show that $r = r_G$ which would then imply that $m = m_G$. We have $1 < m \equiv 1(p)$, $m_G \equiv 1(p)$, $r < p$ and $r_G|(p-1)$. Since $|G|$ is the product of the relevant three terms, this gives

$$(k'p+1)r_G = (kp+1)r$$

for some integers k, k' yielding $r_G \equiv r(p)$. Since both these terms are at most $p-1$, we get $r = r_G$ and so $m = m_G$.

We now want to show G is simple. Assume

$$1 \lneq H \lhd G$$

and so we want to show $H = G$. Lemma 8.1 tells us that H is transitive.

Since H is transitive, we have $p||H|$ and hence the Sylow p-subgroups of H have the same size as those of G, namely p. Since $H \lhd G$, it follows that all of the Sylow p-subgroups of G are contained in H (Why?). We conclude using the same notation as above for H that $m_H = m_G = m$. Now writing, with all the same definitions but with respect to H,

$$|H| = pm_Hr_H$$

and hence $r_H|r$. Since $m > 1$ by assumption, we get $m_H > 1$ and hence $r_H > 1$ by Lemma 8.10 applied to H. Since r is prime, we conclude that $r_H = r$ and hence $|H| = |G|$.
QED

Remark: Theorem 8.4 is true without the 2-transitivity assumption. Since p is prime, we can conclude that G is primitive and we mentioned earlier that Lemma 8.1 is still true under the weaker assumption of primitivity. Lemma 8.1 was the only place where the 2-transitivity assumption was used.

8.3 Simplicity of $PSL(3,4)$

This result is "standard" (not meaning easy) and is in many books. If one is happy just taking this as "known" or "on faith", one can skip this subsection. However for those who want to see the proof, it is "sort of" done here. We repeat the statement.

Theorem 8.11. *If F is a field on q elements, then $PSL(n,F)$ is a simple group except when (n,q) is $(2,2)$ or $(2,3)$. (For the results later on, we will just need that $PSL(3,4)$ is simple.)*

Although not the original way in which simplicity was proven for these groups, some books now use the following theorem which provides an intrinsic sufficient condition for simplicity due to Iwasawa. I would guess that Iwasawa pulled out the essential arguments in the original proofs and then axiomitized them.

Lemma 8.12. *(Iwasawa). Let G act 2-transitively (or in fact even primitively) on a set X. We assume that (1) $G' = G$ and (2) there exists $x \in X$ so that the point stabilizer G_x contains an abelian (or in fact even solvable) subgroup U with $U \lhd G_x$ and such that the G-conjugates of U generate G. Then G/K is simple where K is the kernel of the action.*

Proof.
Assume that $K \lneqq N \lhd G$. It follows from a small variant of Lemma 8.1 with the same proof that N is transitive. (The difference is moving from a permutation group to an action.) Lemma 8.3 then tells us that $G = NG_x$. We next claim that $NU \lhd G$. Since $N \lhd G$ and $U \lhd G_x$, we have $G_x \subseteq N_G(NU)$. Also $N \subseteq N_G(NU)$ since $NU = UN$. It follows that $NG_x \subseteq N_G(NU)$ and since $G = NG_x$, we obtain that $NU \lhd G$.

$NU \lhd G$ now implies that NU contains all of the G-conjugates of U and hence $NU = G$ by assumption (2). The (in some places called) second isomorphism theorem now gives us that

$$G/N = NU/N \cong U/(N \cap U)$$

which is abelian since U is. It follows from Theorem 2.6 that $G' \subseteq N$ and hence by (1) that $N = G$, as desired.
QED

Exercise: Use Iwasawa's Theorem to prove that A_5 is simple.

We will only prove things in the $n = 2$ case (which is a little cheating since we are going to use $PSL(3, 4)$ is simple; this what the "sort of" refers to at the start of this section) but the general proof is not much different. On the other hand, the $n = 2$ case has the advantage that it allows us to see the problem with the two cases which are not simple. We now apply Iwasawa's Theorem to obtain:

Proof of Theorem 7.2 (the $n = 2$ case).

Let $SL(2, F)$ act on the 1-dimensional subspaces of F^2. By Theorem 7.3, this action is 2-transitive and the kernel of this action is the set of scalar matrices contained in $SL(2, F)$. Observe that

$$G_{(1,0)} = \{\begin{bmatrix} a & b \\ 0 & 1/a \end{bmatrix} : \{a, b\} \subseteq F, a \neq 0\}$$

Let

$$U = \{\begin{bmatrix} 1 & * \\ 0 & 1 \end{bmatrix}\}$$

It is easy to check (do it!) that U is abelian and that $U \lhd G_{(1,0)}$ (even though one can check that U is not normal in $SL(2, F)$). The elements in U and their conjugates in $SL(2, F)$ are called *transvections.*

To conclude simplicity of $PSL(2, F) = SL(2, F)/F^*$, we need to prove (1) that U and its G-conjugates generate G and (2) that $(SL(2, F))' = SL(2, F)$. Only (2) will require the assumption that F has at least four elements.

Proof of (1). We will prove the stronger fact that every element of $SL(2, F)$ is product of at most three transvections. The key computation is done in the following lemma.

Lemma 8.13. *For each nonzero $v = (x, y) \in F^2$, there is a transvection or a product of two transvections taking $(1, 0)$ to v. (Only one will be needed if $(1, 0)$ and v are linearly independent.)*

Proof of lemma.
We need the form of a general transvection. If $g = \begin{bmatrix} \alpha & \beta \\ \gamma & \sigma \end{bmatrix} \in SL(2, F)$, an

easy compution (left to the reader) gives that

$$g \begin{bmatrix} 1 & \lambda \\ 0 & 1 \end{bmatrix} g^{-1} = \begin{bmatrix} 1-\alpha\gamma\lambda & \alpha^2\lambda \\ -\gamma^2\lambda & 1+\alpha\gamma\lambda \end{bmatrix}$$

where α and γ cannot both be zero. Now we want

$$\begin{bmatrix} 1-\alpha\gamma\lambda & \alpha^2\lambda \\ -\gamma^2\lambda & 1+\alpha\gamma\lambda \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$$

This gives us the system

$$\begin{bmatrix} 1-\alpha\gamma\lambda = x \\ -\gamma^2\lambda = y \end{bmatrix}$$

If $y \neq 0$, one can take $\gamma = 1$, $\lambda = -y$ and $\alpha = \frac{1-x}{\lambda} = \frac{x-1}{y}$ and we have our desired transvection.

If $y = 0$, one first notes that

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

is a transvection satisfying

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ 0 \end{bmatrix} = \begin{bmatrix} x \\ x \end{bmatrix}.$$

We can then, by the first case, find a transvection taking (x, x) to $(1, 0)$. Composing these two transvections takes us from $(x, 0)$ to $(1, 0)$.
QED

We now show that every element M of $SL(2, F)$ is product of at most three transvections. Let $v = M \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Choose, by the lemma, h which is a product of at most two transvections taking $(1, 0)$ to v. Then $h^{-1}M$ takes $(1, 0)$ to itself and hence

$$h^{-1}M = \begin{bmatrix} 1 & * \\ 0 & * \end{bmatrix}.$$

Since this must have determinant 1, this is in fact a transvection (even in U) and hence M is a product of at most three such.

We now move to (2). We first argue that $U \subseteq (SL(2, F))'$ for $|F| \geq 4$. One can calculate the following commutator

$$\begin{bmatrix} a & 0 \\ 0 & 1/a \end{bmatrix} \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1/a & 0 \\ 0 & a \end{bmatrix} \begin{bmatrix} 1 & -b \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & b(a^2-1) \\ 0 & 1 \end{bmatrix}$$

There is a nonzero a with $(a^2 - 1) \neq 0$ if and only if $|F| \geq 4$. In this case, we can take such an a and get, by varying b, all elements of U. Hence $U \subseteq (SL(2,F))'$.

