Construction and Simplicity of the Large Mathieu Groups

Simran Tinani
IISER Mohali

A group is called *simple* when its only normal subgroups are the trivial subgroup and the whole group. A group that is not simple can be broken into two smaller groups, a normal subgroup and the quotient group, and the process can be repeated. Simplicity of a group is a property that can, in fact be compared to the primeness of a number. There is no clear classification of infinite simple groups. However, the finite simple groups are finite in number (unlike infinite primes) and are classified as per *The Classification Theorem for Finite Simple groups*. Here we have a look at the Large Mathieu Groups, which fall into the last category, viz. Sporadic Groups, as per this theorem.

In order to construct the large Mathieu groups, we require quite a few preliminary lemmas and theorems. M_{24} is viewed as an automorphism group of the 24 Golay Code. The construction of the Golay code itself requires a thorough understanding of finite fields, vector spaces, their bases and their properties, linear transformations, matrices, the classical groups, multiple transitivity of group actions and of basic projective geometry.

Fields

A field F is a set together with two laws of composition: $+ : F\times\ F \to F$ and $\cdot : F\times\ F \to F$ called addition and multiplication respectively, satisfying the axioms :

1. Addition makes F into an Abelian group $F+$, its identity element being denoted by 0.

2. Multiplication is commutative, and it makes the set of nonzero elements of F into an Abelian group $F^\times$ whose identity element is denoted by 1.

3. Distributive law. $\forall\ a,\ b \in F,\ \ a(b+c) = ab = ac$

Theorem: Let p be a prime integer. Every nonzero congruence class modulo p has a multiplicative inverse. Therefore F_p is a field of order p.

Proof: Let a be a congruence class modulo p. Because (a,p)=1, $\exists$ integers b and c such that $ba + cp = 1 \Rightarrow ab \equiv 1(\text{modp})$. Thus the congruence class modulo p of b is the multiplicative inverse of a. Associativity, commutativity and distributivity of modulo p addition and multiplication follow from the same properties of normal integer addition and multiplication. So all the field axioms are satisfied.

The **characteristic** of a field is defined to be the order of 1 as an element of the additive group F^+, provided that the order is finite. If 1 has order infinity, the field is said to have characteristic 0.

Theorem (Prime Fields): The characteristic of any field is either zero or a prime number.

Proof: Assume a field with a composite characteristic n. Then there exist

$$a, b < n$$

such that

$$n = ab$$

The nonzero congruence classes modulo p of a and b lie in the multiplicative group, but their product $ab \equiv 0$ mod n, does not. This contradicts the field axioms. Thus n has to be prime.

Theorem (Prime Field Uniqueness): Up to isomorphism, there is only one field F_p with p elements. Proof: An isomorphism can be drawn, mapping the additive identities to each other and the (nonzero) multiplicative generators to each other.

Theorem (Primitive Elements): Given any field F_q with q elements, the nonzero elements of F_q form a multiplicative cyclic group $F_q{}^* = \{0, 1, \alpha, \alpha^2, \ldots, \alpha^{q-2}\}$. Consequently F_q has $\phi(d) \geq 1$ elements of multiplicative order d that divides $q - 1$, no elements of any other order. In particular, F_q has $\phi(q) \geq$ primitive elements. The proof of the above theorem involves the definition of polynomials over the field, of factoring of these polynomials, the definition of prime polynomials, and a description of arithmetic mod a monic

prime polynomial over the field. Using the definitions of mod-g(x) arithmetic for a prime monic polynomial g(x) of degree m, we define two operations, and subsequently show that the set $R_{F,m}$ of remainder polynomials mod g(x) actually forms a finite field of size $|\mathbf{F}|^m$, where F is the field over which the polynomials are being defined. This field is denoted by $F_{g(x)}$. Since a polynomial with degree m can have at most m degree-1 factors, the proof of *The Fundamental Theorem of Algebra* is clear. This theorem states that over any field **F**, a monic polynomial $f(x) \in$ F[x] of degree m can have no more than m roots in **F**. If it does have m roots $\{\beta_1, \beta_2, \ldots, \beta_m\}$ then the unique factorization of f(x) is

$$f(x) = (x - \beta_1)(x - \beta_2) \ldots (x - \beta_m)$$

Also every nonzero element in the field satisfies

$$x^{q-1} = 1$$

because they belong to the multiplicative group of order $q - 1$. Since the polynomial

$$x^{q-1} - 1 = 0$$

has at most q-1 roots in the field, it must have q-1 distinct roots and these are all the nonzero elements of the field. So

$$x^{q-1} - 1 = \prod_{\beta \in \mathbf{F}} (x - \beta)$$

Hence, once we have shown that in any field the multiplicative group of nonzero elements has at most one cyclic subgroup of any given order n (consisting of roots of

$$x^n - 1 = 0$$

By a simple and elegant counting of elements of **F** it can be shown that **F*** contains precisely $\phi(q-1)$ primitive elements. By the cyclic groups theorem, $\{F_q^*\}$ has at most one cyclic subgroup of each size d, where d is the order of an element $\beta \in \{F_q^*\}$ and by Lagrange's theorem, d divides $q - 1$. Also, the number of elements in a cyclic subgroup of size d having order d is $\phi(d)$. So, the number 'n' of elements in $\{F_q^*\}$ with order $q - 1$ is at most

$$n \leq \sum_{d:d|q-1, d \neq q-1} \Phi(d)$$

But,

$$q - 1 = \sum_{d:d|q-1} \Phi(d) \Rightarrow q - 1 \geq n + \Phi(q - 1)$$

Hence the number of elements with order $q-1$

$$= (q-1) - n$$

$$\geq \Phi(q-1)$$

But, F_q^* has at most $\Phi(q-1)$ elements of order $q-1$ (0 if F_q^* is not cyclic, and $\Phi(q-1)$ if F_q^* is cyclic). Thus F_q^* has $\Phi(q-1)$ elements of order $q-1$. Thus all inequalities must be satisfied with equality. Thus the number of elements with order less than q is precisely

$$\sum_{d:d|q-1, d\neq q-1} \Phi(d)$$

Hence for every divisor d of $q-1$, F_q^* has precisely $\Phi(d)$ elements of order d, or F_q^* has exactly one cyclic subgroup of order d. In particular, F_q^* is cyclic.

If a field has characteristic p, then F_q has a prime subfield F_p with p elements. Now, $0, 1, -1 \in F_p$, so $x^q - x$ can be regarded as a polynomial in $F_p[x]$. By unique factorization (the fact that the prime factorization of a polynomial is unique; proof- by contradiction, assuming a polynomial of least degree m with a non-unique factorization, one obtains a polynomial of smaller degree and a non-unique factorization) $x^q - x$ factors over F_p into a unique product of prime (monic and irreducible) polynomials $g_i(x) \in F_p[x]$. Each $g_i(x)$ is also a monic polynomial in $F_q[x]$ since each coefficient of $g_i(x)$ is an element of F_q. Hence, again by unique factorization,

$$x^q - x = \prod_{\beta \in F_q} (x - \beta) = \prod_i g_i(x)$$

and each $g_i(x)$ must be reducible over F_q. The prime polynomials in F_p $g_i(x)$ are called minimal polynomials of F_q. Also each element of F_q is a root of exactly one minimal polynomial of F_q, and this partitions the elements of F_q into disjoint sets.

Lemma: The minimal polynomial of $\beta \in F_q$ is the monic polynomial of least degree in $F_p[x]$ such that $g(\beta) = 0$. Moreover, for any $f(x) \in F_p[x]$, $f(\beta) = 0$ iff $g(x)$ divides $f(x)$.

