On Simplicity of the Projective Special Linear Groups

Filip Nygren

2018-01-17

Populärvetenskaplig sammanfattning

Den form av matematik som de flesta av oss är vana vid handlar om heltal, negativa tal, bråktal och irrationella tal som pi och roten ur 2, samt de fyra räknesätten. Dessa tal kallas för de reella talen och de fyra räknesätten är egentligen bara två räknesätt. Den matematiska strukturen ovan kallas för elementär algebra. Men kan vi skapa andra sådana strukturer? Svaret är ja, och området där man studerar olika algebraiska strukturer kallas abstrakt algebra. En viss typ av struktur kallas för grupp, och har endast ett räknesätt, som brukar kallas för operation inom matematiken. Grupper är användbara för att beskriva olika situationer och system i naturen och i samhället, och de studeras inom gruppteori. En algebraisk struktur kan byggas upp av matematiska objekt som inte är tal. Inom linjär algebra studerar man istället vektorer och matriser, som kan skrivas som tabeller av tal. Dessa vektorer och matriser har visat sig vara, förutom inom matematiken, väldigt användbara bland annat inom geometri och fysik.

I detta arbete kommer vi arbeta med grupper som består av matriser. Dessa matrisgrupper har ett ändligt antal element, och kallas därför ändliga grupper. Om en grupp är ett slutet matematiskt system med en operation på en mängd element, kan det finnas mindre grupper i en grupp? Ja, och dessa kallas för undergrupper. En viss typ av undergrupp kallas för normal undergrupp, och dessa är väldigt viktiga i detta arbete.

För heltal så har primtalen en speciell roll. De fungerar nämligen som byggstenar för heltalen. Inom gruppteorin verkar det som att vissa grupper har en liknande roll. Dessa kallas för enkla grupper, och de kännetecknas av att de inte har några normala undergrupper.

Om vi kan hitta alla enkla grupper, som är byggstenarna för alla ändliga grupper, då vet vi väldigt mycket om ändliga grupper. Så började arbetet med att hitta alla enkla grupper

Grupperna som studeras i detta arbete är en familj av matrisgrupper som heter Projective Special Linear Groups, eller PSL. Det visade sig att nästan alla dessa grupper i denna familj är enkla, och familjen har därför varit viktig i arbetet att hitta alla ändliga grupper.

Dessa grupper har också visat sig användbara inom något som kallas projektiv geometri, en ny typ av geometri som studeras inom matematiken.

Abstract

In this thesis, we will be proving that the Projective special linear group $PSL(m, K)$ is a simple group for all dimensions m and finite fields K, with a few exceptions. First, we prove simplicity of $PSL(2, K)$ (Jordan-Moore), and then move on to proving simplicity of $PSL(m, K)$ (Jordan-Dickson) for all dimensions $m \geqslant 3$. The most important tool used to prove this is by using certain linear transformations known as (elementary) transvections.

Contents

1 Introduction

As the title of this thesis suggest, we will be dealing with the Projective special linear group, denoted PSL. This is in fact not only one group, but a whole family of groups. To find a particular Projective special linear group, denoted $PSL(m, K)$, we need to fix a dimension $m \geqslant 2$ and a finite field K. But to understand what $PSL(m, K)$ is, we need to use some concepts from linear algebra, finite fields and group theory. Therefore, in Section 2 we will review some basics in these topics. We will also be using these to create some more specific tools and results needed for the rest of the thesis.

In Section 3, we will use our theory, expanding it where needed, with the aim of proving that $PSL(2, K)$ is a simple group for all finite fields K, otherwise known as Jordan-Moore. In this chapter, we will mainly be considering $PSL(2, K)$ as a group of 2×2 matrices with elements from a finite field K.

In Section 4, we begin the work of proving that $PSL(m, K)$ is a simple group for all dimensions $m \geqslant 3$ and finite fields K. But here we will prefer to work with a general m-dimensional vector space V over a finite field K. In fact, we can choose if we want to work with $PSL(m, K)$ or $PSL(V)$, since they are isomorphic to each other.

2 Background

It should be known to the reader that a field is a commutative ring with identity, where each non-zero element has a multiplicative inverse in the field. If the field has a finite number of elements, we call it a finite field, and the number of elements in the field is called the order of a field.

Definition 2.1. *The number of elements of a field K is called the* ***order*** *of K, denoted $|K|$.*

It is known that the order of a finite field must be a prime power, i.e. $|K| = p^n = q$ for some prime p and $n \geqslant 1$. Throughout this text, K will always denote a finite field with order $p^n = q$ unless it is stated otherwise.

Now we will remind the reader of what a group is.

Definition 2.2. *A* ***group*** *$G = (S, \cdot)$ is a set S of elements and an operation $(\cdot)$ on the elements of S which satisfies $(i) - (iv)$:*

(i) If $a, b \in G$ then $a \cdot b \in G$ for all $a, b \in G$.

(ii) There exist an element $e \in G$, called the ***identity*** *of G, that satisfies $a \cdot e = e \cdot a = a$ for all $a \in G$.*

(iii) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ for all $a, b, c \in G$.

(iv) For all $a \in G$ there exist a unique element $a^{-1} \in G$ such that $a \cdot a^{-1} = a^{-1} \cdot a = e$.

The number of elements of a group G is called the ***order*** *of a group, denoted $|G|$.*

It is customary to omit the $(\cdot)$, and simply write the product of a and b as ab instead of $a \cdot b$. If we have a field K with $p^n = q$ elements, then the elements of K form a group under the multiplication operation of K if we remove the element 0_K. This group is denoted $K^\times$ and has $q - 1$ elements. Now, it could happen that a certain subset H of a group also satisfy all the group axioms. This is then known as a subgroup.

Definition 2.3. *If H is a subset of a group G, and H satisifes $(i) - (iv)$, then H is called a* ***subgroup*** *of G, denoted $H \leqslant G$.*

Certain groups have special properties. One of the most important of these groups in this text is the normal subgroup, which is defined below.

Definition 2.4. *Let H be a subgroup of G. Then H is called a* ***normal*** *subgroup in G if*

$$ghg^{-1} \in H$$

for all $g \in G$, $h \in H$.

This definition tells us that if H is a normal subgroup in G, then $ghg^{-1} = h_1$ for some $h_1 \in H$. Multiplying this equation by g from the right gives us $gh = h_1 g$. An equivalent definition of a normal subgroup is that the left and right cosets are equal, i.e. $gH = Hg$. This means that the set of all elements we get

if we multiply an element $g \in G$ by all the elements of H from the right, is the same set we would get it we would multiply from the left instead. Sometimes we will use this equivalent definition instead. A group G will always have at least two normal subgroups, the group $\{e\}$ consisting only of the identity element, and G itself. These are called the **trivial** normal subgroups. Next, we have another important group, the simple group.

Definition 2.5. *A group G is called a **simple** group if it does not contain any non-trivial normal subgroups.*

Now that we have defined normal subgroups, we can introduce the concept of quotient groups. If H is a normal subgroup in G, then we can form the quotient group, denoted G/H. Before we define quotient groups, however, we must define what a coset is. If H is a subgroup of G, then the (right) **coset** Ha of H in G is the set $\{ha : h \in H\}$. We can think of this as the congruence class of a modulo H.

Definition 2.6. *Let H be a normal subgroup in G. Then the **quotient group** G/H is the group of all right cosets of H in G.*

Less important, but needed at certain points, is the commutator subgroup G' of G.

Definition 2.7. *The **commutator subgroup** G' of a group G is the subgroup of G containing all elements of the form $aba^{-1}b^{-1}$ for all $a, b \in G$, as well as all products of such elements.*

As opposed to fields, groups are not commutative in general. A commutative group, i.e. a group where all elements commute with each other, is known as an **Abelian** group. However, it could be that certain elements of a group commute with all elements of the group. We know that the set of such elements is non-empty, since by the group axioms, e commutes with all elements of a group.

Definition 2.8. *The **center** of a group G, denoted $Z(G)$, is the subset of elements of G that commute with all elements of G.*

It can be shown that the center of a group is in fact always a normal subgroup. Next we have the centralizer of a subset of a group.

Definition 2.9. *Let G be a group and A be a subset of G. The **centralizer** $C_G(A)$ is the subgroup of G containing all the elements of G that commute with all the elements of A.*

We will not prove that the commutator subgroup and the centralizer subgroup actually are subgroups.

Now, the groups that we will be dealing with in this text are groups consisting of $m \times m$ matrices with matrix multiplication or linear transformations on an m-dimensional vector space V over a finite field K. Luckily, these two types of groups are isomorphic, so we can choose which one to work with. Matrices and linear transformations do not generally have inverses, so in order to make groups of them, we can only work with invertible matrices and linear transformations.

Definition 2.10. *The* ***General linear group*** $GL(V)$ *is the group of all invertible linear transformations on an* m*-dimensional vector space* V *over a field* K*, with function composition* $(\circ)$ *as operation.*

The General linear group $GL(m, K)$ *is the group of all invertible* $m \times m$ *matrices with elements from the field* K *with matrix multiplication as operation.*

Definition 2.11. *The* ***Special linear group*** $SL(V)$ *is the group of all invertible linear transformations on an* m*-dimensional vector space* V *over a field* K *whose associated matrices have determinant 1, with function composition* $(\circ)$ *as operation.*

The Special linear group $SL(m, K)$ *is the group of all invertible* $m \times m$ *matrices with elements from the field* K *whose determinants are equal to 1, with matrix multiplication as operation.*

Again, we will not prove that these actually are groups. Now we are ready to define the Projective special linear group PSL. The center of SL is denoted by SZ. Since the center of a group is a normal subgroup, it is natural to form the quotient group SL/SZ.

Definition 2.12. *The* ***Projective special linear group****, denoted* $PSL(m, K)$*, is the group* $SL(m, K)/SZ(m, K)$*.*

In the same way, it is possible to create the Projective general linear group $PGL = GL/GZ$, but this is not included in this text. As it turns out, the group PSL has interesting properties, which gave rise to this topic of study.

Now we have defined what we need concerning groups, and we are now ready to define a very important tool, the (elementary) transvections.

Definition 2.13. *An* ***elementary transvection*** $B_{ij}(\lambda)$*, where* $i \neq j$*, is the matrix obtained from the identity matrix by inserting* $\lambda \in K^\times$ *in the* (i, j)*-position.*

Elementary transvections have determinant 1. To see this, note that we can always find a row or column consisting of a 1 and the rest zeros, and by finding the determinant using the cofactor expansion method, we always end up with determinant 1. Thus, if $B_{ij}(\lambda)$ is an elementary transvection, then $B_{ij}(\lambda) \in SL(m, K)$.

Definition 2.14. *A* ***transvection*** T *is a matrix that is conjugate to an elementary transvection, i.e.* $T = SB_{ij}(\lambda)S^{-1}$ *for some* $S \in GL(m, K)$*.*

Transvections also have determinant 1. By the rule for determinants, we see that $\det T = \det(SB_{ij}(\lambda)S^{-1}) = \det S \det B_{ij}(\lambda) \det S^{-1} = \det S(\det S)^{-1} = 1$.

Remark: Note that multiplying a matrix $A \in GL(m, K)$ from the left by an elementary transvection $B_{ij}(\lambda)$ adds λ times row j to row i. For example, we see that $B_{12}(\lambda)$ adds λ times row 2 onto row 1:

$$B_{12}(\lambda)A = \begin{bmatrix} 1 & \lambda & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} a + \lambda d & b + \lambda e & c + \lambda f \\ d & e & f \\ g & h & i \end{bmatrix}.$$

Note also that the inverse of an elementary transvection is another elementary transvection. In particular, $B_{ij}(\lambda)^{-1} = B_{ij}(-\lambda)$:

$$B_{12}(\lambda)B_{12}(-\lambda) = \begin{bmatrix} 1 & \lambda & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -\lambda & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & \lambda - \lambda & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Now we will state and prove our first lemma.

Lemma 2.15. *Let $A \in GL(m, K)$. Denote $\mu = \det(A)$, and*

$$D(\mu) = \operatorname{diag}\{1, \ldots, 1, \mu\}.$$

Then $A = UD(\mu)$, where U is a product of elementary transvections.