Since $U \subseteq (SL(2,F))'$ and $(SL(2,F))' \lhd SL(2,F)$, we get that all conjugates of elements of U are contained in $(SL(2,F))'$. Therefore, by (1), we must have $(SL(2,F))' = SL(2,F)$.
QED

Remark: For $n \geq 3$, one has more freedom and there are more transvections to choose from which partially explains why there are no restrictions on F in that case.

8.4 "Almost" Preservation of Simplicity under transitive extensions

The following is our crucial theorem which allows us to often prove simplicity of permutation groups from the simplicity of their stabilizers. Here is the precise statement.

Note for the reader: The reader is advised to first (and maybe instead) read the more streamlined proof of this result given in Subsection 8.5. The present subsection might put things in a more general context but it might be easier to digest the proof in the next subsection.

Theorem 8.14. *Assume $k \geq 2$. Let G be a k-transitive subgroup of S_n and assume one (or equivalently every) stabilizer G_x is simple.*
1. If $k \geq 4$, then G is simple.
2. If $k \geq 3$, then either (1) $n = 2^\ell$ for some integer $\ell \geq 2$, (2) $n = 3$ and $G = S_3$ or (3) G is simple.
3. If $k \geq 2$, then either (1) $n = p^\ell$ for some prime p and some integer ℓ or (2) G is simple.

Exercise: Assuming A_5 is simple, prove using the above theorem that A_n is simple for all $n \geq 6$.

We will need to develop a number of lemmas to prove Theorem 8.14. One of the key ones is the following lemma which is an interesting description of the restrictions that arise when one group acts upon another group by automorphisms.

Lemma 8.15. *Let G and H be any two groups and assume that G acts on H by automorphisms.*
(i). If the G-action restricted to $H \backslash \{1\}$ is transitive, then H is an elementary

abelian group; i.e., isomorphic to $(Z/p)^n$ for some prime p and integer n.
(ii). If furthermore the G-action restricted to $H\backslash\{1\}$ is 2-transitive, then either (1) the p in part (i) is 2 and $n \geq 2$ or (2) $H = Z/3$.
(iii). If furthermore the G-action restricted to $H\backslash\{1\}$ is 3-transitive, then $H = Z/2 \times Z/2$.
(iv). Such a G-action cannot be 4-transitive.

Proof.
(i). First, by transitivity, for each nonidentity elements $x, y \in H$, there must be an automorphism taking x to y and hence they must have the same order. Let p be any prime dividing $|H|$. By Cayley's Theorem, there exists $x \in H$ of order p and hence by the above, every element has order p. This does not in itself imply that H is abelian but once we establish that, it will follow that $H \cong (Z/p)^n$ (if you want by the fundamental theorem of finitely generated abelian groups). However, we now know that H is a p-group and hence by Theorem 2.7 has a nontrivial center. It follows from the transitivity and the fact that group automorphisms preserve the center that the center has to be everything. Hence H is abelian.
(ii). 2-transitivity implies that $|H| \geq 3$. If $H \neq Z/3$, then $|H| \geq 4$. If n in (i) is 1, then we must have $p \geq 4$ and it is clear that the set of automorphisms of Z/p for $p \geq 4$ is not 2-transitive. Hence $n \geq 2$. Now if $p \neq 2$, if we take any $x \neq 1$, we will have $x^{-1} \neq x$ and we can choose y different from x, x^{-1} and the identity. It is clear there is no group automorphism taking x to x and x^{-1} to y since the pair (x, x^{-1}) is "fundamentally different" than the pair (x, y). (If you are not convinced by that, you can do the argument more rigorously.) This contradicts 2-transitivity. Hence p must be 2 and $n \geq 2$.
(iii). 3-transitivity implies $|H\backslash\{1\}| \geq 3$ and so by (ii), we know $H = (Z/2)^n$ with $n \geq 2$. If $n \geq 3$, we can find distinct nonidentity elements x, y, w with $w \neq xy$. (This cannot be done if $n = 2$.) Now there is no automorphism taking (x, y, xy) to (x, y, w) (Why?) and this contradicts the 3-transitivity.
(iv). By (iii), we must have $|H\backslash\{1\}| = 3$ which does not allow any 4-transitive action on it.
QED

Definition 8.16. *A permutation group is regular if all the stabilizers are trivial.*

Lemma 8.17. *Let G be a 2-transitive permutation group such that all the stabilizers are simple. Then either*
(1) G is simple
or
(2) G contains a normal subgroup which is transitive and regular.

Remark: The proof holds under the weaker assumption of primitivity.

Proof.
If G is not simple, let N be a nontrivial normal subgroup of G. By Lemma 8.1, N is transitive. Fix x. $N \lhd G$ easily implies that $N \cap G_x \lhd G_x$ (this is true with G_x replaced by any subgroup of G) and since G_x is simple, either $N \cap G_x = 1$ or $N \cap G_x = G_x$. If the latter occurs, then $G_x \subseteq N$ which implies by maximality of G_x (see Lemma 8.2) that $N = G_x$ or $N = G$. The second is false by assumption and the first contradicts the transitivity of N. Therefore we must have $N_x = N \cap G_x = 1$ and hence N is regular.
QED

At this stage, the reader might wonder about the relevance of Lemma 8.15 since why should results specifically concerning actions by automorphisms be relevant to the behavior of general actions. The next lemma in some sense addresses this point.

Lemma 8.18. *Let G be a transitive subgroup of S_Ω and let H be a nontrivial normal (and hence transitive) regular subgroup of G. Fix $x \in \Omega$. Then G_x acting on $\Omega \backslash \{x\}$ is equivalent to G_x acting on $H \backslash \{1\}$ by conjugation ($g(h) = ghg^{-1}$).*

Proof.
We map $H \backslash \{1\}$ to $\Omega \backslash \{x\}$ by $h \to hx$. Since H is regular, this maps into $\Omega \backslash \{x\}$ as needed. It is easy to check that this map is a bijection. Finally, for the equivalence of the actions, this follows from

$$(ghg^{-1})(x) = g(h(x))$$

being true for all $g \in G_x$ and $h \in H \backslash \{1\}$.
QED

Proposition 8.19. *Let G be a k-transitive subgroup of S_Ω and H a nontrivial normal regular subgroup of G.*
(i). If $k \geq 2$, then $H = (Z/p)^\ell$ for some prime p and integer ℓ (which implies $|\Omega| = p^n$).
(ii). If $k \geq 3$, then either (1) p in (i) is 2 or (2) $H = Z/3$.
(iii). If $k \geq 4$, then $H = Z/2 \times Z/2$.
(iv). If $k \geq 5$, then this is not possible.

Proof.
Fixing $x \in \Omega$, Lemma 8.18 gives us a $(k-1)$-transitive action of G_x on $H \backslash \{1\}$

by automorphisms. We can now simply apply Lemma 8.15 and obtain the conclusion.
QED

We finally move to the

Proof of Theorem 8.14. We will combine Lemma 8.17 and Proposition 8.19 to obtain this. In each of the three cases, Lemma 8.17 tells us that if G is not simple, then it contains a normal, transitive regular subgroup H. Regularity also gives in all cases $|H| = n$.
1. For $k \geq 4$, if we have such an H, then Proposition 8.19 tells us that $H = Z/2 \times Z/2$ and hence $n = 4$. It follows that $G = S_4$ and therefore $G_x = S_3$ which is not simple. Hence G must be simple.
2. For $k \geq 3$, if we have such an H, then Proposition 8.19 tells us that $H = Z/3$ or $H = (Z/2)^\ell$. If the latter case, we get $n = 2^\ell$. In the former case, $n = 3$ and so G must be S_3, completing this case.
3. For $k \geq 2$, if we have such an H, then Proposition 8.19 tells us that $H = (Z/p)^\ell$ and hence $n = p^\ell$, completing this case as well.
QED

We end this subsection by showing how each of the cases in Proposition 8.19 can in fact occur.