Define the map $m_\beta : F_p[x] \to F_q$

$$m_\beta(f(x)) = f(\beta)$$

The image of this map is, by definition, the subset of elements $G_\beta \subseteq F_q$ that can be expressed as linear combinations over F_p of powers of β. It is easily

proved using the Euclidean division algorithm the equality of the two sets:

$$G_\beta = \{f(\beta) = \sum_i f_i \beta^i, f(x) \in F_p[x]\} = \{r(\beta), r(x) \in R_{F_p,m}\}$$

Also, it is found that the map drawn as a bijection also preserves addition and multiplication, i.e. it is an isomorphism. Hence, G_β is a field isomorphic to $F_g(x)$ and the correspondence is given by $r(\beta) \in G_\beta \leftrightarrow r(x) \in R_{F_p,m}$ (here $g(x)$ is the minimal polynomial of β).

Theorem: Every finite field is isomorphic to a field $F_g(x)$. In particular, every finite field has order a power of some prime p.

We have proved that every subfield generated by an element β of a finite field F_q (i.e. the field of the linear combinations of β over F_p) must be isomorphic to a field $F_g(x)$ where $g(x)$ is the minimal polynomial of β. Now every finite field contains a primitive element α. The subfield generated by α, G_α must be the whole field F_q. This proves the theorem.

Lemma: Every prime polynomial $g(x) \in F_p[x]$ of degree m divides $x^{p^m} - x$.

Consider any prime polynomial $g(x)$ of degree m in $F_p[x]$. The set $R_{F_p,m}$ with mod-g(x) arithmetic forms a field $F_g(x)$ with p^m elements. The remainder polynomial $x \in R_{F_p,m}$ is a field element $\beta \in F_{g(x)}$. Evidently, $g(\beta) = 0$, but $r(\beta) \neq 0$ if deg(r(x)) ¡ m. So g(x) is the minimal polynomial of β. Because

$$\beta^{p^m-1} = 1$$

(because β is an element of a field of size p^m), β is a root of

$$x^{p^m-1} - 1$$

which means that $g(x)$ divides x^{p^m-1} and hence $x^{p^m} - x$. (In short, for every prime polynomial of degree m, there is a field element β in $F_{g(x)}$ which has g(x) as its minimal polynomial. The result then follows from the properties discussed).

Theorem: All finite fields of the same size are isomorphic.

We have proven that every prime polynomial $g(x)$ over F_p of degree m (provided such a polynomial exists) divides $x^{p^m} - x$ (and is the minimal

polynomial of some field element β). By unique factorization, every field of size p^m includes m elements whose minimal polynomial is $g(x)$. Now choose one of these elements, say β. The subfield generated by β is isomorphic to $F_{g(x)}$, whose elements are remainder polynomials mod-g(x). But the order of $F_{g(x)}$ is p^m, so the subfield must be the whole field F_q. But, the field chosen with p^m elements was arbitrary, and so is the prime polynomial $g(x)$. Hence the theorem follows. (Note that the number of primitive elements in the field $\Phi(p^m - 1)$ is greater than or equal to m, since each of the m powers of p, p^0, p^1, ..., p^{m-1} is co-prime to $p^m - 1$).

Theorem: The polynomial $x^{p^m} - x$ factors over F_p into the product of minimal polynomials whose degrees divide m, with no repetitions. One proves first that the roots of a minimal polynomial of a field form a cyclotomic set of the form

$$\{\beta, \beta^p, \beta^{p^2}, \ldots \beta^{p^{n-1}}\}$$

where n divides m. The definition of formal derivatives (and subsequently the connection with repeated factors) facilitates the proof that $x^{p^m-1} - x$ has no repeated factors, which was earlier proved with the assumption of the existence of a field $\mathbf{F}$ with p^m elements or of a prime polynomial over F_p of degree m.

Theorem: Finite fields F_p^m exist for all prime p and m $\geq$ 1.

It is further shown that there do not exist enough prime polynomials of degree $< m$ that their product could have degree p^m, thereby proving that there must exist a prime polynomial of degree m, and hence a field of size p^m.

Vector Spaces

Definition: A vector space V over a field F is a set together with two laws of composition:

- Addition $V \times V \longrightarrow V$
- Scalar multiplication by elements of the field $F \times V \longrightarrow V$.

These laws are required to satisfy the following axioms:

- Addition makes V into a commutative group V^+ with identity denoted by 0.
- $1 * v = v$ for all v in V
- Associative law: $(ab)v = a(bv)$ for all a, b in F, v in V.
- Distributive laws: $(a + b)v = av + bv$; $a(v + w) = av + aw$ for all a, b in F and v, w in V.

A set of vectors that is linearly independent and that spans the vector space V is said the be the **basis** of V. Equivalently, if every vector in V can be expressed uniquely as a linear combination of a certain set of vectors, then this set is the basis of V. Most results about bases(Eg. their formation by the addition of elements to an independent subset of V, or by removal of elements from a subset spanning V, uniqueness of the size of the basis for V) are fairly straightforward. The existence of a basis for every vector space is, however, not so simple to prove. Clearly, an n-dimensional vector space over a field **F** is isomorphic to F^n, the set of column vectors with n entries over **F**.

Linear Transformation: A linear transformation is a map between vector spaces V and V' over a field F, Φ: $V \longrightarrow V'$ satisfying $\Phi(x + y) = \Phi(x) + \Phi(y)$ and $a\Phi(x) = \Phi(ax)$ for all x, y $\in V$ and $a \in F$.

The representing matrix of a linear transformation T: $V \longrightarrow V$ with respect to basis β is the $n \times n$ matrix with the i^{th} column as v_iT, where the basis of V is $\beta = \{v_1, v_2, \ldots v_n\}$, and right multiplication is used to denote the action of the transformation on vectors in V.

$$[T]_\beta = \begin{pmatrix} -v_1T- \\ \vdots \\ -v_nT- \end{pmatrix}$$

For a given vector $x \in V$,

$$x = x_1 v_1 + x_2 v_2 + \ldots + x_n v_n$$

for scalars $x_i \in F$. So,

$$xT = (x_1 v_1 + x_2 v_2 + \ldots + x_n v_n)T = \sum x_i (v_i T) = [T]_\beta \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

So the linear transformation can be viewed in terms of a matrix, the action on a given vector being given by normal matrix multiplication. It can easily be proved that a linear transformation is bijective iff the n vectors v_1, ..., v_n are linearly independent, or in other words if they form a basis of V. By the invertible matrix theorem, this is the same as saying that a linear transformation is bijective iff its representing matrix is invertible.

Linear Groups and the Simplicity of $PSL_3(F_4)$

We denote by $GL_n(F)$ (the General Linear Group) the group of bijective linear transformations over an n-dimensional vector space V over a field F. Thus $GL_n(F)$ can be viewed as the set of invertible $n \times n$ matrices over field F. Define $SL_n(F)$ (the Special Linear Group) as the subgroup of $GL_n(F)$ whose matrices have determinant 1.

Let E_{ij} denote a matrix with entry all entries except the $(i,j)^{th}$ as 0, and with the $(i,j)^{th}$ entry as 1. We define three types of matrices, called the elementary matrices. Left multiplication by these matrices corresponds to the elementary row transformations.

- **Transvection:** A matrix $T_{ij} = I + bE_{ij}$, where I is the identity matrix, and b is any non-zero field element. It is easy to verify that left multiplication by T_{ij} with a matrix A adds b times thei^{th} column of A to the j^{th} column, whereas right multiplication by T_{ij} with a matrix A adds b times thej^{th} row of A to the i^{th} row.

- $P_{ij} = I - E_{ij} + E_{ji} - E_{ii} - E_{jj}$ The matrix P_{ij} interchanges the negated j^{th} row of A with the i^{th} row of A. A P_{ij} matrix is invertible and $P_{ij}^{-1} = P_{ji}$

- $U_i(d) = I + (d^{-1} - 1)E_{ii} + (d-1)E_{(i+1)(i+1)}$, $d \in F^*$

Now, T_{ij}, being a triangular matrix, has determinant the product of its diagonal entries, which are all 1, and hence is in $SL_n(F)$. Moreover, it can be shown that each P_{ij} and $U_i(d)$ can be expressed as products of transvections, hence all 3 elementary matrices lie in $SL_n(F)$.