Proof. We start out with a general matrix

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1m} \\ \alpha_{21} & \alpha_{22} & \ldots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \ldots & \ldots & \alpha_{mm} \end{bmatrix} \in GL(m, K).$$

Since $A \in GL(m, K)$, at least one element in each column must be non-zero. By the above remark, multiplying from the left by an elementary transvection will add a multiple of a row to another row, and doing this will not change the determinant. First, we wish to make $\alpha_{11} = 1$. By the following discussion, we will see that this can always be done.

If $\alpha_{11} = 1$, we are done. If $\alpha_{11} \neq 1$ and there exist a j such that $\alpha_{j1} \neq 0$, then we can make $\alpha_{11} = 1$ by adding a suitable multiple of row j to row 1. If $\alpha_{11} \neq 1$ and α_{11} is the only non-zero entry in the first column, we can add the first row to some other row j and then add a suitable multiple of row j to row 1, making $\alpha_{11} = 1$.

If $\alpha_{11} = 0$, then we can add a suitable multiple of some other row j to row 1, where $\alpha_{j1} \neq 0$, making $\alpha_{11} = 1$. Hence, A can be transformed to

$$\begin{bmatrix} 1 & \alpha_{12} & \ldots & \alpha_{1m} \\ \alpha_{21} & \alpha_{22} & \ldots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \ldots & \alpha_{mm} \end{bmatrix}.$$

With $\alpha_{11} = 1$, we can add suitable multiples of row 1 to the other rows, making $\alpha_{21}, \alpha_{31}, \ldots, \alpha_{m1}$ all zero, resulting in the matrix

$$\begin{bmatrix} 1 & \alpha_{12} & \ldots & \alpha_{1m} \\ 0 & \alpha_{22} & \ldots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & \ldots & \alpha_{mm} \end{bmatrix}.$$

Now we leave column 1 and focus on column 2. By applying the above procedure to rows 2 to m, it is clear that we can transform our matrix to a matrix of the form

$$\begin{bmatrix} 1 & \alpha_{12} & \alpha_{13} & \dots & \alpha_{1m} \\ 0 & 1 & \alpha_{23} & \dots & \vdots \\ 0 & 0 & \alpha_{33} & \dots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \alpha_{m3} & \dots & \alpha_{mm} \end{bmatrix}.$$

We can also make $\alpha_{12} = 0$ by adding a suitable multiple of row 2 to row 1. By repeatedly applying the above steps, we eventually arrive at a matrix of the form

$$\begin{bmatrix} 1 & 0 & \dots & 0 & \alpha_{1m} \\ 0 & 1 & \dots & 0 & \alpha_{2m} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \dots & 1 & \alpha_{(m-1)m} \\ 0 & 0 & \dots & 0 & \alpha_{mm} \end{bmatrix}.$$

Since $A \in GL(m, K)$, it has a non-zero determinant, and so, at least one entry in column m is non-zero. Thus, we can assume that α_{mm} is non-zero. By adding suitable multiples of row m to the other rows, we can make $\alpha_{1m}, \alpha_{2m}, \dots, \alpha_{(m-1)m}$ all zero, finally arriving at a matrix of the form

$$\begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & \dots & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & \alpha_{mm} \end{bmatrix}.$$

Since multiplying a matrix with unimodular matrices will not change its determinant, the above matrix has determinant μ. Now, by the remark, adding a multiple of a row to some other row in some matrix $B \in GL(m, K)$ can be done by multiplying B from the left by an elementary transvection, and we see that we have

$$PA = D(\mu)$$

where P is a product of elementary transvections. Since elementary transvections are invertible, we now multiply by P^{-1} from the left on both sides, which gives

$$A = P^{-1}D(\mu).$$

By the remark, the inverse of an elementary transvection is another elementary transvection, so this concludes the proof.

□

Lemma 2.16.

(i) Each $T \in GL(m, K)$ can be uniquely decomposed as a product

$$T = UD(\mu)$$

where $U \in SL(m, K)$ and $\mu \neq 0$.

(ii) $SL(m, K)$ is generated by elementary transvections.

Proof. First we prove (i). By Lemma 2.15, each $T \in GL(m, K)$ has a decomposition $T = UD(\mu)$, where U is a product of elementary transvections. Since every elementary transvection has determinant 1, then products of elementary transvections also have determinant 1, and thus, $U \in SL(m, K)$. Now, since $U \in SL(m, K)$ and $T = UD(\mu)$, we get

$$\det(T) = \det(U) \cdot \det D(\mu) = 1 \cdot \mu = \mu.$$

Since T was given, we see that μ is unique, and thus, $D(\mu)$ is also unique. By multiplying both sides by $D(\mu)^{-1} = D(-\mu)$ we get

$$U = TD(\mu)^{-1}.$$

Since $D(\mu)$ is unique for a given T, then so is $D(\mu)^{-1}$. It follows that U is unique because $U = TD(\mu)^{-1}$ for a given T and $D(\mu)^{-1}$ is unique. Thus, T has a unique decomposition $T = UD(\mu)$.

Now we prove (ii). Again, by Lemma 2.15, if $A \in GL(m, K)$ and $\det(A) = \mu$, then $A = UD(\mu)$, where U is a product of elementary transvections. Since $SL(m, K)$ is a subgroup of $GL(m, K)$, the previous statement also applies to $B \in SL(m, K)$, i.e.

$$B = U_1 D(\mu_1),$$

where U_1 is a product of elementary transvections. Taking the determinant on both sides we get

$$1 = \det(B) = \det(U_1) \cdot \det D(\mu_1) = 1 \cdot \mu_1$$

and so $\mu_1 = 1$ and $D(\mu_1)$ is the identity matrix. Hence, $B = U_1$ for any given $B \in SL(m, K)$ and we see that $SL(m, K)$ is generated by elementary transvections. □

Before we state and prove the next theorem, we must introduce some notation. First, let V be an m-dimensional vector space over a field K. Let $Z(V)$ denote the subgroup of $GL(V)$ consisting of all scalar transformations, and let $SZ(V)$ denote the subgroup of $Z(V)$ consisting of all scalar transformations with determinant 1.

Also, we will claim without proof that $GL(V) \cong GL(m, K)$ and $SL(V) \cong SL(m, K)$. One can see that this follows due to the invertibility of the elements in these groups.

Now, let $Z(m, K)$ denote the subgroup of $GL(m, K)$ consisting of all $m \times m$ scalar matrices αI with $\alpha \in K^{\times}$, and let $SZ(m, K)$ denote the subgroup of

$Z(m,K)$ consisting of all scalar matrices αI with $\alpha^m = 1$. In fact, we have $Z(m,K) \cong Z(V)$ and $SZ(m,K) \cong SZ(V)$.

It can be shown that these are in fact subgroups.

Theorem 2.17. *Let V be an m-dimensional vector space over a field K.*
(i) $Z(V)$ is the center of $GL(V)$.
(ii) $SZ(m,K)$ is the center of $SL(m,K)$.

Proof. We begin by proving (i). Let $T \in GL(V)$ be a non-scalar linear transformation. We want to start by showing that we can find a $v \in V$ such that $\{v, Tv\}$ are linearly independent, i.e. v is not a multiple of Tv. Let $\{v_1, \ldots, v_m\}$ be a basis for V. We start by letting $v = v_1$ and applying T to v_1. If Tv_1 is not a multiple of v_1, we are done, so assume Tv_1 is a multiple of v_1, say, $Tv_1 = \alpha_1 v_1$. We then proceed to do the same with the second basis vector v_2.

If we are unlucky, this happens for all basis vectors, in which case we would have $Tv_1 = \alpha_1 v_1, \ldots, Tv_m = \alpha_m v_m$. But since T was a non-scalar linear transformation, we must have $\alpha_i \neq \alpha_j$ for some i, j. Now we apply T to $v_i + v_j$, and we get

$$T(v_i + v_j) = \alpha_i v_i + \alpha_j v_j.$$

Since $\alpha_i \neq \alpha_j$ then $T(v_i + v_j)$ is not a multiple of $v_i + v_j$ and we have found a vector in V such that $\{v, Tv\}$ are linearly independent. Having found the set $\{v, Tv\}$ of two linearly independent vectors, we extend this set with the vectors $u_3, \ldots, u_m$, such that $\{v, Tv, u_3, \ldots, u_m\}$ is linearly independent, and thus forms a basis for V.

Now, since $\{v, Tv, u_3, \ldots, u_m\}$ forms a basis for V, we can in fact show that $\{v, v + Tv, u_3, \ldots, u_m\}$ also forms a basis for V. To see this, note that $Tv = (v + Tv - v) \in span\{v, v + Tv, u_3, \ldots, u_m\}$, and thus,
$span\{v, v + Tv, u_3, \ldots, u_m\} = span\{v, Tv, u_3, \ldots, u_m\}$. Since both these sets of vectors span V, they are both bases for V.

Let S be the linear transformation such that $Sv = v, Su_i = u_i$ for all $i \geqslant 3$ and $ST(v) = v + Tv$. This gives us that $TS(v) = Tv$ and $ST(v) = v + Tv$, and thus $TS(v) \neq ST(v)$, i.e. S and T do not commute and T cannot be in the center of $GL(V)$. Since T was chosen to be any non-scalar linear transformation, the center or $GL(V)$ can only contain scalar linear transformations.

We will now show that scalar linear transformations commute with all elements of $GL(V)$. Let $M = \alpha 1_V \in GL(V)$ and $N \in GL(V)$. Then we get

$$(\alpha 1_V \circ N)(v) = (\alpha 1_V)(N(v)) = \alpha(1_V)(N(v)) = \alpha N(v)$$

and

$$(N \circ \alpha 1_V)(v) = N((\alpha 1_V)v) = N(\alpha v) = \alpha N(v).$$

Thus, the center of $GL(V)$ is $Z(V)$.

Now we prove (ii). Let $R \in SL(V)$ be a non-scalar linear transformation with determinant 1, and let the linear transformation S be constructed in the same way as in (i). In fact, we have that $S = B_{12}(1)$ relative to the basis $\{v, Tv, u_3, \ldots, u_m\}$. Thus, $\det(S) = 1$ and $S \in SL(V)$.

By the same argument as in (i), R is not in the center of $SL(V)$, so the center of $SL(V)$ can only consist of scalar transformations αI. The center of $SL(V)$ is a subgroup of $SL(V)$, so the elements in the center must have determinant 1. Thus, $\det(\alpha I) = \alpha^m$ gives that $\alpha^m = 1$, and we see that the center of $SL(m, K)$ is $SZ(m, K)$.

□

We now proceed to find the order of these groups.

Theorem 2.18. *Let K be the field with $q = p^n$ elements, $m \geqslant 2$ and $d = (m, q-1)$.*

(i) The order of $GL(m, K)$ is $(q^m - 1)(q^m - q) \cdots (q^m - q^{m-1})$.
(ii) The order of $SL(m, K)$ is $(q^m - 1)(q^m - q) \cdots (q^m - q^{m-2})q^{m-1}$.
(iii) The order of $SZ(m, K)$ is d.
(iv) The order of $PSL(m, K)$ is $(q^m - 1)(q^m - q) \cdots (q^m - q^{m-2})q^{m-1}/d$.

Proof. Let C be the family of all ordered bases of V. Before we start, we will show that there is a bijection $W : GL(V) \to C$. If $\{v_1, \ldots, v_m\}$ is an ordered basis for V, $T \in GL(V)$ and $v = \alpha_1 v_1 + \cdots + \alpha_m v_m \in V$, then

$$T(v) = T(\alpha_1 v_1 + \cdots + \alpha_m v_m) = \alpha_1 T(v_1) + \cdots + \alpha_m T(v_m) \in V$$

by linearity of T, and so, $T(v) \in span\{T(v_1), \ldots, T(v_m)\}$. Since $T \in GL(V)$, T is invertible, and thus, $T : V \to V$ is a surjection; all elements in the codomain V are elements of the form $T(u)$ for some $u \in V$. Thus, $\{T(v_1), \ldots, T(v_m)\}$ is also a basis for V. So for a fixed ordered basis $\{v_1, \ldots, v_m\}$, each $T \in GL(V)$ maps to exactly one basis in C.