For (i), let p be prime and ℓ an integer. The affine group $AGL(\ell, p)$ is the affine group for the vector space $(Z/p)^\ell$. More precisely, for $(v, \gamma) \in (Z/p)^\ell \times GL(\ell, p) = AGL(\ell, p)$, $(v, \gamma)(w) = v + \gamma(w)$. $AGL(\ell, p)$ can also be described as a semi-direct product

$$(Z/p)^\ell \times_\theta GL(\ell, p)$$

where the second factor acts naturally on the first factor. $(Z/p)^\ell$ is then a normal regular subgroup and since the stabilizer of 0 is $GL(\ell, p)$ which is transitive off of 0, the action of the full group, being obviously transitive, is 2-transitive.

For the first possibility in (ii), with $p = 2$ and ℓ any integer, we can do the exact same construction. However, now, as long as $\ell \geq 2$, $GL(\ell, 2)$ is 2-transitive off of 0 and so the action of the full group, being obviously transitive, becomes 3-transitive. For the second possibility, S_3 (as a subset of itself) is 3-transitive and $\langle(123)\rangle$ is a normal regular subgroup.

For (iii), we use the same example as the first part of (ii) with ℓ also being 2. Now the 3-transitivity implies 4-transitivity since the degree of the action is 4.

For $\ell \geq 2$, one never has sharp 2-transitivity. For $\ell = p = 2$, one has sharp 3 (or 4)-transitivity. For $p = 2$ and $\ell \geq 3$, one does not have sharp 3-transitivity. For $p \geq 2$ and $\ell = 1$, one has sharp 2-transitivity.

Exercise: Sort out if the last paragraph covers all cases of sharp transitivity.

8.5 A more direct proof of the result in the previous section

We give another proof of Theorem 8.14 whose statement we repeat here. The steps of the proof are essentially the same but the proof might feel more direct and might be easier for the reader to digest.

Theorem 8.20. *Assume $k \geq 2$. Let G be a k-transitive subgroup of S_n and assume one (or equivalently every) stabilizer G_x is simple.*
1. If $k \geq 4$, then G is simple.
2. If $k \geq 3$, then either (1) $n = 2^\ell$ for some integer $\ell \geq 2$, (2) $n = 3$ and $G = S_3$ or (3) G is simple.
3. If $k \geq 2$, then either (1) $n = p^\ell$ for some prime p and some integer ℓ or (2) G is simple.

Proof.
For the first half of the proof, all the three cases will be dealt with at the same time. Assume G is not in fact simple. Then there exists a normal subgroup N with $1 \lneq N \lneq G$.

Step 1: N is sharp transitive (i.e. is transitive and regular). This implies that $|N| = n$.
By Lemma 8.1, N is transitive. Fix x. $N \lhd G$ easily implies that $N \cap G_x \lhd G_x$ (this is true with G_x replaced by any subgroup of G) and since G_x is simple, either $N \cap G_x = 1$ or $N \cap G_x = G_x$. If the latter occurs, then $G_x \subseteq N$ which implies by maximality of G_x (see Lemma 8.2) that $N = G_x$ or $N = G$. The second is false by assumption and the first contradicts the transitivity of N. Therefore we must have $N_x = N \cap G_x = 1$ and hence N is regular.

Step 2: For fixed x, the action of G_x on $N \backslash \{1\}$ by conjugation ($g(n) = gng^{-1}$) is $(k-1)$-transitive.
Fix x. We know by Lemma 2.5 that the action of G_x on $\{1, \ldots, n\} \backslash \{x\}$ is $(k-1)$-transitive. It suffices to show that these two actions are equivalent. To see this, we map $H \backslash \{1\}$ to $\{1, \ldots, n\} \backslash \{x\}$ by $h \to hx$. Since H is regular, this does map into $\{1, \ldots, n\} \backslash \{x\}$ as needed. It is easy to check (using Step 1) that this map is a bijection. Finally, for the equivalence of the actions,

this follows from

$$(ghg^{-1})(x) = g(h(x))$$

being true for all $g \in G_x$ and $h \in H\backslash\{1\}$.

We now continue with the three different parts of the statement of the theorem.

Step 3: Statement 3 of the theorem.
We have by Step 2 that G_x acts transitively on $N\backslash\{1\}$ by automorphisms. It follows that all elements of $N\backslash\{1\}$ have the same order. Let p be any prime dividing $|N|$. By Cayley's Theorem, there exists $x \in N$ of order p and hence by the above, every element has order p. Therefore N is a p-group and hence by Theorem 2.7 has a nontrivial center. Since the center is invariant under all automorphisms, transitivity implies that the center has to be everything and hence N is abelian. It follows from the fundamental theorem of finitely generated abelian groups that $N \cong (Z/p)^\ell$ for some integer ℓ and hence, by the last statement of Step 1, $n = p^\ell$ as claimed.

Step 4: Statement 2 of the theorem.
We have by Step 2 that G_x acts 2-transitively on $N\backslash\{1\}$ by automorphisms. Hence $|N| \geq 3$. If $|N| = n = 3$, then $G = S_3$ since it is 3-transitive. Otherwise $|N| \geq 4$.

Next, by the proof of Step 3, we have $N \cong (Z/p)^\ell$ for some prime p and integer ℓ. If $\ell = 1$, then N is cyclic in which case the automorphisms of N cannot be 2-transitive unless $p = 3$ but now we have that $|N| \geq 4$. Hence we must have $\ell \geq 2$. Now if $p \neq 2$, if we take any $x \neq 1$, we will have $x^{-1} \neq x$ and we can choose y different from x, x^{-1} and the identity. It is clear there is no group automorphism taking x to x and x^{-1} to y which contradicts 2-transitivity. Hence p must be 2 and $\ell \geq 2$.

Step 5: Statement 1 of the theorem.
We have by Step 2 that G_x acts 3-transitively on $N\backslash\{1\}$ by automorphisms. Hence $|N| \geq 4$. By Step 4, we know that $N = (Z/2)^\ell$ with $\ell \geq 2$. If $\ell \geq 3$, we can find distinct nonidentity elements x, y, w in N with $w \neq xy$. (This cannot be done if $\ell = 2$.) Now there is no automorphism taking (x, y, xy) to (x, y, w) contradicting 3-transitivity. Hence $\ell = 2$ and so $n = |N| = 4$. Since G is 4-transitive, we must have $G = S_4$ and hence $G_x = S_3$, contradicting G_x being simple. Hence G must have been simple to begin with.
QED

8.6 Putting it all together: Simplicity of the 5 Mathieu groups

The simplicity of all five of the Mathieu groups will now follow from combining these last four subsections.

Theorem 8.21. *The five Mathieu groups are simple.*

Proof.
M_{11} is simple by Corollary 8.5. Since M_{12} is 5-transitive and has M_{11} as its 1-point stabilizer (i.e., is a transitive extension), Theorem 8.14(1) tells us that M_{12} is simple.

Next, since $PSL(3,4)$ is simple, M_{22} is 3-transitive and has $PSL(3,4)$ as its 1-point stabilizer (i.e., is a transitive extension), it follows from Theorem 8.14(2) that M_{22} is simple since 22 is not of the form 2^k and (of course) M_{22} is not S_3. Since M_{23} and M_{24} are 4 and 5-transitive respectively and have M_{22} and M_{23} as their respective 1-point stabilizers, Theorem 8.14(1) allows us to deduce that M_{23} is simple from the simplicity of M_{22} and that M_{24} is simple from the now established simplicity of M_{23}.
QED

Note that while we used Corollary 8.5 for the simplicity of M_{11}, the simplicity of M_{23} proved in that corollary was not used in the above argument.

