Why is PSL(2, 7) ≅ GL(3, 2)?

Ezra Brown and Nicholas Loehr

1. INTRODUCTION. The groups of invertible matrices over finite fields are among the first groups we meet in a beginning course in modern algebra. Eventually, we find out about simple groups and that the unique simple group of order 168 has two representations as a group of matrices. And this is where we learn that the group of 2×2 unimodular matrices over a seven-element field, with I and $-I$ identified, is isomorphic to the group of invertible 3×3 matrices over a 2-element field. In short, it is a fact that $\mathrm{PSL}(2, 7) \cong \mathrm{GL}(3, 2)$.

Many of us are surprised by this fact: why should a group of 2×2 matrices with mod-7 integer entries be isomorphic to a group of 3×3 binary matrices?

There are a number of proofs of this remarkable theorem. Dickson [**1**, p. 303] gives a proof based on his general theorem giving uniform sets of generators and relations for the family of groups $\mathrm{SL}(2, q)$, where q is any prime power. One checks that the relations appearing in Dickson's presentation of PSL(2, 7) are satisfied by certain generators of GL(3, 2), implying that these groups have the same presentations and are therefore isomorphic. Dummit and Foote [**2**, p. 207–212] show that every simple group of order 168 is necessarily isomorphic to the automorphism group $\mathrm{Aut}(\mathcal{F})$ of the Fano plane $\mathcal{F}$. They then show that $\mathrm{Aut}(\mathcal{F}) \cong \mathrm{GL}(3, 2)$ and that PSL(2, 7) is a simple group of order 168; the isomorphism theorem follows. Rotman gives the result as an exercise [**5**, Exercise 9.26, p. 281]. A hint is to begin with a simple group G of order 168 and use the seven conjugates of a Sylow 2-subgroup P of G to construct a seven-point projective plane; the proof is similar to Dummit and Foote's proof. Jeurissen [**4**] proves the result by showing that both PSL(2, 7) and GL(3, 2) are subgroups of index 2 of the automorphism group of a Coxeter graph. Elkies [**3**] gives a clever proof that uses the automorphism group G of the 3-(8, 4, 1) Steiner system—also known as the Steiner $S(3, 4, 8)$ design. He shows that PSL(2, 7) is contained in G, which in turn maps homomorphically onto GL(3, 2). The result follows from the simplicity of the two groups and the fact that they are both of order 168. We remark that there do exist non-isomorphic simple groups of the same order. For example, Schottenfels showed that PSL(3, 4) and A_8 are non-isomorphic simple groups of order 20,160 [**5**, Theorem 8.24, p. 233].

The aim of this paper is to give a proof that $\mathrm{PSL}(2, 7) \cong \mathrm{GL}(3, 2)$ that is elementary in the sense that it uses neither simplicity, nor projective geometry, nor block designs. We will *not* prove the fact that any two simple groups of order 168 are isomorphic, nor will we use this fact in our proof. What makes our proof work is that: (a) we can identify GL(3, 2) with the set of invertible $\mathbb{F}_2$-linear transformations on the finite field with eight elements; (b) $7 = 2^3 - 1$; (c) the nonzero squares mod 7 are precisely the powers of 2 mod 7; (d) squaring mod 2 is additive (the Freshman's Dream); and (e) the mapping $k \mapsto -1/k$ mod 7 translates to a bit-switch mod 2 — which is linear. We begin by giving functional descriptions for both groups, determining their sizes,

doi:10.4169/193009709X460859

and exhibiting sets of generators for them. After this we define a mapping between the groups and prove that the mapping is a bijective group homomorphism.

Let's begin with GL(3, 2).

2. THE GROUP GL(3, 2). Let $\mathbb{F}_2 = \{0, 1\}$ be the field with two elements. The group GL(3, 2) consists of all invertible 3×3 matrices with entries in $\mathbb{F}_2$. To construct our isomorphism, we need three basic facts about this group.