Now assume that two linear transformations $T_1, T_2 \in GL(V)$ maps an ordered basis to the same basis in C, i.e., $T_1 v_1 = T_2 v_1, \ldots, T_1 v_m = T_2 v_m$, and we try to show $T_1 = T_2$. We get that

$$\begin{aligned} T_1(v) &= T_1(\alpha_1 v_1 + \cdots + \alpha_m v_m) = \alpha_1 T_1 v_1 + \cdots + \alpha_m T_1 v_m \\ &= \alpha_1 T_2 v_1 + \cdots + \alpha_m T_2 v_m = T_2(\alpha_1 v_1 + \cdots + \alpha_m v_m) = T_2(v) \end{aligned}$$

and thus, since T_1 and T_2 have the same effect on a vector v, they are the same linear transformation, i.e. $T_1 = T_2$. Thus, W is a bijection, so $|GL(V)| = |C|$. Therefore, to find the order of $GL(V)$ we can try to find the number of ordered bases $|C|$.

We begin by proving (i). Let $\{v_1, \ldots, v_m\}$ be an ordered basis for V. We want to count how many ways we can choose the ordered basis $\{v_1, \ldots, v_m\}$. For v_1, we can choose any of the q^m vectors of V except the zero vector. Hence, v_1 can be chosen in $q^m - 1$ ways. For v_2, we may not choose a vector that is a multiple of v_1, or the zero vector. There are $q - 1$ non-zero multiples of v_1 and one zero vector. Hence, there are $q^m - q$ choices for v_2. By similar arguments, we have $q^m - q^2$ choices for v_3, $q^m - q^3$ choices for v_4, and so forth, and finally $q^m - q^{m-1}$ choices for v_m. Thus, we have $(q^m - 1)(q^m - q) \cdots (q^m - q^{m-1})$ choices for an ordered basis $\{v_1, \ldots, v_m\}$ of V, and thus, $|GL(V)| = (q^m - 1)(q^m - q) \cdots (q^m - q^{m-1})$.

Now we can prove (ii). If $A \in GL(m,K)$, then $\det(A)$ is some non-zero element of K, i.e.

$$\det : GL(m,K) \to K\backslash\{0\} = K^\times.$$

From linear algebra we know that $\det(AB) = \det(A)\det(B)$. The determinant function is also surjective: for each $\alpha \in K^\times$, we can choose $B = diag\{\alpha, 1, \ldots, 1\}$ so that $\det(B) = \alpha$. Thus, the determinant function above is a surjective homomorphism of groups, where $K^\times$ is a group under multiplication. It is clear that the kernel of det is $SL(m,K) \leqslant GL(m,K)$.

Now, by the First Isomorphism theorem, we have

$$\frac{GL(m,K)}{\ker(\det)} = \frac{GL(m,K)}{SL(m,K)} \cong K^\times.$$

Thus, these groups have the same order, so we have

$$\left|\frac{GL(m,K)}{SL(m,K)}\right| = |K^\times| = q-1.$$

By rewriting this and using (i), we get

$$\begin{aligned}|SL(m,K)| &= |GL(m,K)|/(q-1) = (q^m-1)(q^m-q)\cdots(q^m-q^{m-1})/(q-1)\\ &= (q^m-1)(q^m-q)\cdots(q^m-q^{m-2})q^{m-1}.\end{aligned}$$

Before we prove (iii), we will show that $\alpha^m = 1$ if and only if $\alpha^d = 1$ for $\alpha \in K^\times$, where $d = (m, q-1)$. Note that m is a multiple of d, say $m = dk$ for some $k \in \mathbb{Z}$. If $\alpha^d = 1$, then $\alpha^m = \alpha^{dk} = (\alpha^d)^k = 1$. Now assume $\alpha^m = 1$. There are integers a, b such that $d = am + b(q-1)$ due to the Euclidean algorithm, so we get

$$\alpha^d = \alpha^{am+b(q-1)} = \alpha^{am}\alpha^{b(q-1)} = \alpha^{am}(\alpha^{q-1})^b = \alpha^{am} = (\alpha^m)^a = 1$$

since $a^{q-1} = 1$ due to Lagrange's theorem. Recall that we had defined $SZ(m,K)$ as $\{\alpha I : \alpha^m = 1\}$ and so we get that

$$SZ(m,K) \cong \{\alpha \in K^\times : \alpha^m = 1\} = \{\alpha \in K^\times : \alpha^d = 1\}.$$

Now, if π is a generator of $K^\times$, then for each $\alpha \in K^\times$ there exists an i such that $\alpha = \pi^i$. We take the dth power to get $1 = \alpha^d = \pi^{id}$. The order of π is $q-1$ because it is a generator of $K^\times$, so di must be a multiple of $q-1$, say, $di = (q-1)k$, with $k \in \mathbb{Z}$. This gives $i = (\frac{q-1}{d})k$. We substitute this into our initial form to get

$$\alpha = \pi^i = \pi^{(\frac{q-1}{d})k} = \left(\pi^{\frac{q-1}{d}}\right)^k.$$

We see that when $k \in \mathbb{Z}$ takes values from 1 to d, α assumes distinct values. Hence, $|SZ(m,K)| = |\pi^{\frac{q-1}{d}}| = d$.

Finally, we can prove (iv). We defined $PSL(m,K)$ as $\frac{SL(m,K)}{SZ(m,K)}$. Taking the order and using (ii),(iii) we get

$$|PSL(m,K)| = \frac{|SL(m,K)|}{|SZ(m,K)|} = \frac{(q^m-1)(q^m-q)\dots(q^m-q^{m-2})q^{m-1}}{d}.$$

□

3 Simplicity of PSL(2,K)

In this section, we will focus on PSL when the dimension is 2. The first lemma is quite strong, since it allows us to have some control over the elements in $SL(2,K)$.

Lemma 3.1. *Every element $A \in SL(2,K)$ is a conjugate in $SL(2,K)$ to one of the following:*

(i) $\begin{bmatrix} a & 0 \\ 0 & a^{-1} \end{bmatrix}$, $a \neq 0$,

(ii) $\pm \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix} = \pm B_{12}(a), a \neq 0$,

(iii) $\begin{bmatrix} 0 & -a^{-1} \\ a & b \end{bmatrix}$, $a \neq 0$.

Proof. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL(2,K)$. Let $e_1 = (1,0)$ and $e_2 = (0,1)$. Let $v_1 = e_1$ and multiply v_1 by A from the left to get

$$Av_1 = \begin{bmatrix} a \\ c \end{bmatrix} = ae_1 + ce_2.$$

Now, we have two cases. Either e_1 is not an eigenvector of A or it is an eigenvector of A.

In the first case, assume $e_1 = v_1$ is not an eigenvector of A. Then $c \neq 0$. So $Av_1 = v_2$ will not be a scalar multiple of v_1 and $\{v_1, v_2\}$ forms a basis for K^2. In the basis $\{v_1, v_2\}$ the matrix A has the form

$$\begin{bmatrix} 0 & -1 \\ 1 & r \end{bmatrix}.$$

To see this, note that the first column of A in the basis $\{v_1, v_2\}$ should be the coordinates of v_2 in the basis $\{v_1, v_2\}$, and from $Av_1 = v_2 = 0 \cdot v_1 + 1 \cdot v_2$, we can see that the first column should be $[0,1]$. The second column must by necessity be $[-1, r]$ for some $r \in K$ since $\det A = 1$. Now, let $S \in GL(2,K)$ be the change of basis matrix $\{e_1, e_2\} \to \{v_1, v_2\}$, so that

$$\begin{bmatrix} 0 & -1 \\ 1 & r \end{bmatrix} = S \begin{bmatrix} a & b \\ c & d \end{bmatrix} S^{-1}.$$

Since $S \in GL(2, K)$, S has some determinant $\mu \in K^{\times}$. Let

$$D(\mu) = \text{diag}\{1, , \ldots, 1, \mu\}.$$

Then $T = D(\mu)^{-1}S$ has determinant 1. Furthermore, since $T = D(\mu)^{-1}S$ we get $S = D(\mu)T$ if we multiply by $D(\mu)$ from the left. Thus, we have

$$\begin{bmatrix} 0 & -1 \\ 1 & r \end{bmatrix} = D(\mu)T \begin{bmatrix} a & b \\ c & d \end{bmatrix} T^{-1}D(\mu)^{-1}.$$

By multiplying by $D(\mu)^{-1}$ from the left and $D(\mu)$ from the right, we get

$$D(\mu)^{-1} \begin{bmatrix} 0 & -1 \\ 1 & r \end{bmatrix} D(\mu) = T \begin{bmatrix} a & b \\ c & d \end{bmatrix} T^{-1}.$$

Now, if we simplify the left-hand side, we get

$$\begin{aligned} \begin{bmatrix} 1 & 0 \\ 0 & \mu^{-1} \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & r \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \mu \end{bmatrix} &= \begin{bmatrix} 0 & -1 \\ \mu^{-1} & \mu^{-1}r \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \mu \end{bmatrix} \\ &= \begin{bmatrix} 0 & -\mu \\ \mu^{-1} & r \end{bmatrix} = \begin{bmatrix} 0 & -s^{-1} \\ s & r \end{bmatrix}. \end{aligned}$$

Hence, since $T \in SL(2, K)$, A is conjugate in $SL(2, K)$ to a matrix of the form (iii).

In the second case, assume v_1 is an eigenvector of A, which means $c = 0$ and $Av_1 = v_2$ is a multiple of v_1. We have

$$A = \begin{bmatrix} a & b \\ 0 & d \end{bmatrix} \in SL(2, K).$$

Also, since $A \in SL(2, K)$, d must be equal to a^{-1} where $a \neq 0$. If $a = \pm 1$ we get a matrix of the form (ii). Now assume $a \neq \pm 1$. We want to show that we can always find a $t \in K$ such that

$$\begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ 0 & a^{-1} \end{bmatrix} \begin{bmatrix} 1 & -t \\ 0 & 1 \end{bmatrix}$$

is a matrix of the form (i). We see that

$$\begin{aligned} \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ 0 & a^{-1} \end{bmatrix} \begin{bmatrix} 1 & -t \\ 0 & 1 \end{bmatrix} &= \begin{bmatrix} a & b + ta^{-1} \\ 0 & a^{-1} \end{bmatrix} \begin{bmatrix} 1 & -t \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} a & -ta + b + ta^{-1} \\ 0 & a^{-1} \end{bmatrix} = H. \end{aligned}$$

The matrix H is of the form (i) if

$$b - ta + ta^{-1} = 0.$$

By solving the above equation for t we get

$$t = \frac{b}{a - a^{-1}}.$$

It is clear that t is well-defined for all b and H is of the form (i). Thus, our matrix A is conjugate in $SL(2, K)$ to a matrix of the form (i).