9 Some other fun permutation group results

9.1 Jordan's (other) Theorem.

If $G \subseteq S_n$ contains an n-cycle and a transposition, it might or might not be all of G; it depends on the exact relationship between the n-cycle and the transposition. For example $\langle (1\ 2 \ldots\ n), (1\ 2) \rangle = S_n$ (exercise) but the dihedral group $D_{2n} \subseteq S_n$ contains an n-cycle and a transposition but is not equal to S_n if $n \geq 4$. If G is 2-transitive and contains a transposition, then by conjugation it contains all transpositions and hence is S_n. The following stronger statement is interesting.

Theorem 9.1. *If $G \subseteq S_n$ is primitive and contains a transposition, then $G = S_n$.*

Proof.
Define an equivalence relation on $\{1, \ldots, n\}$ by $x \sim y$ if $x = y$ or the transposition $(x\ y)$ is in G. The only point which requires verification is transitivity and then only in the nontrivial case $x \sim y \sim z$ and $x \neq y \neq z$. In that case, since the transpositions $(x\ y)$ and $(y\ z)$ are in G, we also have

$$(x\ z) = (y\ z)(x\ y)(y\ z) \in G.$$

Next G preserves this equivalence relation since if $g \in G$

$$x \sim y \to (x\ y) \in G \to g(x\ y)g^{-1} \in G \to (gx\ gy) \in G \to g(x) \sim g(y).$$

The latter implies that each equivalence class is a block. (Why?) We have not used *either* of the two assumptions yet! But we do now. Choose now a transposition and consider the equivalence class containing the two points of this transposition. Since this is a block with at least two points, primitivity implies that this equivalence class is all of $\{1, \ldots, n\}$ and hence all transpositions are in G.
QED

9.2 Exotic Embeddings of S_5 into S_6 and non-inner automorphisms for S_6

The following topic seems to be a must in any set of notes like this. Recall that an *inner* automorphism of a group is an automorphism which comes from conjugation. That is, it is the map given by

$$h_g(x) = gxg^{-1}$$

for a fixed $g \in G$.

First, we recall an easy background lemma which follows from the first isomorphism theorem applied to the mapping $g \to h_g$.

Lemma 9.2.

$$G/C(G) \cong Inn(G)$$

where $Inn(G)$ is the collection of inner automorphisms which is a (normal) subgroup of $Aut(G)$ under composition.

There are many automorphisms of a group which are not inner. For example, abelian groups have no inner automorphisms. But this question becomes interesting when applied to S_n.

Question: For S_n, are their automorphisms which are not inner automorphisms?
The interesting answer is given by the following theorem.

Theorem 9.3. *The only n's such that there is a non-inner automorphism is $n = 6$. (We will prove this result for $n = 6$ and discuss/outline the result for other n.)*

Remark: It is good to have the following picture in mind. An inner automorphism preserves each conjugacy class, just permuting the elements within a conjugacy class. A general automorphism need not do that but it does have to permute the *family* of conjugacy classes; i.e. the image of a conjugacy class under an automorphism must be a conjugacy class but it does not have to be the one you started with. Interestingly, there are however non-inner automorphisms of some groups which do in fact preserve each conjugacy class.

Proof. $n = 1, 2$ have no automorphisms at all.

We first show that S_6 has a non-inner automorphism. To do this, the first step we extract as a lemma, not because it is hard, but to emphasize its interest. The following states that there are so-called *exotic* embeddings of S_5 into S_6, meaning one is not just doing something trivial like taking the stabilizer of a point.

Lemma 9.4. *There exists a transitive subgroup of S_6 which is isomorphic to S_5.*

Proof of lemma. The number of Sylow 5-subgroups of S_5 is six since these subgroups are disjoint other than 1 (as 5 is prime), there are 4! 5-cycles altogether and each Sylow 5-subgroup contains 4 nontrivial elements. Next S_5 acts on these six Sylow 5-subgroups by conjugation and the Sylow Theorems tell us that this is a transitive action. Hence this gives rise to a map from S_5 to S_6 whose image is a transitive subgroup. The size of the stabilizer of any one of these subgroups is $5!/6 = 20$ and hence the kernel of the action has size at most 20. However, the normal subgroups of S_5 are just A_5, which has size 60, or the trivial ones. Hence the kernel is trivial and the above map is an embedding. QED

We now construct a non-inner automorphism of S_6. Letting H denote this exotic transitive subgroup of S_6 which is isomorphic to S_5, we have a transitive action of S_6 on the left cosets of H by translations.

$$g_1(g_2H) = g_1g_2H.$$

Clearly H as index six in S_6 and hence the above action gives us a homomorphism from S_6 to itself. The kernel of this action is contained in H and hence has size at most $5! < 6!/2$. Since S_6 also just has three normal subgroups, the kernel must be trivial yielding us an automorphism of S_6.

We claim it is not inner. It is clear that an inner automorphism of S_n must preserve transitivity (check this); i.e. a transitive subgroup must be mapped to a transitive subgroup under an inner automorphism since the conjugacy is just a renaming of the elements in $\{1, \ldots, n\}$. However the above automorphism takes the transitive subgroup H and maps it exactly to the stabilizer of the coset H which is not a transitive subgroup. Hence the map cannot be inner.

Let's move on to other values of n and discuss what is happening. Let T_k denote the conjugacy class of permutations which consist of k transpositions. Note that the union of these T_k's are of course the involutions of S_n. Since automorphisms must send a conjugacy class to another conjugacy class and of course send involutions to involutions, any automorphism of S_n must send T_1 to some T_k. In such a case, we must have of course $|T_1| = |T_k|$. Interestingly, one can check that for $n = 6$, $|T_1| = 15 = |T_3|$ and in fact (if I remember correctly!) our non-inner automorphism maps T_1 to T_3.

If we move to $n \neq 1, 2, 6$, a combinatorial check shows that there is no $k \neq 1$ with $|T_1| = |T_k|$ and as a result an automorphism must send the transpositions to themselves. One can then show that this then implies that the permutation is inner. I haven't had the energy to organize this argument yet (and might not ever!) S_3 is slightly degenerate here since

$T_2 = \emptyset$. This case is easy to deal with directly. First note that since the center is trivial, we have 6 inner automorphisms and so we just have to show that the total number of all automorphisms is at most 6. The last claim is easy since an automorphism permutes the three transpositions (since these are the only things of order 2) and the three transpositions generate S_3 and so any automorphism is determined by what it does to the three transpositions. QED

Exercise: (i) Show there is an *exotic* embedding of S_5 into S_{20} as a transitive subgroup.
Hint: There is a canonical copy H of S_3 inside of S_5. Let S_5 act on the cosets of H by translation.
(ii) Can S_5 be embedded into S_{20} as a 2-transitive subgroup?
(iii) Can one generalize exercise (i)?

9.3 Consequences of solvability and nilpotence for the degree of permutation groups.

Theorem 1.2 puts certain limitations on the values of n for when one can have a transitive subgroup G of S_n which is sharp 4 or 5-transitive. Interestingly, if one makes certain group-theoretic assumptions about the group G, there are also certain restrictions on the values of n for when G can sit inside S_n as a primitive transitive subgroup, without assuming any higher order transitivity. The following theorem gives us such a result.

Theorem 9.5. *1. If G is a transitive primitive subgroup of S_n which is solvable, then $n = p^k$ for some prime p and integer k.*
2. If G is a transitive primitive subgroup of S_n which is nilpotent, then $n = p$ for some prime p.

Remarks:
(i). Part 1 cannot be true without the solvablility assumption by considering any of the Mathieu groups.
(ii) Part 2 cannot be true just under the solvablility assumption by considering S_4 (with $n = 4$), which, together with S_3, is a basic example of a nonnilpotent solvable group.
(iii) If we only assume transitivity and not primitivity, then this question is uninteresting since Cayley's Theorem gives us an embedding of any group G into $S_{|G|}$ as a transitive subgroup.