Theorem: $SL_n(F)$ **is generated by the matrices** $T_{ij}(b)$**, where** $b \in F^*$ **and** $i \neq j$**.**
Proof: We must show that if $A \in SL_n(F)$ then there exist matrices P, Q such that P, Q are products of transvections and $PA = I \Leftrightarrow A = P^{-1}$. Let $A \in SL_n(F)$. By Gaussian elimination, we must multiply A with $T_{ij}(b)$ and P_{ij} matrices to row reduce it and may now assume its reduced row Echelon form. We know that the reduced row Echelon form of an invertible matrix is I. On putting P as the product of these elementary matrices, the desired result is obtained.

Theorem: If $n \geq 3$ **or if** $n = 2$ **and** $|F| > 3$**, then** $SL_n(F)$ **is identical to its commutator subgroup** $SL_n(F)'$**.**
Proof: Let $G = SL_n(F)$. If we can show that the generators $T_{ij}(b)$ are contained in the commutator subgroup G', we are done. If $n \geq 3$, then let $i \neq j$, $k \neq i, j$, we have:

$$T_{ik}(b)T_{kj}(1)T_{ik}(b)^{-1}T_{kj}(b)^{-1} = T_{ik}(b)T_{kj}(1)T_{ik}(-b)T_{kj}(-1)$$

$= (I + bE_{ik})(I + bE_{kj})(I - bE_{ik})(I - bE_{kj}) = I + bE_{ij} = T_{ij}(b) \Rightarrow T_{ij}(b) \in G'$
Now let n = 2.

$$U_1(d)^{-1}T_{12}(c)U_1(d)T_{12}(-c) = \begin{pmatrix} 1 & c(d^2-1) \\ 0 & 1 \end{pmatrix}$$

Now $|F^*| > 2$. If ¡d¿ = F*, d $\neq 1$ and $d^2 \neq 1 \Rightarrow d^2 - 1 \neq 0 \Rightarrow (d^2-1)^{-1} \in F^* \Rightarrow c = b(d^2-1)^{-1}$. Referring to our matrix above, we see that $T_{12}(b) \in G'$. A similar result holds for $T_{21}(b)$.

Lemma: If $n \geq 3$ or if $n = 2$ and $|F| > 3$, then $SL_n(F) = GL_n(F)'$. The proof of this lemma used he first isomorphism theorem and the onto homomorphism, $det : GL_n(F) \longrightarrow F^*$. The kernel of this homomorphism is $SL_n(F)$ by definition, thus the quotient group $G :_n (F)/SL_n(F)$ is isomorphic to the Abelian group F^* and is this Abelian. So, $GL_n(F) \subseteq SL_n(F)$. The reverse inclusion is obvious, since $SL_n(F)' = SL_n(F)$.

Theorem: The centre of the $GL_n(F)$ consists of the set of scalar matrices over F^*, the same being true for $SL_n(F)$, whose centre consists of scalar matrices αI, where $\alpha^n = 1$.
The proof is simple and involves using the fact that a matrix in the centre must commute with a transvection.

We define the Projective General linear group,

$$PGL_n(F) = GL_N(F)/Z(GL_N(F)$$

and the Projective Special Linear Group,

$$PSL_n(F) = SL_N(F)/Z(SL_N(F)$$

Projective Geometry

A projective n-space over field **F**, $P^n(F)$ is defined to be the set of lines through the origin.

Lemma: For v, w $\in V^{\#}$, define $v \quad w$ if there exists an $\alpha \in$ such that $w = \alpha v$. Then is an equivalence relation on $V^{\#}$. (Here $V^{\#} = F^n - \{0\}$)

The properties of a field of identity, inverses and of associativity of scalar multiplication facilitate the proof of reflexivity, symmetry and transitivity respectively.

Then the projective (n-1)-space is defined as:

$$P^{n-1}(F) = \mathbf{\{[v]:\ v} \in V^{\#} \mathbf{\}}, \text{ where } \mathbf{[v] = \{}\alpha\mathbf{v:}\ \alpha \in F^{\times}\mathbf{\}}.$$

If there is a point $(f_0, f_1, \ldots f_n)$ on a line through the origin with $f_0 \neq 0$, then we see that $f_0 - 1(f_0, \ldots f_n) = (1, f_0^{-1}f_1, \ldots f_0^{-1}f_n)$ is the unique point on the line that it determines with 1 for its first co-ordinate (from the uniqueness of inverses). On the other hand, if 0 is the first co-ordinate for any point on the line (except origin) then 0 is the first co-ordinate for every point on the line. Lines satisfying this property are the lines through the origin in the n-dimensional subspace defined by $f_0 = 0$, the subspace being simply a copy of $P^{n-1}(F)$ with a zero glued to the front. Points constituting this space are

said to lie in the "hyperplane at infinity", H_∞. So, a projective space $P^n(F)$ can be decomposed into two parts: a copy of F^n(the affine space) and one of $P^{n-1}(F)$ (the hyperplane at infinity). Also, if one defines the set

$$U_i = \{[v_1 : \ldots : v_{n+1}] \in p^n(E) | v_i \neq 0\}$$

where the homogeneous co-ordinates of [v] are

$$[v_1 : \ldots : v_n + 1]$$

, then clearly,

$$p^n(E) = U_{n+1} \cup H_\infty$$

Now, as a base case, $\|P^0(F)\| = 1$. Hence, recursively, one may derive

$$|P^n(F)| = q^n + q^{n-1} + \ldots + q + 1$$

where q is the size of the field.

There is a clear bijection between the set of one-dimensional non-zero subspaces of a vector space over a field and between the projective space as defined above. Moreover, the projection under the map that takes every vector to its equivalence class, of a two-dimensional subspace of $V^\#$ is called a **projective plane**, and that of a one-dimensional subspace of $V^\#$ is called a **projective line**.
In general, a projective subspace PW of PV is of the form π(W\0) where π is the residue class map and W is a vector subspace of V.

It is not difficult to verify that $GL_n(F)$ acts on $P^{(n-1)}(F)$ by $[v]A = [vA]$ for $A \in GL_n(F)$ and $[v] \in P^{(n-1)}(F)$, and that the kernel of this action is the centre of $GL_n(F)$, $Z(GL_n(F))$. An analogous result is true for $SL_n(F)$.

Theorem: For n $\geq$ 2, $SL_n(F)$ acts 2-transitively on $P^{(n-1)}(F)$.
Proof: Let n $\geq$ 2. By the equivalent conditions for 2-transitivity, it suffices to show that for a distinct pair $[v_1]$ and $[v_2]$ there exists matrix A $\in SL_n(F)$ such that $[v_1]A = [e_1]$ and $[v_2]A = [e_2]$. Since $[v_1] \neq [v_2]$, i.e. $v_1 \neq \alpha\ v_2$ for any scalar α, which means the set $\{v_1, v_2\}$ is linearly independent and hence can be extended to a basis $\{v_1, \ldots, v_n\}$. Consider the change of basis matrix A', i.e. the representing matrix of the linear transformation T that maps v_i to e_i. Let det(A') = α, some non-zero scalar. Then consider A" = diag(α^{-1}, 1, ..., 1) and A = A'A". Det(A) = det(A')det(A") = 1, $v_1A = e_1A'' = \alpha^{-1}e_1$ and $v_2A = e_2A'' = e_2$. Hence $[v_1]A = [e_1]$ and $[v_2]A = [e_2]$ as required.