□

Lemma 3.2. *Let K be a finite field of characteristic 2. Then every element of K is a square.*

Proof. Define the map

$$\phi : K \to K$$

by $a \mapsto a^2$. First, we want to show that ϕ is injective. Thus, we assume $a^2 = b^2$ and try to show that this implies $a = b$. If $a^2 = b^2$, then $a^2 - b^2 = 0$. Since K is commutative, we get that $(a - b)(a + b) = 0$. Hence, $a = -b$ or $a = b$, but since K has characteristic 2 then $1 = -1$. Thus, $a = -a$ and we get that $a = b$, and ϕ is injective. If $a \in K$ then $a^2 \in K$ because K is a field. We see that the $q = |K|$ distinct elements in K are mapped by ϕ to q distinct squares in K, and so every element in K must be a square. □

Theorem 3.3. *Let K be a finite field and let H be a normal subgroup of $SL(2, K)$. If H contains an elementary transvection then $H = SL(2, K)$.*

Proof. Note first that if H contains an elementary transvection $B_{21}(\lambda)$, then H also contains $B_{12}(-\lambda)$. To see this, consider the product

$$UB_{21}(\lambda)U^{-1}$$

where $U = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \in SL(2, K)$. We get

$$\begin{aligned} UB_{21}(\lambda)U^{-1} &= \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \lambda & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} -\lambda & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -\lambda \\ 0 & 1 \end{bmatrix} = B_{12}(-\lambda). \end{aligned} \tag{3.1}$$

Since H is normal in $SL(2, K)$, we have that $B_{12}(-\lambda) \in H$. Now, we want to show that H contains every elementary transvection. By (3.1) we may assume that $B_{12}(\lambda) \in H$ for some $\lambda \in K^{\times}$. Now let $S = \begin{bmatrix} a & b \\ 0 & a^{-1} \end{bmatrix}$, where $a \neq 0$, and form the conjugate

$$\begin{aligned} SB_{12}(\lambda)S^{-1} &= \begin{bmatrix} a & b \\ 0 & a^{-1} \end{bmatrix} \begin{bmatrix} 1 & \lambda \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a^{-1} & -b \\ 0 & a \end{bmatrix} \\ &= \begin{bmatrix} a & a\lambda + b \\ 0 & a^{-1} \end{bmatrix} \begin{bmatrix} a^{-1} & -b \\ 0 & a \end{bmatrix} \\ &= \begin{bmatrix} aa^{-1} & -ba + a^2\lambda + ba \\ 0 & a^{-1}a \end{bmatrix} \\ &= \begin{bmatrix} 1 & a^2\lambda \\ 0 & 1 \end{bmatrix} = B_{12}(\lambda a^2). \end{aligned} \tag{3.2}$$

These conjugates are also in H since H is normal in $SL(2, K)$, so $B_{12}(\lambda a^2) \in H$. If K has characteristic 2, then by Lemma 3.2, all elements in K are of the form a^2 with $a \in K$ since K is finite, and hence, all elements in K are of the form λa^2 with $a \in K$. By (3.1), if $B_{12}(\lambda) \in H$ then $B_{21}(-\lambda)$, and since K has characteristic 2, then $B_{21}(\lambda) \in H$. Thus H contains all elementary transvections, and by Lemma 2.16, $H = SL(2, K)$.

Now assume K does not have characteristic 2. By (3.2) we can assume $B_{12}(\lambda c^2) \in H$ for some $c \in K^\times$. Since H is a group, every element of H has an inverse in H. The inverse of $B_{12}(\lambda c^2)$ is $B_{12}(-\lambda c^2)$:

$$\begin{aligned} B_{12}(\lambda c^2) \cdot B_{12}(-\lambda c^2) &= \begin{bmatrix} 1 & \lambda c^2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \lambda c^{-2} \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & \lambda c^2 - \lambda c^{-2} \\ 0 & 1 \end{bmatrix} = I. \end{aligned}$$

Hence, the product

$$B_{12}(\lambda a^2) \cdot B_{12}(-\lambda c^2) = B_{12}(\lambda(a^2 - c^2))$$

is also in H. Now, a and c can be chosen to be any elements in $K^\times$, so we want to show that all the elements in K can be represented as $\lambda(a^2 - c^2)$. First, we claim that every element $\mu \in K$ can be written as a difference of two squares, if the characteristic of K is not 2, i.e.

$$\mu = \left(\frac{1}{2}(\mu + 1)\right)^2 - \left(\frac{1}{2}(\mu - 1)\right)^2.$$

To see this, simply expand the right-hand side to get

$$\begin{aligned} \left(\frac{1}{2}(\mu + 1)\right)^2 - \left(\frac{1}{2}(\mu - 1)\right)^2 &= \frac{1}{4}(\mu^2 + 2\mu + 1) - \frac{1}{4}(\mu^2 - 2\mu + 1) \\ &= \frac{1}{2}\mu + \frac{1}{2}\mu = \mu. \end{aligned}$$

Since all the elements of K are of the form $a^2 - c^2$, they are also of the form $\lambda(a^2 - c^2)$, and thus, H contains all the elementary transvections $B_{12}(\mu)$, where

$\mu \in K^\times$. By (3.1), H also contains all the elementary transvections $B_{21}(\mu)$, where $\mu \in K^\times$. Finally, since H contains all elementary transvections, we get that $H = SL(2, K)$ by Lemma 2.16. □

Theorem 3.4. *Let G, R be groups, and let*

$$\phi : G \to R$$

be a group homomorphism. Let M be a normal subgroup of R and define

$$H = \{x \in G : \phi(x) \in M\}.$$

Then H is a normal subgroup of G.

Proof. Since ϕ is a group homomorphism from G to R, it will map e_G to e_R, that is, $\phi(e_G) = e_R$. Thus, $e_R \in M$ and by the construction of H, we get that $e_G \in H$, so H is non-empty. If $x \in H$, then $\phi(x) \in M$. Since M is a group, we get that $(\phi(x))^{-1} = \phi(x^{-1}) \in M$ and thus $x^{-1} \in H$ by construction of H. Now, if $x, y \in H$, then $\phi(x), \phi(y) \in M$. Also, since

$$\phi(x)\phi(y) = \phi(xy) \in M$$

we have that $xy \in H$ by construction of H.

Now we want to show normality of H, i.e. we want to show that if $a \in G$ and $x \in H$, then $axa^{-1} \in H$, $\forall a \in G$ and $\forall x \in H$. We take some $a \in G$ and some $x \in H$ and apply ϕ. We get

$$\phi(axa^{-1}) = \phi(a)\phi(x)\phi(a)^{-1}.$$

We know that $\phi(a) = r$ for some $r \in R$ and $\phi(a^{-1}) = r^{-1}$ for some $r^{-1} \in R$ and since M is normal in R, we get that

$$r\phi(x)r^{-1} = \phi(a)\phi(x)\phi(a)^{-1} \in M.$$

Thus, $\phi(axa^{-1}) \in M$ and we have that $axa^{-1} \in H$ by the construction of H. The elements $a \in G$ and $x \in H$ were chosen generally, so we get that $axa^{-1} \in H, \forall a \in G$ and $\forall x \in H$, and hence, H is normal in G and the theorem is established.

□

Now we are ready to prove that $PSL(2, K)$ is a simple group for all finite fields K where $|K| \geqslant 4$. First, however, we will show that $PSL(2, 2)$ and $PSL(2, 3)$ are NOT, in fact, simple groups, since they can be shown to contain a non-trivial normal subgroup.

First, let's look at $PSL(2, 2)$. Now, we have seen that the center of SL is $SZ = \{\alpha I : \alpha^m = 1\}$ which is just $\{I\}$ in this case. Thus, $PSL(2, 2) = SL(2, 2)/I = SL(2, 2)$. It is easy to list all elements of $SL(2, 2) = PSL(2, 2)$:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

We realize that all these have determinant 1, since $-1 = 1$ in $\mathbb{Z}_2$, and any other matrix would have determinant 0 or $2 = 0$.

Now, any group of order 6 must be isomorphic to $\mathbb{Z}_6$ or S_3. It is known that $\mathbb{Z}_6$ is cyclic and therefore Abelian, while S_3 is neither cyclic nor Abelian. So, if we can show that $PSL(2,2)$ is not Abelian, then it must be isomorphic to S_3. We can simply take two matrices and see if they commute:

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

So, we have produced two elements in $PSL(2,2)$ that do not commute, so $PSL(2,2) \cong S_3$. It is also known that A_3 is a non-trivial normal subgroup in S_3, and so, $PSL(2,2)$ must also contain a non-trivial normal subgroup, and thus, $PSL(2,2)$ is not simple.

For the case $PSL(2,3)$, we need a bit more, but with a similar idea. First, we note that $|SL(2,3)| = 24$ and $|SZ(2,3)| = 2$ so $|PSL(2,3)| = 12$. This can also be seen by verifying that $A \in SL(2,3)$ and $-A \in SL(2,3)$ will belong to the same coset, i.e. A and $-A$ will both be mapped to the same element $B \in SL(2,3)/SZ(2,3)$, so we would expect that the order of $PSL(2,3)$ is half of the order of $SL(2,3)$, which is 24. We will show that $PSL(2,3) \cong A_4$. Note that $|A_4| = 4!/2 = 24/2 = 12$, so the orders match.

Before we move on, we will define **group actions**. A group G is said to act on a set X if

(i) $e_G \cdot x = x$,
(ii) $g_1 \cdot (g_2 \cdot x) = (g_1 \cdot g_2) \cdot x$

are satisfied for all $g_1, g_2 \in G$ and $x \in X$. Now, let $X = \mathbb{P}^1(\mathbb{Z}_3) = \{0, 1, 2, \infty\}$ and let $Y = PSL(2,3)$. We will define a map σ that lets an element of $PSL(2,3)$ map an element of X to an element of X. Define

$$\sigma : Y \times X \to X$$

by

$$\left(\begin{bmatrix}a & b\\ c & d\end{bmatrix}, x\right) \mapsto \frac{ax+b}{cx+d}.$$

We denote this by $\begin{bmatrix}a & b\\ c & d\end{bmatrix} \cdot x = \frac{ax+b}{cx+d}$. We will show that this is in fact a group action. We see that (i) is satisified because

$$e_G \cdot x = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} \cdot x = \frac{1 \cdot x + 0}{0 \cdot x + 1} = x.$$

Next, let $g_1 = \begin{bmatrix}a_1 & b_1\\ c_1 & d_1\end{bmatrix}$ and let $g_2 = \begin{bmatrix}a_2 & b_2\\ c_2 & d_2\end{bmatrix}$. Then

$$\begin{aligned} g_1 \cdot (g_2 \cdot x) &= \begin{bmatrix}a_1 & b_1\\ c_1 & d_1\end{bmatrix} \cdot \frac{a_2x+b_2}{c_2x+d_2} = \frac{a_1\frac{a_2x+b_2}{c_2x+d_2} + b_1}{c_1\frac{a_2x+b_2}{c_2x+d_2} + d_1} \\ &= \frac{(a_1a_2 + b_1c_2)x + a_1b_2 + b_1d_2}{(c_1a_2 + d_1c_2)x + c_1b_2 + d_1d_2}. \end{aligned}$$

while

$$(g_1 \cdot g_2) \cdot x = \begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix} \begin{bmatrix} a_2 & b_1 \\ c_1 & d_1 \end{bmatrix} \cdot x = \begin{bmatrix} a_1a_2 + b_1c_2 & a_1b_2 + b_1d_2 \\ a_2c_1 + c_2d_1 & b_2c_1 + d_2d_1 \end{bmatrix} \cdot x$$
$$= \frac{(a_1a_2 + b_1c_2)x + a_1b_2 + b_1d_2}{(c_1a_2 + d_1c_2)x + c_1b_2 + d_1d_2}.$$

Thus, σ satisfies (ii) and hence, σ is a group action. Now, if we fix some $A \in PSL(2,3)$, we can investigate what element in X that A maps to for all $x \in X$. For example, let $B = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}$. Then $B(0) = 1$, $B(1) = \infty$, $B(2) = 2$, and $B(\infty) = 0$. In fact, B represents the permutation $(0\ 1\ \infty) \in A_4$. By checking all 12 elements in $PSL(2,3)$ we will see that they correspond to the 12 permutations in A_4. We say that $Y = PSL(2,3)$ acts faithfully on X. Indeed, the only way that A and A' could correspond to the same $\frac{ax+b}{cx+d}$ would be if $A = \beta A'$ for some non-zero scalar $\beta \in \mathbb{Z}_3$. But the only other non-zero scalar in $\mathbb{Z}_3$ is $2 = -1$, and we already showed that A and $-A$ is represented by the same elements in $PSL(2,3)$. So $PSL(2,3) \cong A_4$.

Now, A_4 does contain the normal subgroup $\{e, (12)(34), (13)(24), (14)(23)\}$ and thus, is not simple. It then follows that $PSL(2,3)$ also contains a non-trivial normal subgroup, and so, $PSL(2,3)$ is not simple.

However, $PSL(2,q)$ is simple for $q \geqslant 4$, which the next theorem will show.