Proof of 2.
Since nilpotent groups have nontrivial centers (a fact you can look up in a

standard book), we can (by Cauchy's Theorem) choose $g \in C(G)$ of order some prime p. Letting $H := \langle g \rangle$, we have $H \lhd G$ and hence by Lemma 8.1 (or actually the remark afterwards that the lemma holds under the weaker assumption of primitivity) H is also transitive. It follows that n divides $|H|$ and hence n is prime.

For the proof 1., we need the following lemma whose proof we give afterwards.

Lemma 9.6. *If G is a solvable group, then every (nontrivial) minimal normal subgroup is an elementary abelian p-group.*

The lemma tells us that a solvable group contains a normal elementary p-group H and then, as in the proof of 2., H must be transitive and hence n divides $|H|$. It follows that n is a prime power. QED

Proof of Lemma 9.6.
Let H be a (nontrivial) minimal normal subgroup of the solvable group G. It is possible that H properly contains nontrivial normal subgroups (subgroups which are normal in H but will not be normal in G) since "being a normal subgroup" is not a transitive relation. However, H cannot contain any nontrivial characteristic subgroups since such a subgroup would be normal in G. (Recall a characteristic subgroup of a group is one which is invariant under all automorphisms of the larger group; normal subgroups are, by definition, those invariant just under automorphisms coming from conjugation.)

Since G is solvable, H, being a subgroup, must also be solvable which implies that H' (the derived group of H) cannot equal H. Since H' is characteristic in H (since it is "intrinsically defined"), H' must be trivial which implies that H is abelian. Next, if p is any prime dividing H, then the set of elements of order p is a nontrivial characteristic subgroup of H and hence equals H. Thus H is an elementary abelian p-group.
QED

Theorem 9.5 has a nice consequence which doesn't mention permutation groups.

Corollary 9.7. *1. If G is a solvable group, then every maximal subgroup has prime power index.*
2. If G is a nilpotent group, then every maximal subgroup has prime index.

Proof:
1. Let G be solvable and H be a maximal subgroup. We get a transitive action of G on G/H by translation in the usual way. Since the stabilizer of the coset H is H which is assumed to be a maximal subgroup, the action is

primitive by the remark after Lemma 8.2. By Theorem 9.5.1, $|G/H|$ must be a prime power.
2. This is proved in the same way using Theorem 9.5.2 instead.
QED

9.4 Restrictions on the degree of permutation groups under transitivity assumptions

If G is a subgroup of S_n, we say that n is the *degree* of the permutation group G. Let us recall some of the results we have concerning the possible degree of certain permutation groups. This will look a little bit like the questions addressed in Proposition 8.19 but here we are looking at *sharp* transitivity and there is no assumption concerning a normal subgroup which is transitive and regular.

First recall that Theorem 1.2 included the following three statements.

- There is no sharp 6-transitive subgroup of S_n for any n.
- If G is a sharp 5-transitive subgroup of S_n, then $n = 12$.
- If G is a sharp 4-transitive subgroup of S_n, then $n = 11$.

There are a number of questions left of this type which are the following.

- (i) What are the possible degrees for sharply transitive permutation groups?
- (ii) What are the possible degrees for sharply 2-transitive permutation groups?
- (iii) What are the possible degrees for sharply 3-transitive permutation groups?
- (iv) What are the possible degrees for (nontrivial) primitive permutation groups?
- (v) What are the possible degrees for (nontrivial) 2-transitive permutation groups?
- (vi) What are the possible degrees for (nontrivial) 3-transitive permutation groups?

- (vii) What are the possible degrees for (nontrivial) k-transitive permutation groups for $k \geq 4$?

I will now comment on each of these questions.

(i) is easy. Cayley's Theorem tells us that any group G with $|G| = n$ can be injectively mapped into S_n as a sharp 1-transitive subgroup (the sharpness refers to the permutation group being regular meaning all of the stabilizers are trivial). Hence any integer n can arise as the degree of some sharply 1-transitive permutation group.

(ii) will be dealt with in Theorem 9.8 below.

(iii) will be dealt with in Theorem 9.9 below. Recall that Theorem 1.5(4) answered this (and even gives an exact characterization of such permutation groups) but that was not proved in these notes. Theorem 9.9 will prove the possible degrees but will not characterize the groups and hence will not prove Theorem 1.5(4).

(iv) This question which seemed to have interested Jordan seems difficult and there is a lot to say about it. Interestingly, for example, there seem to be an infinitely many n which do not appear as the degree of a (nontrivial) primitive permutation group but this might require the classification theorem. (This result immediately implies the same statement if "primitive" is replaced by k-transitive for any $k \geq 2$.) The O'Nan-Scott Theorem arises here since the characterization of such subgroups is the content of the famous O-Nan-Scott Theorem which places them into five isomorphism classes. Looking into a **very** small part of this might be a nice project for a student.

(v) It seems this question is discussed in Dixon and Mortimer.

(vi) ???

(vii) This was completely answered in Theorem 1.4 where it was stated that the classification theorem is needed.

The following theorem gives the exact answer for the sharply 2-transitive case.

Theorem 9.8. *There is a sharply 2-transitive subgroup of S_n if and only if $n = p^k$ for some prime p and integer k.*

Proof.
First, Theorem 4.1(1) showed that for any $n = p^k$, there is a sharply 2-transitive subgroup of S_n.

Our goal is now to prove the converse. So let G be a sharply 2-transitive subgroup of S_n. The sharp 2-transitivity implies that $|G| = n(n-1)$. Furthermore, letting G_1 be all the elements in G which have 1 fixed point and G_0 be all the elements in G which have 0 fixed points, the sharp 2-transitivity implies that $G = G_1 \cup G_0 \cup \{1\}$ since only the identity fixes two or more points.

Now, the orbit stabilizer theorem tells us that $|G_a| = |G|/n = n-1$. Also, the sharp 2-transitivity clearly gives

$$G(1) = \cup_{a\in\{1,\ldots,n\}} G_a \backslash \{1\}$$

with this being a disjoint union. Hence $|G(1)| = n(n-2)$ and hence $|G(0)| = |G| - |G(1)| - 1 = n-1$.

To show that $n = p^k$, we need to show that n is divisible but at most one prime number. Letting p be a prime divisor of n, we also have that p divides $|G|$ and we let, by Cayley's Theorem, $\pi \in G$ have order p. We will now show that $G(0)$ is precisely the conjugates of π and hence every element of $G(0)$ has order p. Clearly, there is at most one p with this property.

Next, clearly π must consist of p-cycles and 1-cycles. Since p divides n, p must also divide the number of 1-cycles. However, π has at most one 1-cycle by sharp 2-transitivity and hence there are no 1-cycles. So π just consists of p-cycles and so belongs to G_0.

Note that for every a,

$$G_a \cap C_G(\pi) \subseteq G_a \cap \pi G_a \pi^{-1} = G_a \cap G_{\pi a} = 1,$$

the last equality following from sharp 2-transitivity and that $\pi a \neq a$.

The above implies that

$$[G : C_G(\pi)] \geq |G_a| = n-1$$

and hence the number of conjugates of π is at least $n-1$ (G acts on G by conjugation and the orbit of π is the index of its stabilizer which is just its centralizer). Since all the conjugates clearly are in $G(0)$ which has size $n-1$, this shows that $G(0)$ is precisely the conjugates of π.
QED

This allows us to easily conclude

Theorem 9.9. *There is a sharply 3-transitive subgroup of S_n if and only if $n = p^k + 1$ for some prime p and integer k.*

Proof.
First, Theorem 4.1(2) showed that for any $n = p^k + 1$, there is a sharply 3-transitive subgroup of S_n. Conversely, if we have a sharply 3-transitive subgroup of S_n, then any of its stabilzers is a sharply 2-transitive subgroup of S_{n-1} and hence by Theorem 9.8, we have that $n - 1$ equals $n = p^k$ for some prime p and integer k.
QED

We would like to now go further here concerning the sharply 2-transitive permutation groups by describing their structure, not only their degrees which was done above.