1. *Functional Description of the Group.* Let $\mathbb{F}_8 = \mathbb{F}_2[X]/\langle X^3 + X + 1\rangle$, and let $x = X + \langle X^3 + X + 1\rangle \in \mathbb{F}_8$. On one hand, $\mathbb{F}_8$ is an eight-element field whose multiplicative group is generated by x. On the other hand, $\mathbb{F}_8$ is a three-dimensional vector space over $\mathbb{F}_2$ with ordered basis $B = (x^0, x^1, x^2)$. Let $\mathrm{GL}(\mathbb{F}_8)$ be the set of all invertible $\mathbb{F}_2$-linear transformations of this vector space. This means that $\mathrm{GL}(\mathbb{F}_8)$ consists of all bijections $L : \mathbb{F}_8 \to \mathbb{F}_8$ such that $L(u + v) = L(u) + L(v)$ for all $u, v \in \mathbb{F}_8$. We note that $L(cu) = cL(u)$ holds automatically, since the only available scalars c are 0 and 1. Let $[L]_B$ denote the matrix of L relative to the ordered basis B. Then the map $L \mapsto [L]_B$ defines an isomorphism between $\mathrm{GL}(\mathbb{F}_8)$ and GL(3, 2). From now on, we identify these two groups by means of this isomorphism.
2. *The Size of* GL(3, 2). The following counting argument proves that $|\,\mathrm{GL}(3, 2)| = 168$. Let us build an invertible 3×3 matrix of 0s and 1s one row at a time. The first row can be any nonzero bit string of length 3; there are seven such bit strings. The second row can be any nonzero bit string different from the first row; there are six such bit strings. When choosing the third row, we must pick a bit string that is not a linear combination of the first two rows. There are four such linear combinations (zero, the first row, the second row, or the sum of the first two rows), so there are $8 - 4 = 4$ choices for the third row. By the product rule,
$$|\mathrm{GL}(3, 2)| = 7 \cdot 6 \cdot 4 = 168.$$
3. *Generators for* GL(3, 2). It will be useful to have a small set of generators for GL(3, 2). Starting with any matrix $A \in \mathrm{GL}(3, 2)$, we can use elementary row operations (Gaussian elimination) to reduce A to the identity matrix. Each elementary operation can be accomplished by multiplying on the left by one of the following nine elementary matrices:
$$E_{12} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad E_{13} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad E_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix},$$
$$E_{21} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad E_{31} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \quad E_{32} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix},$$
$$S_{12} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad S_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad S_{13} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$
For example, multiplying A on the left by E_{12} adds the the second row of A to the first row, whereas multiplying A on the left by S_{13} interchanges the first and third rows of A. Thus, the row-reduction of A to the identity matrix via elementary row operations translates to a matrix equation of the form $E_1 \cdots E_k A = I$,

where each E_i is an elementary matrix. Solving for A and noting that each elementary matrix equals its own inverse, we see that GL(3, 2) is generated by the nine elementary matrices listed above. In fact, many of these matrices are redundant, and the set $\{E_{23}, S_{12}, S_{23}\}$ already generates the whole group. This remark follows from the formulas

$$S_{13} = S_{12}S_{23}S_{12}, \quad E_{13} = S_{12}E_{23}S_{12}, \quad E_{32} = S_{23}E_{23}S_{23},$$
$$E_{21} = S_{13}E_{23}S_{13}, \quad E_{12} = S_{13}E_{32}S_{13}, \quad E_{31} = S_{12}E_{32}S_{12}.$$

Now consider the three matrices

$$A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

We have $E_{23} = A_1A_3$, $S_{12} = A_2^2A_1A_2^3A_3$, $S_{23} = A_1$, and so $\mathrm{GL}(3, 2) = \langle A_1, A_2, A_3\rangle$.

And now, on to a description of PSL(2, 7).

3. THE GROUP PSL(2, 7). Let $\mathbb{F}_7 = \{0, 1, 2, 3, 4, 5, 6\}$ be the field with seven elements. The group SL(2, 7) consists of all 2 × 2 matrices with entries in $\mathbb{F}_7$ and determinant 1. The group PSL(2, 7) is the quotient group $\mathrm{SL}(2, 7)/\{I, -I\}$. To construct our isomorphism, we need three basic facts about this group.