Theorem 3.5. *(Jordan–Moore)*
The group $PSL(2,K)$ is a simple group if $q \geqslant 4$, where $q = |K|$.

Proof. Let $M \neq \{e\}$ be a normal subgroup of $PSL(2,K)$. We want to show that $M = PSL(2,K)$. Denote by ϕ the canonical homomorphism

$$\phi : SL(2,K) \to PSL(2,K)$$

such that $g \mapsto SZ(2,K)g$ for $g \in SL(2,K)$. Let

$$H = \phi^{-1}(M) = \{x \in SL(2,K) : \phi(x) \in M\}.$$

Then H is a normal subgroup of $SL(2,K)$ by Theorem 3.4. Note that ϕ will map $SZ(2,K)$ to $\{e\} \leqslant PSL(2,K)$. Therefore, by construction of H, we have that H in fact properly contains $SZ(2,K)$, since $M \neq \{e\}$. The homomorphism ϕ maps $g \in SL(2,K)$ to the coset in $PSL(2,K)$ which g belongs to modulo $SZ(2,K)$. If $H = SL(2,K)$, we have that ϕ maps H onto $SL(2,K)/SZ(2,K) = PSL(2,K)$, and by construction of H, we get that $M = PSL(2,K)$. By this argument, it suffices to show that $H = SL(2,K)$. Thus, we want to show that H contains an elementary transvection because this implies that $H = SL(2,K)$ by Lemma 2.16, and subsequently, $M = PSL(2,K)$.

Let A be a matrix in H which is different from $\pm I$. By Lemma 3.1, every matrix in $SL(2,K)$ is a conjugate in $SL(2,K)$ to one of the following:

(i) $\begin{bmatrix} a & 0 \\ 0 & a^{-1} \end{bmatrix}$, $a \neq 0$,

(ii) $\pm \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix} = \pm B_{12}(a), a \neq 0$,

(iii) $\begin{bmatrix} 0 & -a^{-1} \\ a & b \end{bmatrix}$, $a \neq 0$.

We want to show that H contains an elementary transvection in all three cases. We begin with Case (iii), so assume A is a conjugate in $SL(2, K)$ to

$$D = \begin{bmatrix} 0 & -a^{-1} \\ a & b \end{bmatrix}.$$

Since H is normal in $SL(2, K)$, we also have $D \in H$. Now define

$$T = \begin{bmatrix} c^{-1} & 0 \\ 0 & c \end{bmatrix} \in SL(2, K),$$

where $c \neq \pm 1$ is an element of $K^\times$ we have yet to choose, and form the product $U = TDT^{-1}D^{-1}$. Since H is a normal subgroup in $SL(2, K)$, we have that $TDT^{-1} \in H$ and $D^{-1} \in H$, so $U \in H$. We expand the product to get

$$\begin{aligned} U = TDT^{-1}D^{-1} &= \begin{bmatrix} c^{-1} & 0 \\ 0 & c \end{bmatrix} \begin{bmatrix} 0 & -a^{-1} \\ a & b \end{bmatrix} \begin{bmatrix} c & 0 \\ 0 & c^{-1} \end{bmatrix} \begin{bmatrix} b & a^{-1} \\ -a & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & -c^{-1}a^{-1} \\ ca & cb \end{bmatrix} \begin{bmatrix} c & 0 \\ 0 & c^{-1} \end{bmatrix} \begin{bmatrix} b & a^{-1} \\ -a & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & -c^{-2}a^{-1} \\ c^2a & cc^{-1}b \end{bmatrix} \begin{bmatrix} b & a^{-1} \\ -a & 0 \end{bmatrix} \\ &= \begin{bmatrix} c^{-2}a^{-1}a & 0 \\ c^2ab - ab & c^2aa^{-1} \end{bmatrix} = \begin{bmatrix} c^{-2} & 0 \\ ab(c^2 - 1) & c^2 \end{bmatrix}. \end{aligned} \tag{3.3}$$

Then U is a matrix of the form

$$L = \begin{bmatrix} d & 0 \\ g & d^{-1} \end{bmatrix}. \tag{3.4}$$

Take $B = B_{21}(1)$ and form the product $E = BLB^{-1}L^{-1} \in H$. We get

$$\begin{aligned} E = BLB^{-1}L^{-1} &= \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} d & 0 \\ g & d^{-1} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} d^{-1} & 0 \\ -g & d \end{bmatrix} \\ &= \begin{bmatrix} d & 0 \\ d + g & d^{-1} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} d^{-1} & 0 \\ -g & d \end{bmatrix} \\ &= \begin{bmatrix} d & 0 \\ d + g - d^{-1} & d^{-1} \end{bmatrix} \begin{bmatrix} d^{-1} & 0 \\ -g & d \end{bmatrix} \\ &= \begin{bmatrix} dd^{-1} & 0 \\ d^{-1}(d + g - d^{-1}) - d^{-1}g & d^{-1}d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 - d^{-2} & 1 \end{bmatrix}. \end{aligned} \tag{3.5}$$

If $d \neq \pm 1$, then E will be an elementary transvection. For our matrix (3.4) above, this means that E is an elementary transvection if c is such that $1 - c^4 \neq 0$, since this means that $1 - d^{-2} \neq 0$. If $q > 5$, then we can always find a c such that $c^4 - 1 \neq 0$ because the polynomial $t^4 - 1$ can have no more than four zeros, so H contains an elementary transvection E for the case $q > 5$. If $q = 4$, let $K = \mathbb{Z}_2[x]/(x^2 + x + 1) = \{0, 1, [x], [x + 1]\}$. We will see that all elements of K satisfy

$$t^4 - t = 0. \tag{3.6}$$

The element 0 and 1 trivially satisfies (3.6). If $t = [x]$ then

$$t \cdot t = [x] \cdot [x] = [x \cdot x] = [x^2] = [x^2 + x + 1] + [x + 1] = [x + 1].$$

Thus,

$$[x]^4 = [x^2] \cdot [x^2] = [x + 1] \cdot [x + 1] = [x^2 + 1] = [x^2 + x + 1] + [x] = [x]$$

and hence, $[x]^4 - [x] = 0$.

If $t = [x + 1]$, then $[x + 1]^2 = [x]$ so $[x + 1]^4 = [x]^2 = [x + 1]$, and hence, $[x + 1]^4 - [x + 1] = 0$. So we see that if $c \neq 1$, then $c^4 \neq 1$. Thus, H contains an elementary transvection E for $q = 4$.

For $q = 5$, we look at our matrix $U = \begin{bmatrix} c^{-2} & 0 \\ ab(c^2 - 1) & c^2 \end{bmatrix}$ again, and note that $K \cong \mathbb{Z}_5$. If $b \neq 0$, then we choose $c = 2 \in \mathbb{Z}_5$. We see that $c^2 - 1 = 3 \neq 0$, and that $c^2 = c^{-2} = 4$ so our matrix U can be written as

$$U_5 = \begin{bmatrix} 4 & 0 \\ 3ab & 4 \end{bmatrix}.$$

By squaring U_5 we get

$$U_5^2 = \begin{bmatrix} 4 & 0 \\ 3ab & 4 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 3ab & 4 \end{bmatrix} = \begin{bmatrix} 16 & 0 \\ 24ab & 16 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 4ab & 1 \end{bmatrix} = B_{21}(4ab) \in H,$$

which is an elementary transvection because $b \neq 0$ and $a \neq 0$. Thus, H contains an elementary transvection for $q = 5$ and $b \neq 0$.

If $b = 0$, then A is a conjugate in $SL(2, K)$ to $D = \begin{bmatrix} 0 & -a^{-1} \\ a & 0 \end{bmatrix}$. Now we form the product

$$S = B_{12}(2a^{-1})DB_{12}(2a^{-1})^{-1} = B_{12}(2a^{-1})DB_{12}(-2a^{-1}) \in H.$$

We get

$$\begin{aligned} S &= \begin{bmatrix} 1 & 2a^{-1} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & -a^{-1} \\ a & 0 \end{bmatrix} \begin{bmatrix} 1 & -2a^{-1} \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2 & -a^{-1} \\ a & 0 \end{bmatrix} \begin{bmatrix} 1 & -2a^{-1} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ a & -2 \end{bmatrix} \in H. \end{aligned}$$

Since $-2 = 2^{-1}$ in $\mathbb{Z}_5$, we see that S is a matrix of the form (3.4), where $d = 2$. Hence, by the calculation (3.5), if $S \in H$ then $E \in H$ where $d = 2$. For this particular d, we will find that

$$E = \begin{bmatrix} 1 & 0 \\ 1 - d^{-2} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 - (2^{-1})^2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 - 4 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \in H.$$

Thus, H contains an elementary transvection E when $q = 5$ and $b = 0$.

Case (i). If A is conjugate in $SL(2, K)$ to a matrix of the form $\begin{bmatrix} a & 0 \\ 0 & a^{-1} \end{bmatrix}$, then A is a conjugate in $SL(2, K)$ to a matrix of the form (3.4) where $g = 0$. By the same arguments, we have seen that this will lead to the conclusion that H contains an elementary transvection. Since $a \neq 0$, the only problem would be if $a = \pm 1$, i.e. that A is conjugate in $SL(2, K)$ to $\pm I$. If that was the case, then

$$SAS^{-1} = \pm I$$

for some $S \in SL(2, K)$. This would have to mean that $SA = \pm S$, which implies that $A = \pm I$. But by assumption, A is a matrix in H which is different from $\pm I$, so this would lead to a contradiction. Thus, $A \neq \pm I$ and H contains an elementary transvection.

Case (ii). If A is conjugate to a matrix of the form $\pm \begin{bmatrix} 1 & f \\ 0 & 1 \end{bmatrix}$, then H already contains an elementary transvection.

We have shown that we can always find an elementary transvection in H, and thus, by Theorem 3.3, $H = SL(2, K)$. Hence, we see that $M = PSL(2, K)$, and since $PSL(2, K)$ contains no non-trivial normal subgroups, $PSL(2, K)$ is simple.

□

4 Simplicity of PSL(m,K)

Since the rows of an invertible matrix form a basis for V, we realise that $B_{ij}(\lambda)$ represents a linear transformation T that fixes all basis vectors but the jth basis vector if we have chosen a suitable basis. In other words, $Tv_k = v_k$ for all $k \neq j$ and $Tv_j = v_j + \lambda v_i$.

Since V is an m-dimensional vector space, the vectors $\{v_k\}$ form a basis for an $(m-1)$-dimensional subspace of V because the ith basis vector was removed. This prompts us to make the following definition.

Definition 4.1. *A **hyperplane** H is an $(m-1)$-dimensional subspace of an m-dimensional vector space over a field K.*

In the previous paragraph, we saw that the linear transformation $B_{ij}(\lambda)$ fixes the $(m-1)$ basis vectors for some $(m-1)$-dimensional hyperplane H. In fact, T will fix all vectors in H, for if $v = \alpha_1 v_1 + \ldots + \alpha_{m-1} v_{m-1} \in H$, then

$$\begin{aligned} T(v) &= T(\alpha_1 v_1) + \ldots + T(\alpha_{m-1} v_{m-1}) = \alpha_1 T v_1 + \ldots + \alpha_{m-1} T v_{m-1} \\ &= \alpha_1 v_1 + \ldots + \alpha_{m-1} v_{m-1} = v. \end{aligned}$$

If we take a vector $w \in V$ where $w \notin H$, then $w + H$ will form an affine hyperplane that is parallel to H. If we take a $\mu \in K$ which is not 0 or 1, then $\mu w + H$ will form a different hyperplane that is parallel to H. So if we let μ vary over K, the union of the hyperplanes $\mu w + H$ will fill up the entire vector space V. In other words, the hyperplanes $\mu w + H$ form disjoint cosets whose union is V. Thus, every vector v in V can be uniquely written in the form

$$v = \mu w + h,$$

where $\mu \in K$ and $h \in H$.

Lemma 4.2. *Let H be a hyperplane in V and let $T \in GL(V)$ be a linear transformation that fixes H pointwise. If $w \in V$ but $w \notin H$, then*

$$T(w) = \mu w + h_0$$

for some $\mu \in K$ and some $h_0 \in H$.