Theorem 9.10. *If G is a sharply 2-transitive subgroup of S_n, then (1) $n = p^k$ for some prime p and integer k, (2) G contains a normal Sylow p-subgroup and (3) G is equivalent to a subgroup of $AGL(k, p)$ which contains its translation subgroup.*

Proof.
Theorem 9.8 proved (1) and we will use parts of that proof. We know that $|G| = p^k(p^k - 1)$. Let P be a Sylow p-subgroup of G.

Step 1: $P \lhd G$.
If $g \in P\backslash\{1\}$, then g has order p^k. Considering the possible cycle structure of g, $n = p^k$ and the sharp 2-transitivity telling us that there is at most one fixed point, we can conclude $g \in G(0)$. Since $P\backslash\{1\}$ and $G(0)$ have the same cardinality, we conclude that $P\backslash\{1\} = G(0)$. Clearly $G(0)$ is invariant under conjugation by elements in G (and even by elements in S_n) proving $P \lhd G$.

Let N be a minimal normal subgroup of G which is contained in P. Note that we need Step 1 for the existence of N.

Step 2: N is abelian.
Since N is minimal normal, it cannot contain any nontrivial characteristic subgroups (since such subgroups would be normal in G). The center of N is a characteristic subgroup which is nontrivial since N is a p-group. Hence $C(N) = N$ and N is abelian.

Step 3: N is elementary abelian.
The set of order p elements in N is a characteristic subgroup and being nontrivial by Cayley's Theorem, we conclude this set is all of N; i.e. N is elementary abelian. (Even if we did not already know that N is a p-group, we could have reached the same conclusion just assuming abelian and "characteristically simple".)

Step 4: N is transitive and regular and hence $N = p^k$.
Lemma 8.1 gives us transitivity of N. Therefore all of the N-stabilizers are conjugate within N and hence are the equal since N is abelian. They therefore must be trivial since any element of one of them must fix everything.

Step 5: For any a, $G = NG_a$. (Note the regularity of N gives $G_a \cap N = 1$.)
Given $g \in G$, choose by transitivity of N, $h \in N$ so that $h(a) = g(a)$. Then $h^{-1}g \in G_a$ and so $g \in NG_a$.

So we have G is a semidirect product of N and G_a. The natural action of G on N can then be seen to be equivalent to the original action. (Exercise!).
QED

Although one did not need it in the above proof, one can argue that N above is the unique minimal normal subgroup of G as follows. One first shows that N is its own centralizer in G. N being abelian gives one direction and for the the other direction, if $g \in C_G(N)$, write $g = g_a n$ with $g_a \in G_a$ and $n \in N$. We want to show $g_a = 1$ and we clearly have $g_a \in C_G(N)$. Given any b, choose $n' \in N$ so that $n'(a) = b$. Then

$$g_a(b) = g_a n(a) = n g_a(a) = n(a) = b$$

and so $g_a = 1$. Now if there were another minimal normal subgroup N', then $N \cap N' = 1$ which implies, as usual, that N and N' commute. Hence $N' \subseteq C_G(N) = N$, a contradiction.

Remark
One can check that the proof of the above theorem (and paragraph after the proof) gives that if a primitive permutation group contains a minimal normal subgroup which is abelian, then it follows N is elementary abelian, N is transitive and regular and hence $n = p^k$, N is the unique minimal subgroup and G is a semi-product as described above.

10 Designs and Steiner systems

10.1 Designs and Steiner systems: the general setup

The following combinatorial concept has been an object of very much study. If one is *only* interested in the connections with the Mathieu groups, one does not in fact need much more from this section than the definition of a design given next.

Definition 10.1. *An (n, k, t, λ)-design is a set X with n elements, a collection $\mathcal{B}$ of subsets (called boxes) of X each of size k such that each subset Y of X with t elements is contained in λ many boxes in $\mathcal{B}$.*

Remark: (only for those who this means something to): This is the same (by definition!) as a k-uniform hypergraph on n vertices such that each t-set is contained in λ edges.

The following questions immediately present themselves.
1. Given (n, k, t, λ), does there exist an (n, k, t, λ)-design?
2. If there is one, how many are there up to isomorphism? (There is an obvious notion of isomorphism for such objects and so we are asking about the number of isomorphism classes.)
3. Given an (n, k, t, λ)-design, what is the automorphism group of the design? (An automorphism f of the design is a bijection from X to X which preserves the box structure $\mathcal{B}$; i.e. $B \in \mathcal{B}$ if and only if $f(B) \in \mathcal{B}$.) If there is more than one design (up to isomorphism) for (n, k, t, λ), we certainly expect that the automorphism group depends on the isomorphism class.

It can be interesting to start off and list some examples of what is known and what is not known for one specific class. For each n, we consider the design with parameters $(n^2 + n + 1, n + 1, 2, 1)$ which we denote by $P(n)$. (So a $P(n)$ design is a design on a set with $n^2 + n + 1$ points.) We will look at this class, which are called projective planes (which explains the notation $P(n)$) in more detail later but here is what is known about them (which I hope is up to date). To the right is the number of isomorphism classes; 0 simply means that there is no such design.

$P(2)$	1
$P(3)$	1
$P(4)$	1
$P(5)$	1
$P(6)$	0
$P(7)$	1

$P(8)$	1
$P(9)$	4
$P(10)$	0
$P(11)$	≥ 1, not known if > 1.
$P(12)$	unknown if there exist any

These are known to exist when n is a prime power and we can use finite fields for their construction. There is no known n which is not a prime power where such a design is known. There are various other results known but we stop here for this example.

It is appropriate to mention a big theorem from 2015 by P. Keevash which is considered the largest progress in the field in the last 100 years. Until recently, it was not known if there were any (nontrivial) designs with $t \geq 6$ and if there were infinitely many with $t = 4, 5$. There are some, as we will see, trivial divisibility conditions that the parameters (n, k, t, λ) must satisfy in order for there to exist a design with those parameters.

Theorem 10.2. *(Keevash) For all k, t, and λ, there exists $n = n(k, t, \lambda)$ such that for all $n \geq n_0$, there exists an (n, k, t, λ)-design assuming the "trivial divisibility conditions".*

The proof of this which is considered exceedingly difficult uses, among other things, the so-called "probabilistic method".

There are two other natural parameters associated to a design which are $b = |\mathcal{B}|$, the number of boxes, and r which is the number of boxes containing a fixed point. It turns out that r is independent of the point and b and r are functions of the initial four parameters. This follows from the following theorem, where the first equation yields r in terms of the 4 parameters (and proves its independence of the point) and then the second equation allows us to then compute b.

Theorem 10.3. *Assume there exists an (n, k, t, λ)-design and define b and r as above. Then.*
1. $\binom{n-1}{t-1} \lambda = r \binom{k-1}{t-1}$.
2. nr=bk.
3. If $\lambda = 1$, then $b = \frac{\binom{n}{t}}{\binom{k}{t}}$.

Proof.
1. Fix $a \in X$. Consider the set $\{(S, B) : |S| = t - 1, B \in \mathcal{B}, \{a\} \cup S \subseteq B\}$ and double count (i.e. Fubini).

2. Consider the set $\{(x, B) : x \in B \in \mathcal{B}\}$ and double count.
3. Consider the set $\{(Y, B) : Y \subseteq B \in \mathcal{B}, |Y| = t\}$ and double count. QED

While part 3 can easily be deduced from the first two parts, since we will often be in the case that $\lambda = 1$, we wanted to emphasize the direct formula for b and the simplicity of the proof in this special case.

Corollary 10.4. *If there exists an (n, k, t, λ)-design, then both of the following are integers:*

$$\frac{\lambda\binom{n-1}{t-1}}{\binom{k-1}{t-1}} \qquad \frac{n\lambda\binom{n-1}{t-1}}{k\binom{k-1}{t-1}}$$

So, for example, there is no $(11, 6, 2, 2)$-design since the second fraction is not an integer (the first one is).