1. *Functional Description of the Group.* Let

$$\overline{\mathbb{F}_7} = \mathbb{F}_7 \cup \{\infty\} = \{0, 1, 2, 3, 4, 5, 6, \infty\}.$$

As in complex analysis, we define a *linear fractional transformation* on $\overline{\mathbb{F}_7}$ to be a function $f : \overline{\mathbb{F}_7} \to \overline{\mathbb{F}_7}$ of the form

$$f(k) = \frac{ak + b}{ck + d} \qquad (k \in \overline{\mathbb{F}_7}), \tag{1}$$

where $a, b, c, d \in \mathbb{F}_7$ are constants such that $ad - bc \neq 0$. (The same definition works for any field $\mathbb{F}$.) In the formula for $f(k)$, division by $ck + d$ means multiplication by the inverse of $ck + d$ in the field $\mathbb{F}_7$; any nonzero element divided by 0 is ∞; and anything (other than ∞) divided by ∞ is 0. We have $f(\infty) = a/c$ when $c \neq 0$, and $f(\infty) = \infty$ when $c = 0$. The transformation f is called *special* if a, b, c, d can be chosen so that $ad - bc = 1$. There is a natural map ϕ from SL(2, 7) to the set SLF(7) of special linear fractional transformations on $\overline{\mathbb{F}_7}$, which sends the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ to the function f given in (1). A routine calculation shows that $\phi(A) \circ \phi(B) = \phi(AB)$ (which says composing linear fractional transformations is done by matrix multiplication), so that ϕ is a group homomorphism. Furthermore, one sees that $\phi(A) = \phi(B)$ iff $B = A$ or $B = -A$. It follows that $\ker(\phi) = \{I, -I\}$, so that ϕ induces a group *isomorphism* from PSL(2, 7) onto SLF(7). Henceforth, we identify these two groups by means of this isomorphism.

2. *The Size of* PSL(2, 7). The following counting argument proves that $|\mathrm{SL}(2, 7)| = 336$. Let us build a matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with entries in $\mathbb{F}_7$ such that $ad - bc = 1$.

There are two cases: $c = 0$ or $c \neq 0$. In the case where $c = 0$, choose a to be any nonzero element (6 possibilities); then choose $d = a^{-1}$ to force the determinant to be 1 (one possibility); then choose b to be anything (7 possibilities). This gives 42 upper-triangular matrices in SL(2, 7). In the case where $c \neq 0$, choose c (6 possibilities); then choose a and d arbitrarily (7 possibilities each); then we must choose $b = c^{-1}(ad - 1)$ to get the right determinant. This gives $6 \cdot 7 \cdot 7 = 294$ more matrices, for a total of 336. Taking the quotient by the two-element subgroup $\{I, -I\}$ cuts the number of group elements in half, so $|\operatorname{PSL}(2, 7)| = 336/2 = 168$.

3. *Generators for* PSL(2, 7). We can use the functional description of PSL(2, 7) to find a convenient set of generators for this group. We define three special linear fractional transformations r, t, δ by setting

   - $r(k) = -1/k$ (the "reflection map");
   - $t(k) = k + 1$ (the "translation map"); and
   - $\delta(k) = 2k$ (the "doubling map").

   We will prove that $\operatorname{SLF}(7) = \langle r, t, \delta \rangle$. Consider a special linear fractional transformation $f(k) = (ak + b)/(ck + d)$. If $c = 0$, we must have $ad = 1$ and $d = a^{-1}$, so $f(k) = a^2k + ab$. The nonzero squares mod 7 are 1, 2, 4, so $f = t^{ab} \circ \delta^j$ for a suitable $j \in \{0, 1, 2\}$. For example, given $f(k) = (3k + 6)/5$, we have $f(k) = 2k + 4 = t(t(t(t(\delta(k)))))$. Next, consider $f(k) = (ak + b)/(ck + d)$ where $c \neq 0$. Division gives

$$f(k) = (ac^{-1}) + \frac{bc - ad}{c(ck + d)} = (ac^{-1}) + \frac{-1}{c^2k + cd}.$$

   Writing $c^2 = 2^j$, it is now evident that $f = t^{ac^{-1}} \circ r \circ t^{cd} \circ \delta^j$. Hence, SLF(7) is generated by r, t, and δ.