Furthermore, if $v \in V$, then

$$T(v) = \mu v + h_1$$

for some $h_1 \in H$.

Proof. Since $w \in V$, then $T(w) \in V$ because $T : V \mapsto V$. By the preceeding discussion, all vectors in V can be written in the form $\mu w + h$ for some $h \in H$, so in particular, $T(w) = \mu w + h_0$ for some $\mu \in K$ and some $h_0 \in H$. If we take any $v \in V$, we can write v as $v = \lambda w + h_2$ for some $\lambda \in K$ and $h_2 \in H$ by the preceeding discussion. Due to linearity of T and the fact that T fixes H pointwise, we get

$$\begin{aligned} T(v) &= T(\lambda w + h_2) = \lambda T(w) + T(h_2) \\ &= \lambda(\mu w + h_0) + h_2 = \lambda\mu w + \lambda h_0 + h_2 \\ &= \lambda\mu w + \lambda h_0 + h_2 + \mu h_2 - \mu h_2 = \mu(\lambda w + h_2) + (1-\mu)h_2 + \lambda h_0 \\ &= \mu v + (1-\mu)h_2 + \lambda h_0. \end{aligned} \tag{4.1}$$

Now, it is clear that $(1-\mu)h_2 + \lambda h_0$ is some element in H, say, h_1. Thus, $T(v) = \mu v + h_1$ and we are done. □

In Lemma 4.2 above, we first consider the hyperplane H to be fixed. Then we fix a vector $w \in V$ that is not in H. We then pick a $T \in GL(V)$ that fixes H pointwise, and after picking this T, we arrive at a specific $\mu = \mu(T)$. It turns our that this $\mu(T)$ tells us something about a linear transformation T that fixes our hyperplane H.

Definition 4.3. *Let $T \in GL(V)$ fix a hyperplane H pointwise.*

*(i) If $\mu(T) \neq 1$ then T is a **dilatation**.*

*(ii) If $\mu(T) = 1$ and $T \neq 1_V$, then T is a **transvection**.*

Theorem 4.4. *Let H be a hyperplane in V, let $T \in GL(V)$ be a linear transformation that fixes H pointwise, and let $\mu = \mu(T)$.*

(i) If T is a dilatation, then T is associated with the matrix $D(\mu) = \text{diag}\{1, \ldots, 1, \mu\}$ for a suitable basis for V.

(ii) If T is a transvection, then T is associated with the matrix $B_{21}(1)$ for a suitable basis of V. Also, all eigenvectors of T are in H.

Proof. Since T fixes all vectors in H, i.e. $Tv = v, \forall v \in H$, then all nonzero vectors in H are eigenvectors of T with eigenvalue 1. Now we want to investigate if there are any other eigenvectors of T.

If we choose a $w \in V$ with $w \notin H$, then by Lemma 4.2 we have

$$Tw = \mu w + h_0$$

where $h_0 \in H$. The proof of Lemma 4.2 gives us that if $v \in V$ with $v \notin H$, then

$$Tv = \mu v + h_1 = \mu v + (1 - \mu)h_2 + \lambda h_0.$$

If v is an eigenvector of T, we have $Tv = av$ for some $a \in K$. In fact, v is an eigenvector of T if and only if $a = \mu$ and $\lambda h_0 = (\mu - 1)h_2$, which we will now show.

First assume $a = \mu$ and $\lambda h_0 = (\mu - 1)h_2$. This clearly shows that v is an eigenvector of T. To see this, note that Lemma 4.2 gave us that

$$Tv = \mu v + (1 - \mu)h_2 + \lambda h_0.$$

If $\lambda h_0 = (\mu - 1)h_2$, then $\lambda h_0 + (1 - \mu)h_2 = 0$, since $(\mu - 1) = -(1 - \mu)$. Thus, $Tv = av + 0 = av$ and v is an eigenvector of T with eigenvalue a.

Now assume that v is an eigenvector of T with eigenvalue $a \in K$. Then $Tv = av = \mu v + h_1$. We can rewrite this expression to get $(a - \mu)v = h_1$. Since v is a nonzero vector in V outside of H, then we must have $a = \mu$ because $\langle v \rangle$ and H intersects only at 0. Thus, $h_1 = 0$, which implies $\lambda h_0 = (\mu - 1)h_2$ and we have shown what we wanted.

(i) Assume T is a dilatation, so that $\mu \neq 1$, or equivalently $\mu - 1 \neq 0$. If $\mu - 1 \neq 0$, then $(\mu - 1)^{-1}$ exists. As we saw in the preceeding discussion, if T has any more eigenvectors, we must have $\lambda h_0 = (\mu - 1)h_2$. Since $\mu - 1 \neq 0$, we get $h_2 = \lambda(\mu - 1)^{-1}h_0$. In fact, $v = w + (\mu - 1)^{-1}h_0$ is an eigenvector of T:

$$\begin{aligned} Tv &= Tw + (\mu - 1)^{-1}h_0 = \mu w + h_0 + (\mu - 1)^{-1}h_0 \\ &= \mu(v - (\mu - 1)^{-1}h_0) + h_0 + (\mu - 1)^{-1}h_0 \\ &= \mu v - \mu(\mu - 1)^{-1}h_0 + h_0 + (\mu - 1)^{-1}h_0 \\ &= \mu v - (\mu(\mu - 1)^{-1} - 1 - (\mu - 1)^{-1})h_0 \\ &= \mu v + ((\mu - 1)^{-1}(\mu - 1) - 1)h_0 = \mu v. \end{aligned}$$

Thus, v is an eigenvector of T with eigenvalue μ. By adjoining v to a basis $\{v_1, \ldots, v_{m-1}\}$ of H, we get a basis for V, since v was chosen to be outside of

H and thus will be linearly independent to a basis for H. Hence, the associated matrix for T in this basis is $D(\mu) = diag\{1, \ldots, 1, \mu\}$.

(ii) If T is a transvection, then $\mu = 1$ and thus $\mu - 1 = 0$. Now we choose a $w \notin H$ such that $Tw = \mu w + h_3 = w + h_3$ where $h_3 \in H$ and $h_3 \neq 0$. If $v \notin H$ is to be an eigenvector of T, we must have $Tv = bv = v + h_4$ for some $b \in K$ and some $h_4 \in H$. But this means that $(b-1)v = h_4$. Then we must have $b = 1$ since 0 is the only point where $\langle v \rangle$ intersects H. If so, then $Tv = v$ and T would be the identity transformation. But by the definition of a transvection, T cannot be the identity transformation, and we arrive at a contradiction. Hence, T can have no eigenvectors outside of H. Finally, by taking a basis $\{h_0, u_3, \ldots, u_m\}$ of H and adjoining w as the first vector, this is a basis for V, and the associated matrix for T is $B_{21}(1)$.

□

At this point, we need to introduce some new concepts to move forward. If $T \in GL(V)$ is a transvection that fixes a hyperplane H, then $\mu(T) = 1$ by our definition. The first part of Lemma 4.2 then gives us that $Tw = w + h_0$. Also, the discussion preceeding Lemma 4.2 tells us that if $v \in V$, then v can be uniquely written as $v = \lambda w + h$ where $\lambda \in K$ and $h \in H$. So for a fixed v we get a certain scalar $\lambda \in K$. Now we define the function $\phi : V \to K$ by $\phi(v) = \phi(\lambda w + h) = \lambda$. In fact, ϕ is a linear functional, whose kernel is H. The kernel of ϕ is the set of all $v \in V$ such that $\phi(v) = 0$, i.e. $\lambda = 0$. If $\lambda = 0$, then $\lambda w + h \in H$, and since $w \notin H$, we see that $\lambda w + h \notin H$ if $\lambda \neq 0$.

Now, since T is a transvection, we have $\mu = 1$ so $\mu - 1 = 0$. This implies $Tv = v + \lambda h_0$ from the calculation (4.1) in the proof of Lemma 4.2. Also, from the preceeding discussion, we have

$$Tv = v + \phi(v)h_0 \tag{4.2}$$

for all $v \in V$, where ϕ is a linear functional and $h_0 \in H = \ker \phi$. This holds for each transvection T fixing a hyperplane H. In fact, for each transvection T fixing a hyperplane H, there exists linear functional ϕ and a $h_0 \in H$ such that (4.2) holds, so we can associate a particular ϕ and $h_0 \in H$ to each transvection T. Thus, we can denote each transvection T by

$$T = [\phi, h] : v \mapsto v + \phi(v)h_0.$$

The following lemma gives us some properties of transvections, and how they relate to this new notation.

Lemma 4.5. *Let V be a vector space over K.*

(i) If ϕ and ψ are linear functionals on V, and if $h, l \in V$ satisfy

$$\phi(h) = \psi(h) = \phi(l) = 0,$$

then

$$[\phi, h] \circ [\phi, l] = [\phi, h + l] \tag{4.3}$$

and

$$[\phi, h] \circ [\psi, h] = [\phi + \psi, h] \tag{4.4}$$

where $(\phi + \psi)(v) = \phi(v) + \psi(v)$.

(ii) For all $a \in K^\times$

$$[a\phi, h] = [\phi, ah]. \tag{4.5}$$

(iii) Assume $\phi \neq 0$ and $\psi \neq 0$. Then $[\phi, h] = [\psi, l]$ if and only if there is a scalar $a \in K^\times$ with $\psi = a\phi$ and $h = al$.

(iv) If $S \in GL(V)$, then

$$S[\phi, h]S^{-1} = [\phi S^{-1}, Sh] \tag{4.6}$$

where

$$\phi S^{-1}(v) = \phi(S^{-1}v). \tag{4.7}$$

Proof. (i) We start off by showing (4.3). To do this, we apply $[\phi, h] \circ [\phi, l]$ to a general vector $v \in V$. First we use our definition of $[\phi, h]$ to find what $[\phi, l](v)$ is, and then apply $[\phi, h]$ to that vector. Then, by using the definition again and rewriting the expression we get

$$\begin{aligned}
[\phi, h] \circ [\phi, l](v) &= [\phi, h](v + \phi(v)l) = (v + \phi(v)l) + \phi(v + \phi(v)l)h \\
&= (v + \phi(v)l) + (\phi(v) + \phi(v)\phi(l))h \\
&= (v + \phi(v)l) + \phi(v)h = v + \phi(v)(h + l) = [\phi, h + l](v)
\end{aligned}$$

which concludes the proof of (4.3).

To prove (4.4), we start off in a similar way. We simply use our definition to find what $[\psi, h]$ is, and then apply $[\phi, h]$ to that vector. We get

$$\begin{aligned}
[\phi, h] \circ [\psi, h](v) &= [\phi, h](v + \psi(v)h) = (v + \psi(v)h) + \phi(v + \psi(v)h)h \\
&= v + \psi(v)h + (\phi(v) + \psi(v)\phi(h))h = v + \psi(v)h + \phi(v)h \\
&= v + (\phi(v) + \psi(v))h.
\end{aligned}$$

By our previous discussion, we said that we interpret $[\phi + \psi, h](v)$ as the linear transformation with $v \mapsto v + (\phi(v) + \psi(v))h$, and the result follows.

(*ii*) To prove (4.5), we simply apply the left hand side to a general vector $v \in V$ and use the definition to get

$$[a\phi, h](v) = v + a\phi(v)h = v + \phi(v)ah = [\phi, ah](v).$$

(*iii*) We start by assuming $[\phi, h] = [\psi, l]$ and try to prove that it implies $\phi = a\psi$ and $h = al$ for some $a \in K^\times$. By the definition, we have

$$[\phi, h](v) = v + \phi(v)h$$

and

$$[\psi, l](v) = v + \psi(v)l.$$

Let $H = \ker \phi$. We assumed that $[\phi, h] = [\psi, l]$, and so $[\phi, h]$ and $[\psi, l]$ will have the same effect on a vector $v \in V$ where $v \notin H$. Thus, we get

$$v + \phi(v)h = v + \psi(v)l.$$

which implies $\phi(v)h = \psi(v)l$. Now, $\phi(v) = \lambda$ and $\psi(v) = \mu$ where λ, μ are elements in K. Note that $\phi(v) \neq 0$ implies that $\lambda \neq 0$, so in fact, μ and λ are elements in $K^\times$. Thus, $\lambda \in K^\times$ has an inverse $\lambda^{-1} \in K^\times$. By multiplying by λ^{-1} from the left in the equation above, we get $h = \lambda^{-1}\mu l$ and we see that $h = al$ for some $a \in K^\times$ since $h \neq 0$. Since we know that $\phi(v)h = \psi(v)l$ and $h = al$, then $\phi(v)al = a\phi(v)l = \psi(v)l$ which implies $\psi = a\phi$.