There are other further necessary conditions, for example those in the next theorem.

Theorem 10.5. *Let r_i be the number of boxes containing an i-set (so r_1 is our previous r.) Then for all $i \in \{0, \ldots, n-1\}$, we have*

$$\lambda\binom{n-i}{t-i} = r_i\binom{k-i}{t-i}$$

and hence we must have

$$\binom{k-i}{t-i} \text{ divides } \lambda\binom{n-i}{t-i}$$

Proof: Exercise: Find an appropriate combinatorial collection and apply double counting to prove the equality.
QED

Historically, one often restricted the parameters in a design under consideration. Some of these are given in the following definitions. The terminology is not always so consistent.

Definition 10.6. *A* **general Steiner system** *is an $(n, k, t, 1)$-design. These will be denoted by $S(n, k, t, 1)$ (S for Steiner).*

Definition 10.7. *A* **classical Steiner system** *is an $(n, k, 2, 1)$-design.*

We next have the following three important infinite classes of classical Steiner systems, each having one parameter.

Definition 10.8. *A* **Steiner triple system with parameter** n *is an* $(n, 3, 2, 1)$*-design.*

Definition 10.9. *A* **projective plane with parameter** n *is an* $(n^2 + n + 1, n + 1, 2, 1)$*-design.*

Definition 10.10. *An* **affine plane with parameter** n *is an* $(n^2, n, 2, 1)$*-design.*

10.2 Steiner Triple Systems

The following is the main theorem concerning Steiner triple systems. We will only prove the easy direction.

Theorem 10.11. *A Steiner triple system with parameter n (i.e. an* $(n, 3, 2, 1)$*-design) exists if and only if*

$$n \equiv 1 \text{ or } 3 \ \ (\ mod \ 6).$$

(Easy part of proof:) There are two different constructions depending on whether $n \equiv 1$ or $\equiv 3$ (mod 6). Here we prove that these are necessary conditions. Theorem 10.3 immediately gives us

$$n - 1 = 2r, nr = 3b.$$

The first equality gives that n is odd, ruling out three possible residues mod 6. So, we need to just rule out 5 as a possible residue mod 6. First, since 3 divides $nr = \frac{n(n-1)}{2}$, we have 6 divides $n(n - 1)$. However if $n = 6k + 5$, then

$$n(n - 1) = (6k + 5)(6k + 4) = 36k^2 + 54k + 20$$

which is not divisible by 6.
QED

The number of isomorphism classes for small n is quite fascinating. Ignoring the cases $n = 1$ (which I guess doesn't really exist) and $n = 3$ which is trivial, the first five real cases are $n = 7, 9, 13, 15, 19$. The number of isomorphism classes for these first five cases are
1
1
2
80
11,084,874,829! (the ! is not of course a factorial but for emphasis).

10.3 The projective plane

In this subsection, we discuss the projective plane with parameter n, which is the second classical Steiner system mentioned and is, we recall an

$$(n^2 + n + 1, n + 1, 2, 1) - \text{design}.$$

Using finite fields, it is fairly easy to prove the following theorem.

Theorem 10.12. *If $n = p^\ell$ for some prime p and integer ℓ, then there exists a projective plane with parameter n.*

Proof.
Let F be a finite field with $n = p^\ell$ elements and let $V = F^3$ be a 3-dimensional vector space over F. Let X be the set of 1-d dimensional subspaces in V. It is an easy exercise to show that $|X| = \frac{n^3-1}{n-1} = n^2 + n + 1$. Let $\mathcal{B}$ be the set of 2-d dimensional subspaces in V; i..e., to be more precise, a box is the set of 1-d subspaces which is contained in a given 2-d dimensional subspace. It is another elementary exercise to show that a box has size $\frac{n^2-1}{n-1} = n+1$. Finally, each pair of distinct 1-d dimensional subspaces is contained in a unique 2-d subspace, namely the 2-d subspace that they generate. Hence we have an $(n^2 + n + 1, n + 1, 2, 1)$-design.
QED

As previously mentioned, it is not known if there an n which is not a prime power for which there is a projective plane with that parameter. Let us however note that for $n = 6$, which would be a $(43, 7, 2, 1)$-design, this does not exist. This follows immediately from Corollary 10.4 noting that the second fraction is not an integer (the first one is).

We can immediately determine, using Theorem 10.3, what r, the number of boxes containing a fixed point and b, the number of boxes, are. These are

$$r = n + 1, \quad b = n^2 + n + 1. \tag{3}$$

Exercise: For a projective plane, it is sometimes written as an assumption that the intersection of any two boxes contain one element (colloquially, "two lines intersect in one point"). It is clear that in the construction above using finite fields that this holds since the intersection of two distinct 2-d subspaces is a 1-d space. Prove that the intersection of two boxes has 1 element for every projective plane, not only those as constructed above. Note that this is first relevant for $P(9)$ which is a $(91, 10, 2, 1)$-design since this is the smallest projective plane for which there is more than one isomorphism class.

It might seem sort of arbitrary that we *defined* a projective plane as we did, namely as an $(n^2+n+1, n+1, 2, 1)$-design which of course has a very specific form. However, there is a theorem from combinatorial geometry that motivates this definition. First, it is immediate to check that for a projective plane $P(n)$ with $n \geq 2$, we have
(i) any two distinct points are contained in a unique box,
(ii) the intersection of any two boxes is a single point
(iii) for every point, there are at least 2 boxes that it is not contained in and
(iv) for every box, there are at least 2 points that is not contained in the box.

(i) is true by definition, (ii) is the above exercise and (iii) and (iv) are easily checked by counting. The following theorem is now interesting.

Theorem 10.13. *Assume we have a set X and a collection of subsets (boxes) which satisfy (i), (ii), (iii) and (iv) above. Then this collection is $P(n)$ for some n.*

10.4 The affine plane

In this subsection, we discuss the affine plane with parameter n, denoted $A(n)$, which is the third classical Steiner system mentioned and so is an $(n^2, n, 2, 1)$-design. Using finite fields, it is fairly easy to prove the following theorem.

Theorem 10.14. *If $n = p^\ell$ for some prime p and integer ℓ, then there exists an affine plane with parameter n.*

Proof.
Let F be a finite field with p^ℓ elements and let $X = F^2$ be a 2-dimensional vector space over F so that $|X| = n^2$. Let $\mathcal{B}$ be the set of all translates of 1-d subspaces. Each box has size n. Next, for each pair of points in X, there is a unique element of $\mathcal{B}$ containing them. Hence we have a $(n^2, n, 2, 1)$-design.
QED

We give an alternative proof of this constructing it from any projective plane. Take any projective plane $P(n)$ and "remove a line"; i.e., we remove all the points in one of the boxes B. This leaves a set with n^2 elements. Since each pair of boxes intersect in 1 element (this is clear for our construction of the projective plane but is true for any projective plane by the exercise above), all the remaining boxes have size n. Moreover, each pair of points which remain clearly are still contained in a unique box. This yields an $(n^2, n, 2, 1)$-design.
QED

As is the case for the projective plane, it might seem sort of arbitrary that we *defined* an affine plane as we did, namely as an $(n^2, n, 2, 1)$-design which of course has a very specific form. However, there is again a theorem from combinatorial geometry that motivates this definition. First, it is immediate to check that for an affine plane $A(n)$ with $n \geq 2$, we have
(i) any two distinct points are contained in a unique box,
(ii) if x is a point and B is a box with $x \notin B$, then there exists a unique box B' so that $x \in B'$ and $B \cap B' = \emptyset$ and
(iii) there exist 3 points which are not contained in any box.

This is clear for our first construction of an affine space $A(n)$ but this should be checked for any such affine space. The following theorem is now interesting.