4. THE ISOMORPHISM PSL(2, 7) ≅ GL(3, 2). We now have all the ingredients needed to define the promised group isomorphism between PSL(2, 7) and GL(3, 2). Using the functional descriptions of these groups, it will suffice to define an isomorphism $T : \operatorname{SLF}(7) \to \operatorname{GL}(\mathbb{F}_8)$. We proceed in four stages.

1. *Definition of T.* For each function $f \in \operatorname{SLF}(7)$, we need to define an associated function $T(f) = T_f \in \operatorname{GL}(\mathbb{F}_8)$. How can we use the function f, whose domain is $\overline{\mathbb{F}_7}$, to build a function T_f, whose domain is $\mathbb{F}_8$? To relate these two domains, we define $x^\infty = 0$ and then observe that $\mathbb{F}_8 = \{x^k : k \in \overline{\mathbb{F}_7}\}$. This observation suggests the map $x^k \mapsto x^{f(k)}$ as a possibility for T_f. However, this map is not always *linear*, since zero maps to zero only if $f(\infty) = \infty$. To account for this difficulty, we instead define

$$T_f(x^k) = x^{f(k)} + x^{f(\infty)} \qquad (k \in \overline{\mathbb{F}_7}). \tag{2}$$

   With this definition, $T_f(0) = 0$ always holds, though it is not yet evident that T_f must belong to $\operatorname{GL}(\mathbb{F}_8)$.

   Let us illustrate the formula by computing $T(r)$, $T(t)$, and $T(\delta)$. This computation will reveal that *each of these three functions does indeed lie in* $\operatorname{GL}(\mathbb{F}_8)$. The function $r(k) = -1/k$ is given in two-line form as follows:

$$r = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & \infty \\ \infty & 6 & 3 & 2 & 5 & 4 & 1 & 0 \end{pmatrix}.$$

Since $x^3 + x + 1 = 0$ in $\mathbb{F}_8$, we have

$$x^3 = x + 1, \quad x^4 = x^2 + x, \quad x^5 = x^2 + x + 1, \quad x^6 = x^2 + 1.$$

Let us represent an element $b_2 x^2 + b_1 x^1 + b_0 x^0$ in $\mathbb{F}_8$ by the bit string $b_2 b_1 b_0$. In this notation,

$$x^0 = 001,\ x^1 = 010,\ x^2 = 100,\ x^3 = 011,$$

$$x^4 = 110,\ x^5 = 111,\ x^6 = 101,\ x^\infty = 000.$$

Putting all this information into (2), we conclude that

$$T(r) = \begin{pmatrix} 001 & 010 & 100 & 011 & 110 & 111 & 101 & 000 \\ 001 & 100 & 010 & 101 & 110 & 111 & 011 & 000 \end{pmatrix}.$$

Note that T_r just interchanges the first two bits. Thus, T_r is the invertible linear map on $\mathbb{F}_8$ that interchanges the basis vectors x^1 and x^2, and so the matrix of T_r relative to the ordered basis $B = (x^0, x^1, x^2)$ is

$$[T(r)]_B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = A_1.$$

It is even easier to compute $T(t)$ and $T(\delta)$. We have $T_t(0) = 0$ and, for $k \neq \infty$,

$$T_t(x^k) = x^{t(k)} + x^{t(\infty)} = x^{k+1} = x(x^k).$$

Thus, T_t is simply left-multiplication by x in the field $\mathbb{F}_8$. This map is linear by the distributive law in $\mathbb{F}_8$, and it is invertible since the inverse map is left-multiplication by $x^{-1} = x^6$. The matrix of T_t is

$$[T(t)]_B = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = A_2.$$

Finally, $T_\delta(0) = 0$ and, for $k \neq \infty$, $T_\delta(x^k) = x^{2k} = (x^k)^2$. So T_δ is the squaring map in $\mathbb{F}_8$. This map is linear (and even a ring homomorphism) since $(u + v)^2 = u^2 + 2uv + v^2 = u^2 + v^2$ for $u, v \in \mathbb{F}_8$. The map is one-to-one (hence onto) since the kernel is zero. The matrix of T_δ is

$$[T(\delta)]_B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} = A_3.$$

2. *The key lemma.* Suppose $f, g \in \mathrm{SLF}(7)$ are two functions such that $T(f)$ and $T(g)$ lie in $\mathrm{GL}(\mathbb{F}_8)$. Then $T(f \circ g) = T(f) \circ T(g)$, and hence $T(f \circ g)$ also lies in $\mathrm{GL}(\mathbb{F}_8)$.