Lemma: For $n \geq 2$, let $(e_1, e_2, \ldots e_n)$ be the standard basis for F^n and let $G = SL_n(F)$. Then $Stab_G([e_1])$ contains a normal Abelian subgroup $A(e_1)$ whose conjugates in G generate G.
Let T $\in Stab_G([e_1])$. $[e_1]T = [e_1] \Rightarrow e_1T = \alpha e_1$ for some non-zero scalar α. So the first row of T is αe_1. The normal subgroup specified in the lemma can now be identified with the kernel of the homomorphism that maps a matrix to its submatrix obtained on deleting its first row and column. It can explicitly be shown to be Abelian, and its conjugates can be shown to generate all transvections and hence G.

We define a **block** of a set X that is acted upon by a group G to be a subset B of X such that for any g $\in$ G, either $B \cap Bg = \phi$ or $B = Bg$.
If there are no blocks other than the trivial blocks (ϕ, X and one-point subsets), then G is said to act **primitively** on X.

Lemma: Let G act transitively on X. Then G acts primitively iff Stab(x) is a maximal subgroup for each $x \in X$.
The proof is by contradiction. On assuming Stab(x)$\subset$ H $\subset$ G, one may prove that the set xH is a non-trivial block of X. On assuming that Stab(x) is maximal and that there exists a non-trivial block B, the maximality of Stab(x) is contradicted.

Lemma: If G acts primitively on a set X and normal subgroup H of G is not contained in the kernel of the action, then H acts transitively on X. In particular, G acts transitively on X as long as the action is not trivial.
H also acts on X, and thus partitions X into H-orbits xH (using the right actions notation). If $x \in X$ is fixed and $g \in G$, using the normality of H in G,

$$(xH)g = x(Hg) = (xg)H$$

Thus orbit xH is taken to orbit (xg)H by $g \in G$. Because these orbits are a partition of X, $\forall x \in X$ and $\forall x \in X$, either $(xH)g = xH$ or $(xH)g$
$capxH = \phi$. Thus the H-orbits are blocks of X. Because G acts primitively on X, the orbits are either X or single points of X. Because H is not contained in the kernel, $\exists\ h \in H$ such that $yh \neq y$ for some $y \in X \Rightarrow yH \neq \{y\}$. Thus yH = X for some $y \in X$. Hence H acts transitively on X.

Lemma: If G acts primitively on a set X and H$\subseteq$G acts transitively on X, then $G = Stab(x)H$.

Lemma: If G acts 2-transitively on X, then G acts primitively.

Assume a non-trivial block B of X. There must exist distinct x, y, z in X with x, y ∈ B and z ∈ B. Because G acts 2-transitively, ∃ g ∈ G such that xg = x and yg = z. Thus x ∈ Bg∩B → Bg∩B ≠ ϕ. Also, we have z ∈ Bg → Bg∩B ≠ B. But this contradicts the definition of a block. So B is trivial.

Lemma: Let G act on X and let K be the kernel of the action. If

- G acts primitively on X
- G = G'
- ∃ x ∈ X such that Stab(x) contains a normal abelian subgroup with the property that G is generated by the conjugates $g^{-1}a(x)g$, $g{\in}G$, $a(x){\in}A(x)$

Then, G/K is simple.

Theorem: If n ≥ 3 or if n = 2 and —F— ≥ 3, then $PSL_n(F)$ is simple.
Proof- Let n ≥ 2. Now, $SL_n(F)$ acts 2-transitively on $P^{n-1}(F)$ and the kernel of the action is $Z(SL_n(F))$. So $SL_n(F)$ must act primitively on $P^{n-1}(F)$ as per the lemma. Also, except for n=2 and —F—=2 or 3, $SL_n(F) = SL_n(F)'$. We have $A(e_1) \subset Stab(e_1) \subset SL_n(F)$ where $A(e_1)$ is the normal abelian subgroup whose conjugates generate $SL_n(F)$. Thus $SL_n(F)/Z(SL_n(F))$ is simple.

Theorem(orders of linear groups): Let F_q be a finite field. We have:

$$\#|GL_n(F_q)| = \prod_{i=0} q^n - q^i$$

Follows from the counting of the number of invertible n×n matrices, q^n - 1 choices for the first row, $q^n - q$ choices for the second, since there are q multiples of the first row, $q^n - q^2$ for the third, since there are q^2 linear combinations of the first two rows, and so on.

$$\#|SL_n(F_q)| = \prod_{i=0} q^n - q^i//q - 1$$

Follows from the fact that $SL_n(F)$ is the kernel of the determinant homomorphism, hence $GL_n(F)/SL_n(F) \simeq F^{\times}$.

$$\#|PGL_n(F_q)| = \prod_{i=0} q^n - q^i/q - 1$$

$Z(GL_n(F)) = F^{\times}I$, hence $|PGL_n(F)| \equiv |GL_n(F)|/|F^{\times}I|$

$$\#|PSL_n(F_q)| = (q^n - 1)(q^n - q)\dots(q^n - q^{n-2})q^{n-1}/d,$$

where

$$d = gcd(n, q - 1)$$

$$|Z(SL_n(F_q))| = |\{fI | f \in F^{\times}, f^n = 1 = f^{q-1} \Leftrightarrow f^d = 1\}| = d$$

where d = gcd(n, q-1).

Corollary: $PSL_2(F_2)$ **and** $PSL_2(F_3)$ **are not simple.**
Follows from the fact that their orders are respectively 6 and 12, and the smallest simple group of non-prime order is A_5 (order 60).

$|GL_3(F_4)| = (4^3 - 1)(4^3 - 4)(4^3 - 4^2) = 63 \times 60 \times 48 = 181440$
$|SL_3(F_4)| = |PGL_3(F_4)| = 181440/3 = 60480$ $|SL_3(F_4)| = 60480/3 = 20160$
(gcd(3, 4-1) = 3)

Semi-linear Groups

A bijective ring homomorphism from a field E to itself is called a **Field automorphism**.
We use the notation of right operators, i.e. for an automorphism σ, the mapping of a $\in$ E is denoted as a^{σ}.

Lemma: The automorphisms of a field E form a group, Aut(E) under the operation of function composition.

 # A field E is said to be an extension of a field F if $F \subseteq E$ and this extension field is denoted by E/F. If E is a field, then the intersection of all its subfields is called the prime subfield of E. This is the smallest subfield of E, and consists of 1 and the elements generated by it, i.e. 1, 1+1, 1+1+1, ... It is isomorphic to Z_p if E is finite and to Q if E is infinite.

Every element of $Aut(E)$ fixes the prime subfield P (follows from the fact that an automorphism must fix the identity 1).

We define the **Galois Group, Gal(E/F)** to be the automorphism group of a Galois extension field E/F, where each automorphism fixes every element of the base field F.

Then clearly, if the Galois extension field E/P has prime subfield P, then Aut(E) = Gal(E/F).

Definition: Let V, W be two vector spaces over a field F and let θ be a field automorphism of F. A transformation T: V $\longrightarrow$ W is θ-semi-linear if for all x, y $\in$ V and a $\in$ F,

$$(x+y)T = xT + yT$$

$$(ax)T = a^{\theta}T$$

Theorem: Let V = E^n be a vector space over the Galois extension field E/F. The set of all invertible semi-linear transformations from V to V forms a group, the **General Semi-linear Group, ΓL(V)**.

Outline of proof- If S,T $\in$ ΓL(V) are respectively θ and ϕ semi-linear, then ST is $\phi\theta$-semi-linear (closure). The composition of functions is associative. The identity transformation is defined as the one that maps every element to itself, and this element is 1-semi-linear. By definition the elements are invertible. If T $\in$ ΓL(V) is θ-semi-linear, then T^{-1} is θ^{-1} semi-linear

Lemma: Let V = E^n be a vector space over a Galois extension field E/F. Define f: Gal(E/F) $\rightarrow$ ΓL(V) such that

$$(x_1, x_2, \ldots x_n)f(\sigma) = (x_1^{\sigma}, \ldots, x_n^{\sigma}).$$

Then f is a one-to-one homomorphism.