Now we prove the other direction, so assume that $\psi = a\phi$ and $h = al$ for some $a \in K^\times$. Then we apply $[\psi, l]$ to a general vector $v \in V$ and use the two assumptions to get

$$[\psi, l](v) = v + \psi(v)l = v + a\phi(v)l = v + \phi(v)h = [\phi, h](v)$$

which is what we wanted to show.

(*iv*) To prove (*iv*) we apply $S[\phi, h]S^{-1}$ to a general vector $v \in V$. We do this by applying the rightmost linear transformation to the vector, one at a time. By noting that S is linear and that $\phi(S^{-1}v)$ is a scalar, we get

$$\begin{aligned} S[\phi, h]S^{-1}(v) &= S[\phi, h](S^{-1}v) = S(S^{-1}v + \phi(S^{-1}v)h) \\ &= v + S(\phi(S^{-1}v)h) = v + \phi(S^{-1}v)Sh. \end{aligned}$$

On the other hand, we have

$$[\phi S^{-1}, Sh](v) = v + \phi S^{-1}(v)Sh$$

where ϕS^{-1} is defined in (4.7) and we are done. □

Theorem 4.6. *All transvections are conjugate in* $GL(m, K)$.

Proof. By Theorem 4.4, all transvections have matrix $B_{21}(1)$ for a suitable basis for V. This means that if T_1, T_2 are two transvections, we have that the matrix for T_1 can be written as $A_{T_1} = S_1 B_{21}(1) S_1^{-1}$ for a suitable basis for V, and the matrix for T_2 can be written as $A_{T_2} = S_2 B_{21}(1) S_2^{-1}$ for another suitable basis for V, where $S_1, S_2 \in GL(m, K)$. By rewriting the second equation, we get $S_2^{-1} A_{T_2} S_2 = B_{21}(1)$. Hence, we get

$$A_{T_1} = S_1 S_2^{-1} A_{T_2} S_2 S_1^{-1} = S_1 S_2^{-1} A_{T_2} (S_1 S_2^{-1})^{-1} = S_3 A_{T_2} S_3^{-1}$$

where $S_3 \in GL(m, K)$. Thus, all transvections are conjugate in $GL(m, K)$. □

Theorem 4.7. *Let V be a vector space that is not a two-dimensional vector space over $\mathbb{Z}_2$. Then $SL(V)$ is the commutator subgroup GL' of $GL(V)$.*

Proof. We will prove the theorem by showing that GL' is contained in $SL(V)$, and that $SL(V)$ is contained in GL', so that $SL(V) = GL'$. Now, the commutator subgroup GL' contains all elements of the form $aba^{-1}b^{-1}$, $a, b \in GL(V)$, as well as all products of such elements. But all these elements have determinant 1, since the rule for determinants gives that

$$\begin{aligned} \det(aba^{-1}b^{-1}) &= \det(a)\det(b)\det(a^{-1})\det(b^{-1}) \\ &= \det(a)\det(b)(\det a)^{-1}(\det b)^{-1} \\ &= \det(a)(\det a)^{-1}\det(b)(\det b)^{-1} = 1. \end{aligned}$$

Also, all products of commutators has determinant 1 for the same reason. Hence, $GL' \leqslant SL(V)$.

Now we want to prove that $SL(V) \leqslant GL'$. We denote by γ the natural map $\gamma : GL(V) \to GL(V)/GL'$. We claim now that if T_1, T_2 are any two transvections, then $\gamma(T_1) = \gamma(T_2) = \delta$, where δ is the coset that T_1 and T_2 both belong to. To see this, we start by noting that all transvections are conjugate in $GL(V)$ by Theorem 4.6. This means that for any two transvections T_1, T_2, there is a $S \in GL(V)$ such that $T_1 = S T_2 S^{-1}$. By rewriting this, we get $T_2 = S^{-1} T_1 S$. Now, note that $T_2^{-1} = S^{-1} T_1^{-1} S$. By forming the product $T_1 T_2^{-1}$, we find that

$$T_1 T_2^{-1} = T_1 S^{-1} T_1^{-1} S \in GL'$$

Since $T_1 T_2^{-1} \in GL'$, we have that $T_1 \cong T_2 \mod GL'$. Thus, T_1 and T_2 belong to the same coset δ, which means that $\gamma(T_1) = \gamma(T_2)$. In fact, all transvections belong to the same coset, since all transvections are conjugate in $GL(V)$.

Now, fix a hyperplane H in V and choose two non-zero vectors $h, l \in H$ such that $h + l \neq 0$. Such vectors can always be found since V is not a 2-dimensional vector space over $\mathbb{Z}_2$. Also, fix a linear functional ϕ such that $H = \ker \phi$. Then $[\phi, h]$ and $[\phi, l]$ are two transvections fixing a hyperplane H in V. By Lemma 4.5, we have that

$$[\phi, h] \circ [\phi, l] = [\phi, h + l]. \tag{4.8}$$

We want $[\phi, h+l]$ to be a transvection, i.e. that $h+l \neq 0$. By applying γ to (4.8), we get $\delta^2 = \delta$ in GL'. Multiplying by δ^{-1} gives us that $\delta = 1$. Thus, γ maps all transvections to 1 in GL', and so every transvection is in the kernel of γ. Now, note also that the kernel of γ is GL'. Hence, every transvection is in GL'. By Theorem 3.3, $SL(V)$ is generated by (elementary) transvections, so $SL(V) \leqslant GL'$. We found that $GL' \leqslant SL(V)$ and $SL(V) \leqslant GL'$, which means that $SL(V) = GL'$ and we are done.

□

Theorem 4.8. *Let V be an m-dimensional vector space over a field K, where $m \geqslant 3$. Then all transvections are conjugate in $SL(V)$.*

Proof. Let $T_1 = [\phi, h]$ be a transvection fixing a hyperplane H in V, and let $T_2 = [\psi, l]$ be a transvection fixing a hyperplane L in V. Furthermore, let $\{h, h_2, \ldots, h_{m-1}\}$ be a basis for H, and let $\{l, l_2, \ldots, l_{m-1}\}$ be a basis for L. Now, we choose $v, u \in V$ such that $\phi(v) = \psi(u) = 1$. Note that this implies that $v \notin H$ and $u \notin L$. Thus, we have that $B_1 = \{v, h, h_2, \ldots, h_{m-1}\}$ and $B_2 = \{u, l, l_2, \ldots, l_{m-1}\}$ are both bases for V.

Now let $S \in GL(V)$ be the linear transformation mapping B_1 to B_2 such that $S(v) = u$, $S(h) = l$ and $S(h_i) = l_i$ for $i = 2, \ldots, m-1$. Since $S \in GL(V)$, it has some determinant $\mu \in K^\times$. For a fixed basis, by the properties of the determinant, multiplying a column of a matrix A by some value a will result in a matrix with determinant $a \det A$. Thus, we multiply the rightmost column in our matrix S by μ^{-1}, resulting in a matrix S_1 with determinant 1. Note that S_1 also satisfies $S_1(v) = u$, $S_1(h) = l$ and $S_1(H) = L$. Since $m \geqslant 3$, we can always find a basis vector different from v and h. By the above discussion, S_1 also satisifies $S_1(h_{m-1}) = \mu^{-1} l_{m-1}$ and has the same properties of S.

Now, since we want to prove that any two transvections are conjugate in $SL(V)$, we want to show that

$$S_1[\phi, h]S_1^{-1} = [\psi, l]. \tag{4.9}$$

To show this, we will verify that $S_1[\phi, h]S_1^{-1}$ and $[\psi, l]$ have the same effect on all vectors in V.

First, we investigate their effect on the vector u. We get

$$(S_1[\phi, h]S_1^{-1})(u) = S_1([\phi, h](v)) = S_1(v + \phi(v)h) = S_1(v) + S_1(h) = u + l$$

since $S_1^{-1}(u) = v$, $\phi(v) = 1$ and $S_1(h) = l$. Now let's check the right-hand side:

$$[\psi, l](u) = u + \psi(u)l = u + l$$

since $\psi(u) = 1$.

Secondly, we investigate their effect on a general vector $l' \in L$. First, we check the right-hand side:

$$[\psi, l](l') = l'$$

since $[\psi, l]$ fixes all vectors in L. Now, on the left-hand side we get

$$(S_1[\phi, h]S_1^{-1})(l') = S_1([\phi, h](h')) = S_1(h') = l'$$

since S_1^{-1} maps l' to some element $h' \in H$ and $[\phi, h]$ fixes all vectors in H. Since (4.9) is valid on L, it in particular holds on the basis vectors $l, l_2, \ldots, l_{m-1}$ of L. Together with u, (4.9) holds on the basis vectors in B_2, hence (4.9) holds on V by linearity. Thus, since $[\phi, h]$ and $[\psi, l]$ were chosen to be any two transvections in V, we have that all transvections are conjugate in $SL(V)$.

□

Before the next theorem, we will introduce new notation. Let H be a hyperplane in a vector space V. Then let $\tau(H)$ denote the set of all transvections that fix the hyperplane H. Furthermore, let $1_V \in \tau(H)$. The next theorem shows that $\tau(H)$ is a group, and is, in fact, an Abelian subgroup of $SL(V)$.

Lemma 4.9. *Let V be an m-dimensional vector space over a field K, and let H be a hyperplane in V.*

(i) There is a linear functional $\phi \neq 0$ with $\ker \phi = H$ such that

$$\tau(H) = \{[\phi, h] : h \in H\} \cup \{1_V\}.$$

(ii) The group $\tau(H)$ is an Abelian subgroup of $SL(V)$. Furthermore, $\tau(H) \cong H$.

(iii) The centralizer $C_{SL(V)}(\tau(H))$ equals $SZ(V)\tau(H)$.

Proof. Before we start, we will prove that two linear functionsls ϕ and ψ have the same kernel if and only if $\psi = a\phi$ for some $a \in K^\times$.

First, assume $\psi = a\phi$ for some $a \in K^\times$. Now, $\phi(v)$ will take values in K, but since K is a field, it has no zero divisors. Thus, $a\phi$ is zero only when ϕ is zero, and so, ψ and $a\phi$ have the same kernel.

Assume now, instead, that $\ker \phi = \ker \psi = H$. We choose a vector $w \in V$ that is not in H. Then $\phi(w)$ and $\psi(w)$ are some non-zero elements in K. Since K is a field, there is an element $a \in K^\times$ such that $\psi(w) = a\phi(w)$. Now, if $v \in V$, then $v = \lambda w + h$ by the proof of Lemma 4.2, where $\lambda \in K$ and $h \in H$. Thus, we get

$$\psi(v) = \psi(\lambda w + h) = \lambda\psi(w) = a\lambda\phi(w) = a\phi(\lambda w + h) = a\phi(v) \tag{4.10}$$

which is what we wanted to prove.