   Proof: For any $k \in \overline{\mathbb{F}_7}$, we compute

$$T_f \circ T_g(x^k) = T_f(T_g(x^k)) = T_f(x^{g(k)} + x^{g(\infty)})$$

$$\begin{aligned}
&= T_f(x^{g(k)}) + T_f(x^{g(\infty)}) \quad \text{(since } T_f \text{ is linear)}\\
&= (x^{f(g(k))} + x^{f(\infty)}) + (x^{f(g(\infty))} + x^{f(\infty)})\\
&= x^{f(g(k))} + x^{f(g(\infty))} \quad \text{(since } u + u = 0 \text{ for all } u \in \mathbb{F}_8)\\
&= T_{f\circ g}(x^k).
\end{aligned}$$

3. *Proof that T is a homomorphism mapping into* $\mathrm{GL}(\mathbb{F}_8)$. We have seen that each element of SLF(7) can be written as a product of the generators r, t, and δ (using only positive powers, in fact). Since $T(r)$, $T(t)$, and $T(\delta)$ are known to lie in $\mathrm{GL}(\mathbb{F}_8)$, repeated application of the lemma shows that $T(h)$ lies in $\mathrm{GL}(\mathbb{F}_8)$ for *all* $h \in$ SLF(7). Having drawn this conclusion, the lemma now shows that T is a group homomorphism.
4. *Proof that T is a bijection.* So far, we know that T is a group homomorphism mapping SLF(7) *into* $\mathrm{GL}(\mathbb{F}_8)$. T is actually *onto*, since the image of T contains $\langle T(r), T(t), T(\delta)\rangle$, which is the whole group $\mathrm{GL}(\mathbb{F}_8)$. Since SLF(7) and $\mathrm{GL}(\mathbb{F}_8)$ both have 168 elements, T must also be one-to-one.

Our proof that $\mathrm{PSL}(2, 7) \cong \mathrm{GL}(3, 2)$ is now complete. We leave it as a challenge for the reader to find an explicit description of the inverse bijection $T^{-1} : \mathrm{GL}(\mathbb{F}_8) \to$ SLF(7).

REFERENCES

1. L. E. Dickson, *Linear Groups*, B. G. Teubner, Leipzig, 1901.
2. D. S. Dummit and R. M. Foote, *Abstract Algebra*, 3rd ed., John Wiley, Hoboken, NJ, 2004.
3. N. Elkies, Handout for Math 155 (1998), available at `http://www.math.harvard.edu/~elkies/M155.98/h4.ps`.
4. R. H. Jeurissen, A proof by graphs that $\mathrm{PSL}(2, 7) \cong \mathrm{GL}(3, 2)$, *Discrete Math.* **70** (1988) 315–317. `doi:10.1016/0012-365X(88)90008-8`
5. J. J. Rotman, *An Introduction to the Theory of Groups*, 4th ed., Springer-Verlag, New York, 1995.

Department of Mathematics, Virginia Tech, Blacksburg, VA 24061-0123
brown@math.vt.edu
nloehr@vt.edu

Angles as Probabilities

David V. Feldman and Daniel A. Klain

Almost everyone knows that the inner angles of a triangle sum to 180°. But if you ask the typical mathematician how to sum the solid inner angles over the vertices of a tetrahedron, you are likely to receive a blank stare or a mystified shrug. In some cases you may be directed to the Gram-Euler relations for higher-dimensional polytopes [**4, 5, 7, 8**], a 19th-century result unjustly consigned to relative obscurity. But the answer is really much simpler than that, and here it is:

> The sum of the solid inner vertex angles of a tetrahedron T, divided by 2π, gives the probability that the orthogonal projection of T onto a random 2-plane is a triangle.

doi:10.4169/193009709X460868