Proof: It is easy to verify the properties of a semi-linear transformation for f(σ), showing that f is well-defined and that f(σ) is σ-semi-linear. Now let ϕ, σ $\in$ Gal(E/F) and let x $\in$ V. We have

$$x(f(\sigma)f(\phi)) = ((x(f(\sigma)))f(\phi) = ((x_1, \ldots x_n)f(\sigma)f(\phi)) = (x_1^{\sigma}, \ldots x_n^{\sigma})f(\phi))$$

$$= (x_1^{\sigma\phi}, \ldots x_n^{\sigma\phi}) = (x_1, \ldots x_n)f(\sigma\phi)$$

So f is a homomorphism. Now suppose that f(σ_1) = f(σ_2) $\Rightarrow$ xf(σ_1)= xf(σ_2) for all x $\in$ V. For k $\in$ E/F, we have

$(k^{\sigma_1}, 0, \ldots, 0) = (k^{\sigma_2}, 0, \ldots, 0) \Rightarrow k^{\sigma_1} = k^{\sigma_2}$. Because k is arbitrary, this means that $\sigma_1 = \sigma_2$, and thus f is one-to-one.

Theorem: Given a vector space V = E^n over a Galois extension field E/F, we have ΓL(V) = GL(V)f(Gal(E/F)).
Proof: Take T ∈ ΓL(V), where T is σ-semi-linear. Now, f(σ^{-1}) ∈ ΓL(V). One may verify that Tf(σ^{-1}) ∈ $GL(V)$. Since f is a homomorphism, (Tf(σ^{-1})f(σ)) = T(f(σ^{-1})f(σ)) = Tf($\sigma^{-}1\sigma$) = Tf(1) = T ∈ ΓL(V). So every element in ΓL(V) is the product of an element of GL(V) with an element of f(Gal(E/F)). Moreover, one may easily prove that $GL(V) \cap Gal(E/F) = 1$, and hence ΓL(V) is the semi-direct product ΓL(V) = GL(V)Gal(E/F) (semi-direct). Similarly, we find that the **Special Semi-linear group, ∑L(V)** defined as the product ∑L(V) = SL(V)Gal(E/F) is also a semi-direct product.

Theorem: Let V = E^n be a vector space over the Galois extension field E/F. If S ∈ GL(V) and the matrix of S is A, then S^σ:= $F(\sigma)^{-1}Sf(\sigma)$ ∈ GL(V) and $f(\sigma)$ and the matrix of S^σ is A^σ := $(a_{ij})^\sigma$ = (a_{ij}^σ).

Action of the Semi-linear Groups on Projective Space

Lemma: The semi-linear group $\Gamma L(E^n)$ acts on P^{n-1}(E) such that [v]T = [vT] for t ∈ ΓL(E^n) and [v] ∈ P^{n-1} (E). The special linear group ∑L(E^n) acts similarly on P^{n-1} (E).

Lemma: $Z(GL_n(F))$ is a normal subgroup of $\Gamma L_n(E^n)$ and $Z(SL_n(F))$ is a normal subgroup of $\sum L_n(E^n)$.

Lemma: The kernels of the actions of $\Gamma L_n(E^n)$ and $GL_n(E^n)$ are identical, and so are those of $\sum L_n(E^n)$ and $SL_n(E^n)$.

Proof: It suffices to show that the kernel of the action of $\Gamma L_n(E^n)$ on $P^{n-1}(E)$ is contained in $Z(GL_n(E^n))$, the kernel of the action of $GL_n(E^n)$. An element of $\Gamma L_n(E^n)$ is now represented as $\mathrm{S}f(\sigma)$ where $\mathrm{S} \in GL_n(E^n)$ and $f(\sigma) \in$ f(Gal(E/F)). Suppose such an element lies in the kernel of the action. Then,

$$\alpha e_i = e_i(Sf(\sigma)) = (e_iS)(f(\sigma)) = (e_iS)^{\phi} \Rightarrow \alpha^{\phi^{-1}} e_i = (\alpha e_i)^{\sigma^{-1}} = ((e_iS)^{\sigma})^{\sigma^{-1}} = e_iS$$

So, the i^{th} row of S is $\alpha^{\sigma^{-1}} e_i$. Hence S is a scalar matrix. Consider the action of $Sf(\sigma)$ on the row vector $(\lambda, 1, \ldots, 1) \in E^n$, where λ is a non-zero field element. When we equate the transformed row vector to an arbitrary scalar multiple of itself, we conclude that $\sigma = 1$, the identity transformation. Hence the kernel is the same as that of the action of $GL_n(E^n)$. The proof is similar for $\sum L_n(E^n)$ and $SL_n(E^n)$.

Then clearly, the **Projective semi-linear group**

$$P\Gamma L_n(E) := \Gamma L_n(E)/Z(\Gamma L_n(E))$$

and the **Projective special semi-linear group**

$$P\sum L_n(E) := \sum L_n(E)/Z(\sum L_n(E))$$

act faithfully on $P^{n-1}(E)$.

Action of $\Gamma L_n(E)$ on zero sets of Homogeneous Polynomials

A homogeneous polynomial is defined to be a polynomial in $E[x_1, \ldots, x_n]$ where each term is of the same total degree.
Lemma: If f is a homogeneous of degree d and if $\alpha \in$ E and x $\in E^n$, then

$$f(\alpha x) = \alpha^d f(x)$$

Let f: $E^n \to$ E be a homogeneous polynomial in n variables. The **zero set** of f is

$$Z(f) = \{[x] \in P^{n-1}(E) | f(x) = 0\} \subseteq P^{n-1}(E)$$

A **curve in** $P^{n-1}(E)$ is defined as the zero set of a homogeneous polynomial $f \in E[x_1, \ldots x_n]$.

Bilinear form: Let V denote a vector space of dimension n over a field E. A bilinear form is a map

$$B : V \times V \to E$$

that is linear in each variable as the other is held fixed, i.e.

$$B(x + y, z) = B(x, z) + B(y, z)$$

$$B(ax, y) = aB(x, y)$$

$$B(x, y + z) = B(x, y) + B(x, z)$$

$$B(x, ay) = aB(s, y)$$

$$\forall x, y, z \in V, \forall a \in E$$

It can be proven that
B(x,y) = xAy^T is a bilinear map. When the matrix A is I, then this map becomes the regular dot product.

A bilinear form B that satisfies $B(x, y) = B(y, x)$ for all $x, y \in V$ is called symmetric. One that satisfies B(x,x) = for all $x \in V$ is called an alternate or skew-symmetric bilinear form.

Orthogonality with respect to a bilinear form on a vector space is defined as: v is orthogonal to w if B(v,w) =0 for v,w $\in V^{\#}$. If B(v,w) = 0 $\Leftrightarrow$ B(w,v) = 0, we say that B is a reflexive bilinear form.

We define the orthogonal complement of a set S $\subseteq$ V to be the set

$$S^{\perp} = \{v \in V | B(v, w) = 0 \Leftrightarrow B(w, v) = 0, w \in S\}$$

It is easy to verify that dim($S^{\perp}$) = dim(V) - dim(S) and with this result one may prove that $S^{\perp\perp} = S$.

Theorem: The two definitions of lines in $P^2(F_4)$ given agree.
Proof: Let W be a 2D subspace of E^3. By definition, $\pi(W)$ is a projective line. Now, W is a plane through the origin. Also, the dot product is reflexive, which means that $W^{\perp\perp} = W$. Thus W is the set of x $\in E^3$ such that a·x = 0, where a is a nonzero vector orthogonal to W and · is the dot product; a = (a_1, a_2, a_3). Let f = a_1x_1 + a_2x_2 + a_3x_3. We see that $\pi(W)$ = Z(f) Conversely, let Z(f) be the zero set of the linear homogeneous polynomial $f = a_1x_1 + a_2x_2 + a_3x_3$. Because the coefficients of f are not all 0, a=$(a_1, a_2, a_3) \neq$

0. Thus Z(f) is the set of all [y] such that a·y = 0. Let $W = \{y \in V | a \cdot y = 0\}$. Dim(a) = 1 $\rightarrow$ Dim(W) = dim(V) - dim(a) = 3-1 = 2. Hence W is a 2D subspace and Z(f)=$\pi(W)$.