Now, let $[\phi, h], [\psi, l] \in \tau(H)$ be two transvections fixing a hyperplane H in V. Then, by (4.10) and Lemma 4.5, we have that $[\psi, l] = [a\phi, l] = [\phi, al]$. If $[\phi, h] \in \tau(H)$ fixes a hyperplane H, then so does its inverse $[\phi, h] = [\phi, h]^{-1} = [\phi, -h]$, and thus, $[\phi, h]^{-1} \in \tau(H)$. Since $[\psi, l] \in \tau(H)$ fixes H, then so does its inverse $[\psi, l]^{-1} = [\phi, -al] \in \tau(H)$ and $[\phi, h] \circ [\phi, -al] = [\phi, h - al] \in \tau(H)$. So, since $a, b \in \tau(H)$ implied that $ab^{-1} \in \tau(H)$, and since $1_V \in \tau(H)$, we have that $\tau(H)$ is a group. Transvections have determinant 1, so $\tau(H) \leqslant SL(V)$. Also, $\tau(H)$ is Abelian, because (i) in Lemma 4.5 tells us that

$$[\phi, h] \circ [\phi, l] = [\phi, h + l] = [\phi, l + h] = [\phi, l] \circ [\phi, h].$$

(i) By the above discussion, if $[\psi, l]$ is in $\tau(H)$, then ψ has the same kernel as ϕ since they fix the same hyperplane H. Then $[\psi, l] = [\phi, h_1]$ for some $h_1 \in H$, and thus, all transvections in $\tau(H)$ can be written as $[\phi, h_1]$ for some $h_1 \in H$.

(*ii*) Furthermore, this form will be unique. Since all transvections in $\tau(H)$ can be expressed by a single ϕ but for different $h \in H$, there are as many elements of $\tau(H)$ as there are in H. Hence, $\tau(H) \cong H$.

(*iii*) Finally, we want to show that $SZ(V)\tau(H) = C_{SL(V)}(\tau(H))$ by showing that $SZ(V)\tau(H) \leqslant C_{SL(V)}(\tau(H))$ and $C_{SL(V)}(\tau(H)) \leqslant SZ(V)\tau(H)$. Since $\tau(H)$ is Abelian we have that $\tau(H) \leqslant C_{SL(V)}(\tau(H))$. Clearly, we see that $SZ(V) \leqslant C_{SL(V)}(\tau(H))$, so together we have $SZ(V)\tau(H) \leqslant C_{SL(V)}(\tau(H))$.

Now we want to show that $C_{SL(V)}(\tau(H)) \leqslant SZ(V)\tau(H)$. Assume there is an $S \in SL(V)$ that commutes with every transvection $[\phi, h]$. This means that $S[\phi, h] = [\phi, h]S$ for all $h \in H$, i.e. $S[\phi, h]S^{-1} = [\phi, h]$. Now, (*iv*) in Lemma 4.5 tells us that $S[\phi, h]S^{-1} = [\phi S^{-1}, Sh]$. But since $S[\phi, h]S^{-1} = [\phi, h]$, this means that $[\phi, h] = [\phi S^{-1}, Sh]$. We now use (*iii*) in Lemma 4.5 to see that

$$\phi S^{-1} = a\phi \tag{4.11}$$

and

$$h = aSh$$

for some $a \in K^{\times}$. Since $h = aSh$, we see that aS fixes the hyperplane H pointwise, so aS is either a transvection or a dilatation, or the identity.

Assume that aS is a dilatation. Then by Theorem 4.4, given a suitable basis for V, aS has an eigenvector v_1 outside of H with eigenvalue $\mu \neq 1$. This means that $aSv_1 = \mu v_1$, and furthermore, $\mu = \det aS = a^m \det S = a^m$ because $\det S = 1$. Plugging this into the previous equation, we get $aSv_1 = a^m v_1$, which implies $Sv_1 = a^{m-1}v_1$. Now, multiply by S^{-1} and $1/a^{m-1} = a^{-m+1}$ from the left to get $a^{-m+1}v_1 = S^{-1}v_1$. This gives $\phi S^{-1}v_1 = \phi(a^{-m+1}v_1) = a^{-m+1}\phi(v_1)$. Multiplying both sides by a^{m-1} and using (4.11), we get that

$$a^{m-1}\phi S^{-1}v_1 = a^{m-1}a\phi v_1 = a^m \phi v_1 = \phi v_1$$

which implies $a^m = 1$ since $w \notin H$ implies $\phi(w) \neq 0$. But this is a contradiction since $1 \neq \mu = a^m$, and so, aS cannot be a dilatation.

Clearly, the identity operator belongs to $SZ(V)\tau(H)$, so assume that aS is a transvection. Then $aS \in \tau(H)$. Now, note that $a^{-1}1_V \in SZ(V)$ because $(a^{-1})^m = (a^m)^{-1} = 1^{-1} = 1$. Thus, $S = a^{-1}1_V(aS) \in SZ(V)\tau(H)$. So if $S \in SL(V)$ commutes with every element in $\tau(H)$, i.e. $S \in C_{SL(V)}(\tau(H))$, then $S \in SZ(V)\tau(H)$.

Thus, it follows that $C_{SL(V)}(\tau(H)) \leqslant SZ(V)\tau(H)$ and we have that $SZ(V)\tau(H) = C_{SL(V)}(\tau(H))$.

□

Theorem 4.10. *(Jordan–Dickson) Let V be an m-dimensional vector field over a field K, where $m \geqslant 3$. Then the group $PSL(V)$ is a simple group.*

Proof. The proof of this theorem will be similar to the proof Theorem 3.5. Let $\phi : SL(V) \to PSL(V)$ denote the cannonical homomorphism. We want to show that if $PSL(V)$ contains a normal subgroup $M \neq \{e\}$, then we must have $M = PSL(V)$. We will do this by showing that if $N = \phi^{-1}(M)$ is a normal

subgroup of $SL(V)$ that contains some linear transformation A which is not in $SZ(V)$, then N must be $SL(V)$. Note that Theorem 3.4 ensures that if M is normal, then $N = \phi^{-1}(M)$ is normal. Then, when we have showed that $N = SL(V)$, we apply the canonical homomorphism ϕ to $N = SL(V)$, which maps it to $N/SZ(V) = SL(V)/SZ(V) = PSL(V)$. Thus, in the same way as in Theorem 3.5, by controlling our pre-image $N = \phi^{-1}(M)$ and showing that it must be all of $SL(V)$, we force M to be all of $PSL(V)$.

So, assume N is a normal subgroup of $SL(V)$ containing some linear transformation A which is not in $SZ(V)$. We want to show that N contains a transvection. Since N contains a linear transformation A which is not in $SZ(V)$, then there exists a transvection T_0 that does not commute with A. To see this, note that $SZ(V)$ is all the linear transformations in $SL(V)$ that commutes with all elements of $SL(V)$, and $A \notin SZ(V)$. Thus there exists some linear transformation $B \in SL(V)$ that does not commute with A. Also, $SL(V)$ is generated by transvections, so all elements in $SL(V)$ can be written as a product of transvections. In particular, B can be written as a product of transvections. For simplicity, let $B = T_4T_5T_6$. Since $AB \neq BA$, this means that $AT_4T_5T_6 \neq T_4T_5T_6A$. In fact, our linear transformation A can commute with some of these transvections, but not all, else we would have equality. Thus, there exists a transvection T_0 that does not commute with A.

Now, form the commutator $C = T_0^{-1}A^{-1}T_0A$, and note that $C \neq 1$ since A does not commute with T_0. Furthermore, $C \in N$ since $T_0^{-1}A^{-1}T_0 \in N$ due to normality of N and $A \in N$. It is easy to show that a conjugate $A^{-1}T_0A$ of a transvection is a transvection. Set $T_2 = A^{-1}T_0A \in SL(V)$. Since the inverse of a transvection is a transvection, $T_0^{-1} = T_1$ is a transvection and thus, we have that

$$C = T_0^{-1}A^{-1}T_0A = T_1T_2.$$

We saw earlier in this section that if T_1, T_2 are transvections, then $T_1 = [\phi_1, h_1]$ and $T_2 = [\phi_2, h_2]$ where $h_1 \in H_1 = \ker\phi_1$ and $h_2 \in H_2 = \ker\phi_2$. We remind the reader that this means that

$$T_i(v) = v + \phi_i(v)h_i$$

for all $v \in V$ and $i = 1, 2$.

Now, let W be the subspace spanned by h_1, h_2, i.e. $W = \langle h_1, h_2\rangle$. The dimension of W can be no more than 2, so $\dim W \leqslant 2$. By assumption, we have that $\dim V \geqslant 3$, so there exists a hyperplane L in V that contains W. If $l \in L$, then

$$C(l) = T_1T_2(l) = T_2(l) + \phi_1(T_2(l))h_1 = l + \phi_2(l)h_2 + \phi_1(T_2(l))h_1.$$

Note that $(\phi_2(l)h_2 + \phi_1(T_2(l))h_1) \in W \subseteq L$ and $l \in L$, so we see that $C(l) \in L$.

We will now show that $H_1 \cap H_2 \neq \{0\}$. If $H_1 = H_2$, this is clearly true, so assume $H_1 \neq H_2$. Since hyperplanes are, by definition, $(m-1)$-dimensional subspaces of the vector space V, we have $H_1 + H_2 = V$ since $\dim(H_1 + H_2) = m$. By linear algebra we have

$$\dim H_1 + \dim H_2 = \dim(H_1 + H_2) + \dim(H_1 \cap H_2).$$

In our case, the equation becomes

$$(m-1)+(m-1) = m + dim(H_1 \cap H_2)$$

and we see that $\dim(H_1 \cap H_2) = m - 2 \geqslant 1$ since $m \geqslant 3$.

Now, if z is a non-zero vector in $H_1 \cap H_2$, then

$$C(z) = T_1 T_2(z) = z$$

since T_1 fixes all vectors in H_1 and T_2 fixes all vectors in H_2. If $C \in N$ is a transvection, then $N = SL(V)$ because $SL(V)$ is generated by transvections, and we are done. So we assume C is not a transvection, that is, assume $C \notin \tau(L)$. We wish to show that $C \notin SZ(V)\tau(L) = C_{SL(V)}\tau(L)$. For a contradiction, assume that $C = a1_V S$ for some $S \in \tau(L)$, where $a \in K^\times$. Then $z = Cz = aSz$ gives that $Sz = a^{-1}z$, and we see that z is an eigenvector of S. Since all eigenvectors of transvections lie in the hyperplane they fix, and all eigenvalues are 1, we have that $z \in L$ and $a = a^{-1} = 1$, and thus $S = C$. But $C = S \in \tau(L)$ and $S = C \notin \tau(L)$ is a contradiction. Therefore, $C \neq aS$, and it follows that $C \notin SZ(V)\tau(L) = C_{SL(V)}(\tau(L)$ since $SZ(V)$ are all the elements in $SL(V)$ of the form $a1_V$ where $a \in K^\times$.

By the same argument as in the start of this proof, there must therefore exist a transvection $T \in \tau(L)$ that does not commute with C. Thus, we have that the commutator D is not 1:

$$D = TCT^{-1}C^{-1} \neq 1.$$

Now, $C^{-1} \in N$ and again, by normality of N, we have that $TCT^{-1} \in N$, and so $D \in N$. If $l \in L$ is some vector in L, then

$$D(l) = TCT^{-1}C^{-1}(l) = TCT^{-1}(C^{-1}(l)) = TC(C^{-1}(l)) = T(l) = l$$

because $T \in \tau(L)$ fixes all vectors in L and $C^{-1}(l) \in L$. But this means that D fixes all vectors in L, so D is either a transvection or a dilatation. By the rule of determinants, D has determinant 1, so D cannot be a dilatation. Hence, D is a transvection, and we have shown that N contains a transvection. Therefore, $\phi^{-1}(M) = N = SL(V)$ because $SL(V)$ is generated by transvections. By the discussion in the beginning of this proof, this means that the normal subgroup $M \leqslant PSL(V)$ in fact is $PSL(V)$. Thus, $PSL(V)$ has no non-trivial normal subgroups, so $PSL(V)$ is simple, and we are done.

□

References

[1] J. B. Fraleigh, *A First Course in Abstract Algebra*, Seventh Edition, Pearson Education, 2003.

[2] T. W. Hungerford, *Abstract Algebra, An Introduction*, Third Edition, Brooke/Cole Cengage Learning, 2014.

[3] J. J. Rotman, *An Introduction to the Theory of Groups*, Fourth Edition, Graduate Texts in Mathematics vol. 148, Springer Verlag, 1995.