Theorem 10.15. *Assume we have a set X and a collection of subsets (boxes) which satisfy (i), (ii) and (iii) above. Then this collection is $A(n)$ for some n.*

10.5 Projections of general Steiner systems

We have seen that when we have a k-transitive subgroup of S_n, we can easily obtain a $k-1$-transitive subgroup of S_{n-1} by looking at one of the stabilizers. In other words, we saw it was easy to "go down" but extending upwards was more nontrivial and certainly not always possible. We will now see an analogous "going down" and "going up" in the context of general Steiner systems.

Theorem 10.16. *If $S(n, k, t, 1)$ exists with $t \geq 2$, then $S(n-1, k-1, t-1, 1)$ exists.*

Proof: Let $(X, \mathcal{B})$ be an $S(n, k, t, 1)$ system. Fix $a \in X$ and let

$$X' = X \backslash \{a\} \quad \mathcal{B}' := \{B \backslash \{a\} : B \in \mathcal{B}, a \in B\}.$$

Clearly $|X| = n - 1$ and the boxes in $\mathcal{B}'$ all have size $k - 1$. If $Y \subseteq X'$ has $t - 1$ elements, then, since $Y \cup \{a\}$ is contained in one element of $\mathcal{B}$, Y is contained in one element of $\mathcal{B}'$.
QED

The harder question is whether a given general Steiner system $S(n, k, t, 1)$ is the projection of a $S(n+1, k+1, t+1, 1)$ general Steiner system in the sense of the above theorem. The following will provide for us a necessary condition which will be quite useful.

Theorem 10.17. *Assume that the $S(n, k, t, 1)$ system $(X, \mathcal{B})$ is a projection of a $S(n + 1, k + 1, t + 1, 1)$ system $(X', \mathcal{B}')$. Letting b, be as earlier, the number of boxes $|\mathcal{B}|$, we must have*

$$(k + 1) \mid (n + 1)b \tag{4}$$

Proof: Letting r' and b' denote the values of r and b for the larger system (so $b' := |\mathcal{B}'|$ and r' is the number of boxes in $\mathcal{B}'$ containing a fixed point in X'), we have by Theorem 10.3, applied to the extended system, that

$$(n + 1)r' = (k + 1)b'.$$

Then one observes that, by construction, $b = r'$ and the result follows.
QED

10.6 Extensions for the projective planes $P(n)$ and first connection with the Mathieu groups.

The necessary condition of Theorem 10.17 for extension reduces dramatically the number of projective planes which arise as projections.

Corollary 10.18. *If a projective plane $P(n) = S(n^2 + n + 1, n + 1, 2, 1)$ comes from a projection, then $n = 1, 2$ or 4.*

Proof: Theorem 10.17 tells us that a necessary condition is that

$$(n + 2) \mid (n^2 + n + 2)b = (n^2 + n + 2)(n^2 + n + 1) = n^4 + 2n^3 + 4n^2 + 3n + 2$$

where (3) is used for the first equality. Since

$$n^4 + 2n^3 + 4n^2 + 3n + 2 = (n + 2)(n^3 + 4n - 5) + 12$$

, we obtain that $(n + 2) \mid 12$ which gives $n = 1, 2, 4 or 10$. However $P(10)$ is known not to exist. QED

Let's now look at $P(1), P(2)$ and $P(4)$.

Exercise: Show that $P(1)$ is infinitely extendible. What is the trivial reason that lies behind this? It is sort of a degenerate situation.

Let's look at $P(2)$, the Fano plane. One observes that (4) holds since $k = 3$ and $n = 7$. Hence it is possible we can extend. In fact, one can extend obtaining a $S(8, 4, 3, 1)$. This turns out to be a special case of what is called a Hadamard 3-design which we don't discuss. Here is the system if the reader what's to check for him/her-self.

$$\mathcal{B} = \{\{1,2,5,6\}\{3,4,7,8\}\{1,3,5,7\}\{2,4,6,8\}\{1,4,5,8\}\{2,3,6,7\}\{1,2,3,4\}$$
$$\{5,6,7,8\}\{1,2,7,8\}\{3,4,5,6\}\{1,3,6,8\}\{2,4,5,7\}\{1,4,6,7\}\{2,3,5,8\}\}$$

Can we extend $P(2)$ a second time, i.e. can we extend $S(8,4,3,1)$? For this system, $k = 4$, $n = 8$ and $b = 14$. (The last term is either obtained by observing that there are 14 sets listed above or by determining it by using Theorem 10.3(2).) One now observes that (4) fails since 5 does not divide $9 \cdot 14$ and hence there is no extension by Theorem 10.17.

Let us move on to $P(4) = S(21,5,2,1)$.

Exercises:
(i) Compute b for $S(21,5,2,1)$ and verify that $S(21,5,2,1)$ satisfies the necessary condition (4) of Theorem 10.17 and hence it is possible that it can extend.
(ii) Assuming one can extend $S(21,5,2,1)$ to $S(22,6,3,1)$, determine b for this latter system and verify that $S(22,6,3,1)$ satisfies the necessary condition (4) and hence it is possible that it can extend.
(iii) Assuming one can also extend $S(22,6,3,1)$ to $S(23,7,4,1)$, determine b for this latter system and verify that $S(23,7,4,1)$ satisfies the necessary condition (4) and hence it is possible that it can extend.
(iv). Assuming one can then also extend $S(23,7,4,1)$ to $S(24,8,5,1)$, determine b for this latter system and verify that $S(24,8,5,1)$ does *not* satisfy the necessary condition (4) and hence cannot be extended.

The reader has hopefully noticed that the numbers coming up here, 21,22,23 and 24 are precisely the size of the sets that $PSL(3,4), M_{22}, M_{23}$ and M_{24} act on. This is of course *no* coincidence and this is our first suggestion of a relationship between general Steiner systems and the Matheau groups. This will be explained in more detail in Section 11.

10.7 Automorphism groups of designs

The following will be an important general concept and will relate to the Mathieu groups in the next section.

Definition 10.19. *Given a design $(X, \mathcal{B})$, we let $G_{(X,\mathcal{B})}$ be the automorphism group of the design meaning it is the subgroup of S_X consisting of those bijections from X to X such that for any $B \subseteq X$, $f(B) \in \mathcal{B}$ if and only if $B \in \mathcal{B}$.*

We state the following theorem without proof which describes the automorphism group of the Fano plane $P(2)$.

Theorem 10.20. *If $(X, \mathcal{B})$ is the Fano plane, then its automorphism group $G_{(X,\mathcal{B})}$ is isomorphic to $PSL(2, 7)$ (which is also isomorphic to $PSL(3, 2)$).*

Remark: This is the smallest simple group which is not an alternating group.

11 The Mathieu groups as automorphism groups of Steiner systems

The goal of this section is to prove the following theorem which characterizes the Mathieu groups as automorphism groups of general Steiner Systems.

Theorem 11.1. *1. There is a $S(21, 5, 2, 1)$-design whose automorphism group is $SL(3, 4)$.*
2. There is a $S(22, 6, 3, 1)$-design whose automorphism group is M_{22}.
3. There is a $S(23, 7, 4, 1)$-design whose automorphism group is M_{23}.
4. There is a $S(24, 8, 5, 1)$-design whose automorphism group is M_{24}.

It is the case there there are unique designs (up to isomorphism) for the four above general Steiner systems. We will not however prove that.

12 Various other properties of the Mathieu groups

13 References

These notes are a synthesis of material I obtained from various sources. For example, I used different parts of the following books as well as various other sources. Of course, zero of these notes are original.

Permutation Groups by J. D. Dixon and B. Mortimer,
Group theory by M. Hall,
Finite group theory by M. Isaacs,
Basic Algebra by N. Jacobson,
Permutations Groups by D. Passman,
A Course in the the Theory of Groups by D. Robinson
Introduction to Groups by J. Rotman

14 Acknowledgements

I would like to thank John Shareshian, who, in addition to being a great table tennis doubles partner many years ago and a very good friend, has more recently put up with and answered my various questions in group theory.