We define the set $Z(f)^T = \{[x]T \in P^{n-1}(E) | f(x) = 0\}$, which is the orbit of the zero set Z(f) under the action of $\Gamma L_n(E)$ already defined.

If f $\in E[x_1, x_2, \ldots x_n]$ is homogeneous of degree d, and T $\in \Gamma L_n(E)$, we define
$f^T(x) = f(xT^{-1}) \; \forall \; x \in E^n$. Note that if T has a non-trivial field automorphism (i.e. if it is not linear), then f^T is not a polynomial in $[x_1, \ldots, x_n]$.

Lemma: For T $\in \Gamma L_n$(F) and f $\in E[x_1, \ldots, x_n]$, the zero set of f^T is well-defined and equal to $(Z(f))^T$, i.e.

$$Z(f^T) = (Z(f))^T$$

Lemma: If f is homogeneous of degree f and $T = T_1\theta \in \Gamma L_n(E)$ where $T_1 \in GL_n(E)$ and $\theta \in$ Gal(E/F), then :

$$Z(f^T) = Z(f^{T_1})_\theta$$

where f_θ is the polynomial resulting from the application of θ to the coefficients of f.

Theorem: $\Gamma L_n(E)$ acts on the zero sets of homogeneous polynomials. The kernel of the action is the centre of Γ $L_n(E)$. Hence $P\Gamma L_n(E)$ also acts on the zero sets of the homogeneous polynomials.

Let S $\subseteq P^2(E)$ such that S is a point or line in $P^2(E)$. So S = $\pi(W)$ where W $\subseteq E^3$ of dimension 1 or 2, where π is a the residue class map. Define a mapping

$$\phi : P^2(E) \rightarrow P^2(E)$$

such that

$$\phi(S) = \pi(\pi^{-1}S)^{\perp} = \pi(W^{\perp})$$

Theorem: Lines and points are dual in the projective plane $P^2(E)$.
The dot product is a non-degenerate symmetric bilinear form and so

$$dimW^{\perp} = dimE^3 - dimW = 3 - dimW$$

If S is a point, W has dimension 1 and thus $W^{\perp}$ has dimension 2 and $\pi(W^{\perp}$ is a line. Similarly, if S is a line, $\pi(W^{\perp})$ is a point. So, ϕ maps lines to points and points to lines. It is easy to verify that ϕ also preserves incidence, and is bijective, with $\phi^{-1} = \phi$

Corollary: If $|E|$ = q then $P^2(E)$ has $q^2 + q + 1$ lines, and q+1 lines in $P^2(F)$ contain a given point. Since ϕ is a bijection, the number of lines is equal to the number of points, viz. $q^2 + q + 1$.

Lemma: If l: aX + bY + cZ =0, where a, b, c are field elements not all equal to 0, then the solution set of l and αl: αa + αb + αc = 0 for $\alpha \in E^{\times}$ are identical. The projection of this set is a line in $P^2(E)$.

In fact, we may define:

$$[a : b : c]^{\perp} = \{[X : Y : Z] \in P^2(E) | aX + bY + zZ = 0\}$$

and

$$\{[X : Y : Z] \in P^2(E) | aX + bY + zZ = 0\} = [a : b : c]$$

So, lines in $P^2(E)$ are the projections of 2-dimensional subspaces of E^3 which are defines by equations of the form $a_1X + a_2Y + a_3Z = 0$, such that $a_i \in$ E are not all 0.

Theorem: Two distinct lines in $P^2(E)$ intersect in exactly one point.

Theorem: Lines in $P^2(E)$ have 3 forms:

1. {[x : mx + b: 1] — m, b ∈ E} ∪ {[1:m:0] with m, b fixed as x varies}
   # when $a_2 \neq 0$

2. {[b : y: 1] — b ∈ E} ∪ {[0:1:0] with b fixed as x varies} # when $a_2 = 0$, $a_1 \neq 0$ (this includes the line at infinity, [X:Y:0])

3. {[1 : m: 0] — m ∈ E} ∪ {[0:1:0] with m fixed as x varies} # when $a_2 = 0$, $a_1 = 0$

For example, let E = F_4

$$|P^2(F_4)| = 4^2 + 4 + 1$$

Each projective line in $P^2(F_4)$ has $4 + 1 = 5$ points. By the previous theorem, we have 3 types of lines in $P^2(F_4)$. Lines of the form y = mx + b have 4 choices for m and 4 choices for b and are thus 16 in number. Lines of the form x = b have 4 choices for b, and thus 4 lines. There is one line at infinity, L_∞. These account for the 21 lines. Each line of type 1 and 2 has 4 affine points, plus one point at infinty, while the line at infinity has 4 affine slopes, plus the slope ∞.

Hexads in $P^2(F_4)$

A **k- arc** is a set of k points in $P^2(F_q)$ such that no 3 points are collinear.
If k ¡3, this is vacuously true for any set of k points.
If a point is removed from a k-arc, the remaining points form a (k-1)-arc.
For k=2,3,4,5,6, k-arcs are respectively called duads, triads, tetrads, pentads and hexads.

Lemma: If $\{v_i\} \subseteq F_q^3$ for $1 \leq i \leq k$, then the following are equivalent:

1. For any $a_i \neq 0$ in F_q, ever 3-element subset of $\{a_1v_1, \ldots a_kv_k\}$ is linearly independent.

2. Every 3-element subset of $\{v_1, \ldots v_k\}$ is linearly independent.

3. $\{[v_1], \ldots, [v_k]\}$ is a k-arc.

For $1 \Rightarrow 2$, put $a_i = 1$ For $2 \Rightarrow 3$, choose $\{v_1, \ldots v_k\}$ as a representative subset. So $v_i \neq 0$ for $i = 1, 2, 3$ and $[v_1], [v_2], [v_3]$ are the corresponding points in $P^2(F_q)$. Now any two of these vectors must form an independent set, and hence their span is a 2-dimensional subspace of F_q^3, corresponding to a line in $P^2(F_q)$. Because the third vector is not in this subspace, $[v_1], [v_2], [v_3]$ are not collinear. For $3 \Rightarrow 1$, let $[v_1], [v_2], [v_3]$ be elements of a k-arc, and take $a_i \neq 0$ for i =1, 2, 3. Because they are not collinear, any one of them is not on the line defined by the other two, and so $[v_3]$ is not on the line defined by $[v_1]$ and $[v_2]$ (v_1 and v_2 cannot be linearly dependent, for then $[v_1] = [v_2]$) This line corresponds to span $\{a_1v_1, a_2v_2\}$. Thus a_3v_3 is not in span$\{a_1v_1, a_2v_2\}$. So, span$\{a_1v_1, a_2v_2, a_3v_3\}$ is a three-dimensional subspace of F_q^2 and is hence the set $\{a_1v_1, a_2v_2, a_3v_3\}$ is linearly independent.

Lemma: The group $PGL_3(F_q)$ acts transitively on ordered triads in $P^2(F_q)$.
Proof:

$$\beta = \{e_1, e_2, e_3\}$$

is a basis and so

$$([e_1], [e_2], [e_3])$$

is an ordered triad. Let $([v_1], [v_2], [v_3])$ be another ordered triad. The matrix

$S = \begin{pmatrix} -v_1- \\ -v_2- \\ -v_3- \end{pmatrix}$ maps e_i to v_i, i= 1, 2, 3

Because $\{v_1, v_2, v_3\}$ is independent, S is invertible and thus is in an equivalence class [S] in $PGL_3(F_q)$ mapping $\{e_1, e_2, e_3\}$ to $([v_1], [v_2], [v_3])$. This

suffices for the action to be transitive.

Lemma: TFAE:

1. $\{([v_1],[v_2],[v_3],[x])\ \}$
2. $\{v_1,\ v_2\ v_3\ \}$ is an independent set, and x $= a_1v_1 + a_2v_2 + a_3v_3$ such that no a_i is 0.

Theorem: $PGL_3(F_q)$ acts sharply transitively on ordered tetrads in $P^2(F_q)$.
Let $q_1 = \{([v_1],[v_2],[v_3],[v_4])\ \}$ $q_2 = \{([w_1],[w_2],[w_3],[w_4])\ \}$ be ordered tetrads in $P^2(F_q)$. By the lemma,

$$v_4 = a_1v_4 + a_2v_2 + a_3v_3$$

$$w_4 = b_1w_4 + b_2w_2 + b_3w_3$$

such that no a_i or b_i is 0. Because $\{v_1,\ v_2\ v_3\ \}$ and $\{w_1,\ w_2\ w_3\ \}$ are independent, so are $\{a_1v_1,\ a_2v_2\ a_3v_3\ \}$ and $\{a_1w_1,\ a_2w_2\ a_3w_3\ \}$. There exists a T $\in GL_3(F_q)$ such that

$$(a_iv_i)T = b_iw_i$$

for i =1, 2, 3. (follows from the transitivity of the action on triads). Now $\{a_1v_1,\ a_2v_2\ a_3v_3\ \}$ and $\{a_1w_1,\ a_2w_2\ a_3w_3\ \}$ are bases for F_q^3. Take the matrix

$$\begin{pmatrix} -b_1w_1- \\ -b_2w_2- \\ -b_3w_3- \end{pmatrix}$$

(the i^{th} row is b_iw_i written in the basis $\{a_1v_1,\ a_2v_2\ a_3v_3\ \}$). Because these are independent, S is invertible. Now, we have $v_4T = w_4$ (by the linearity of T and its action on the bases). Hence T sends the set $\{a_1v_1,\ a_2v_2\ a_3v_3,\ v_4\ \}$ point-wise to $\{b_1w_1,\ b_2w_2\ b_3w_3,\ w_4\ \}$ Thus $[v_i]T = [v_iT] = [w_i]$ and the group $PGL_3(F_q)$ acts transitively on ordered tetrads in $P^2(F_q)$. Taking $q_0 = \{[100],[010],[001],[111]\}$ as an ordered tetrad, one find that the pointwise stabilizer is trivial in $PGL_3(F_q)$. Hence the action is sharp and transitive.

Lemma: If G acts sharply k-transitively on X, where $|X| = n$ and k $\leq$ n, then G acts faithfully on X.

Thus $PGL_3(F_q)$ acts faithfully, sharply and transitively on ordered tetrads in $P^2(F_q)$. Hence the number of ordered tetrads in $P^2(F_4)$ is equal to

$|PGL_3(F_4)| = 60480$, and the number of unordered tetrads is $60480/4! = 2520$.

From now onwards, the projective plane used is $P^2(F_4)$. The convention used to represent points in the projective plane is to maximize the number of 1's and then the number of ω's.

Lemma: Let

$$q_0 = \{[100], [010], [001], [111]\}$$

be a tetrad in $P^2(F_4)$. Then:

- q_0 is contained in two pentads in $P^2(F_4)$.
- q_0 is contained in one hexad in $P^2(F_4)$, $h_0 = q_0 \cup \{[\omega\omega 1], [\omega\varpi 1]\}$

As a corollary, every tetrad and pentad is contained in only one hexad in $P^2(F_4)$.

Lemma: $PGL_3(F_4)$ acts transitively on ordered hexads in $P^2(F_4)$.
Proof: Let h_1 and h_2 be distinct hexads in $P^2(F_4)$. Delete two points from them to get two tetrads. Now theres exists an element of $PGL_3(F_4)$ that takes one tetrad to the other, Each of these tetrads is in a unique hexad. Therefore $PGL_3(F_4)$ takes a hexad in $P^2(F_4)$ to another.

Number of ordered tetrads that can be formed within a hexad = ${}^6C_4 4! = 360$. Because each tetrad lies in a unique hexad, and there are 60480 ordered tetrads,we have 60480/360 = 168 total hexads.

Theorem: Under the action of $PSL_3(F_4)$, $P^2(F_4)$ has 3 orbits of hexads, all of size 56. Every hexad contains 360 pentads, and G=$PGL_3(F_4)$ acts sharply transitively on the set of tetrads in $P^2(F_4)$. Hence there are 360 elements in G that send a hexad h setwise to itself, and these form a subgroup $H = Stab_G(h)$ of order $360 = |A_6|$. H acts 4-transitively on the 6 points of h, just as A_6 does on {1, 2, 3, 4, 5, 6}, and hence $Stab_G(h) \simeq A_6$. Thus H is simple. Now N = $PSL_3(F_4)$ is a normal subgroup of G. So by the second isomorphism theorem, $H \cap N$ is a normal subgroup of H. So $H \cup N$ = 1 or H. But if $H \cup N = 1$, the orbit size of h turns out to be 20160 ¿ 168, a contradiction to the fact that there are only 168 hexads. Thus $H \cup N$ = H and H lies in N, and the orbit has size 20160/360 = 56.

Using the concept of "hexagrams", it may be proven that **Hexads are in the same $PSL_3(F_4)$ iff they intersect evenly.**

Theorem: Hexads intersecting evenly is an equivalence relation whose equivalence classes are $PSL_3(F_4)$-orbits. Thus hexads are in the same $PSL_3(F_4)$-orbit iff they share an even number of elements.

As a corollary, hexads in different $PSL_3(F_4)$-orbits intersect in 1 or 3 , because if they intersect in 5 points, they'd contain the same pentad. But each pentad lies in a unique hexad.

Theorem: The action of $\mathrm{G} = P\Gamma L_3(F_4)$ on $P^2(F_4)$ preserves the N-orbits of hexads. (where $\mathrm{N} = PSL_3(F_4)$, $\mathrm{G} = P\Gamma L_3(F_4)$)

Let h_1, h_2 be in the same N-orbit and let $T \in \mathrm{G}$. There must exist A in $SL_3(F_4)$ such that $h_1 A = h_2$. $SL_3(F_4)$ is a normal subgroup $\Gamma L_3(F_4)$. Thus A' $= T^{-1}AT \in SL_3(F_4)$. We have $h_2T = (h_1A)T = h_1(AT) = h_1(TA') = (h_1T)A'$. By definition, h_1T and h_2T are in the same N-orbit. Thus $\mathrm{T} \in \mathrm{G}$ preserves the N-orbits of hexads.

Binary Linear Codes

A binary linear code is a k-dimensional subspace C of F_2^n and it is referred to as an (n,k)-code. The vectors of C are called code words.

Self-Orthogonal and Self-dual Codes
If C is an (n,k)-code, then $C^\perp$ is an (n, n-k)-code.
$C^{\perp\perp}$ = C

C is called self-orthogonal if $C^\perp \subseteq$ C, and this also means that k $\leq$ n - k $\Rightarrow$ k $\leq$ n/2.

Following are a few results on (n,k)-codes that are rather straightforward to prove.

If C is self-orthogonal, then C is even (has even weight).
A binary code is doubly even if all its code words have weight divisible by 4.
Let C be a binary (n,k)-code. If the vectors of a generating set S have weight divisible by 4 and are pairwise orthogonal, then C is doubly even.
An (n,k)-code C is called self-dual if C = $C^\perp$.
If C is a self-dual (n,k)-code then n is even, 1 $\in$ C and the weight distribution of C is symmetric.
Let C be a binary (n, k)-code. An automorphism of C is an element of S_n that sends code words to code words:

$$Aut(C) = \{\pi \in S_n | c\pi \in C \forall c \in C\}$$

The Large Mathieu Groups

Action of $P\Gamma L_3(F_4)$ on $PSL_3(F_4)$: Orbits of Hexads

Let G = $P\Gamma L_3(F_4)$ and N = $PSL_3(F_4)$
Let

$$h_0 = \{[100], [010], [001], [111], [\omega\varpi 1], [\varpi\omega 1]\}$$

be a representative hexad for an N-orbit (designated orbit I); let

$$A = \begin{pmatrix} \omega & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Then,

$$h_0' = h_0 A = \{[100], [010], [001], [\omega 11], [1\omega 1], [\varpi\varpi 1]\}$$
$$h_0'' = h_0 A^2 = \{[100], [010], [001], [\varpi 11], [1\varpi 1], [\omega\omega 1]\}$$

are representative hexads of the other two orbits (II and III).

Theorem: G acts on hexad orbits I, I, III.
Proof: The action of G on $P^2(F_4)$ preserves the N-orbits, thus T $\in$ sends hexad orbits to hexad orbits and I $\in$ G sends every orbit to itself. If x $\in$ {I, II, III} and $T_1, T_2 \in$ G, then $(xT_1)T_2 = x(T_1T_2)$. Since x is a subset of $P^2(F_4)$, the action of G on {I, II, III} is the same as the one defined by the action on $P^2(F_4)$.

Lemma: N is a subgroup of the kernel of the action of G on hexad orbits I, II, III, inducing an action G/N on the points I, II, III.
I, II, III are N-orbits, and T $\in$ N sends each of I, II, III to itself. Now the kernel of the action is the subgroup of G fixing I, II, III pointwise and so N is contained in the kernel. Because N is in the kernel, and N is a normal subgroup of G, the induced action of G/N is well-defined.

In fact, one finds that $|G/N| = 6$, and on examining the elements and their action on {I. II. III}, G/N $\simeq S_3$, and the kernel of the action of G/N on this set is trivial, which means that the kernel of the action of G must be N.

The Golay Code C_{24}

Let $X = P^2(F_4) \cup \{I, II, III\}$ (X contains 24 elements).
To each of these 24 elements, we assign a particular number between 1 and

24, with no repetitions. Then a subset of these 24 elements can be associated to a unique vector in F_2^{24}. Suppose the elements of the subset correspond to the numbers i, j, k, ..., m, then the subset containing these points corresponds to the vector $e_i + e_j + \ldots + e_m$.

If l is a line in $X = P^2(F_4)$, then $l \cup \{I, II, III\}$ is called a line octad. If h is a hexad in $X = P^2(F_4)$, then $h \cup \{II, III\}$, if h $\in$ I / $h \cup \{I, III\}$, if h $\in$ II / $h \cup \{I, II\}$, if h $\in$ III, is called a line octad.

We define the Golay code, $C_{24} \in F_2^{24}$ to be the code generated by the 21 line octads and the 168 oval octads.

Theorem: The line octads obtained from the from the following lines of $P^(F_4)$:

$$\{[001]^\perp, [100]^\perp, [101]^\perp, [10\omega]^\perp, [\omega 01]^\perp, [010]^\perp, [011]^\perp, [01\omega]^\perp, [110]^\perp, [\omega 10]^\perp\}$$

form a linearly independent set of 10 vectors. If the oval octads obtained from the following hexads

$$h_0 = \{[100], [010], [001], [111], [\omega\varpi 1], [\varpi\omega 1]\}$$

and

$$h_0' = \{[100], [010], [001], [\omega 11], [1\omega 1], [\varpi\varpi 1]\}$$

are added, they form a linearly independent set of 12 vectors. (This can be verified directly using the map of the oval octads and line octads to F_2^{24}.

Hence $\dim(C_24) \geq 12$.

Theorem: The Golay code is a doubly even code.

We know that in characteristic two, if we denote by $v_i \cap v_j$ the vector that has 1 in every co-ordinate where both v_i and v_j have 1 and 0 in every other co-ordinate, then $v_i \cdot v_j = 0$ iff $|v_i \cap v_j|$ is even. We know also that two lines intersect in exactly one point. So two line octads must intersect in 1 + 3 = 4 points. Also, two hexads in the same orbit intersect in 2 or 4 points, and in different orbits in 1 or 3 points. So the intersection of two oval octads is always in 2+2=4 or 4+2=6 points if they are in the same orbit, and in 1+1 =2 or 3+1 =4 if they are in different orbits.

Now, the Golay code being doubly even must also be self-orthogonal, and so $\dim(C_{24}) \leq 24/2 = 12$. But we have proven that $\dim(C_{24}) \geq 12$. Hence $\dim(C_(24)) = 12$, and $C_{24}^\perp = C_{24}$ and C_{24} is self-dual.

We define the Mathieu Group, M_{24} to be the automorphism group of the Golay code C_{24}.

Theorem: $P\Gamma L_3(F_4)$ is a subgroup of M_{24}.
It can be verified that $P\Gamma L_3(F_4)$ maos line octads to line octads and oval octads to oval octads.

Now, there is a particular co-ordinate transformation π = (12)(34)(9 10)(11 12)(17 19)(18 20)(21 24)(22 23) (where the numbers denote points in ¶$^2(F_4)$ and the corresponding vector in F_2^{24} to i is e_i) that lies in M_24 but not in $P\Gamma L_3(F_4)$.

Define G_{24} = ¡$P\Gamma L_3(F_4)$, π¿.

Theorem: G_{24} acts 5-transitively on X = $P^2(F_4)\cup\{I, II, III\}$, and hence so does M_{24}
We know that $PSL_4(F_4)$ is a normal subgroup of $P\Gamma L_3(F_4)$ and acts 2-transitively on $P^2(F_4)$ and fixes I, II, III pointwise. Also, $P\Gamma L_3(F_4)/PSL_3(F_4)$ is isomorphic to S_3. It suffices to prove that for a, b, c, x, y in X, T in G_{24}, it is possible for T to map a to I, b to II, c to III, x to [010] and y to [100]. Then by 2-transitivity, any 5-tuple and me mapped to this 5-tuple.

Corollary: C_{24} has minimum weight d=8 with no code word of weight 4.
Corollary: Any five points of X lie in exactly one octad of C_{24}. **Corollary:** The octads of C_{24} form a Steiner system S(5, 8, 24).

Theorem: M_24 = Aut(S(5, 8, 24))

Theorem: In G_{24} and M_{24}, the set-wise stabilizer of 3 points is isomorphic to $P\Gamma L_3(F_4)$ and the point-wise stabilizer to $PSL_3(F_4)$.
This follows from the fact that 5-transitivity implies 3-transitivity.

Theorem: G_{24} and M_24 are both groups of order 244823040. This is shown using a theorem on k-transitivity, $|G|$=n(n-1)...(n-k+1)$|Stab(x_1, \ldots, x_k)|$. Hence they are equal.

If one defines M_{24-i} = Stab($x_1, \ldots, x_i$) where 0 $\leq$ i $\leq$ 4. Because M_{24} is 5-transitive, the choice of x_i does not matter.

Theorem: M_{24-i} is a (5-i)-transitive group on (24-i) points for 0 $\leq$ i $\leq$ 4
Follows from the fact that it is a stabilizer on i points.

Theorem: M_{22}, M_{23}, $\in M_24$ are simple groups. Now, M_{21} is a 3-point stabilizer and is hence isomorphic to $PSL_3(F_4)$, and thus has to be simple

as per previous result. A few other simple lemmas on multiple transitivity prove that M_{22}, M_{23}, $\in M_24$ are simple groups.

Theorem: M_{24} **is a simple group.**
Proof: Because M_{23} is a subgroup of M_24 and M_{23} is a stabilizer of one of the 24 points acted on by M_{24}. Because M_{23} is simple, and M_{24} acts 5-transitively on these points, by a lemma on multiple transitively, M_{24} must be simple.

REFERENCES

- Construction and Simplicity of the large Mathieu groups - Robert Peter Hanson, San Jose State University
- MIT OpenCourseware