ILLINOIS JOURNAL OF MATHEMATICS
Volume 22, Number 1, March 1978

# AUTOMORPHISM SEQUENCES OF INTEGER UNIMODULAR GROUPS

BY

JOAN L. DYER

Let $\mathscr{A}(G)$ denote the automorphism group of the group $G$, and let $I: G \to \mathscr{A}(G)$ be the homomorphism which assigns to $g \in G$ the inner automorphism

$$I(g): x \to gxg^{-1} \quad (\text{all } x \in G).$$

This procedure may be iterated to give rise to an *automorphism sequence*

$$G \xrightarrow{I} \mathscr{A}(G) \xrightarrow{I} \mathscr{A}(\mathscr{A}(G)) = \mathscr{A}^2(G) \xrightarrow{I} \cdots.$$

Such a sequence *stabilizes* in finitely many steps if the maps

$$I: \mathscr{A}^i(G) \to \mathscr{A}^{i+1}(G)$$

are isomorphisms for all sufficiently large integers $i$; that is, $\mathscr{A}^i(G)$ has a trivial center and only inner automorphisms for all sufficiently large $i$. Such groups are termed *complete*. Finite stability need not occur, even when $G$ is assumed to be linear. For the infinite dihedral group $D$, each

$$I: \mathscr{A}^i(G) \to \mathscr{A}^{i+1}(G)$$

is a monomorphism with $\mathscr{A}^{i+1}(G)/I(\mathscr{A}^i(G))$ of order two (Hulse [7]). The main result of this paper is:

THEOREM A. *The automorphism sequences of the groups $SL(n, \mathbf{Z})$ and $GL(n, \mathbf{Z})$ stabilize in finitely many steps.*

When $G$ is $SL(n, \mathbf{Z})$ or $GL(n, \mathbf{Z})$, the automorphism group $\mathscr{A}(G)$ is known (Hua and Reiner [5], Wan [17]). Moreover, $G$ has almost all automorphisms inner, and almost has a trivial center. Thus the conclusion of Theorem A is a natural one, and the automorphism sequences of these groups might be expected to stabilize very quickly. However the situation is surprisingly complicated for the general linear group when $n$ is even, as well as for the special linear group when $n = 2$.

We first establish:

THEOREM B. *For $n \geq 2$, $\mathscr{A}(PGL(n, \mathbf{Z}))$ is complete; that is*

$$I: \mathscr{A}(PGL(n, \mathbf{Z})) \to \mathscr{A}^2(PGL(n, \mathbf{Z}))$$

*is an isomorphism.*

Received March 23, 1976.

For $n \geq 3$, Theorem B follows easily from Solazzi's more general results on the automorphisms of projective groups containing enough transvections ([15], or see [13]). These results are obtained by extending O'Meara's residual space techniques, are valid for a broad class of rings, but exclude $n = 2$. The first determination of $\mathscr{A}(PGL(n, \mathbf{Z}))$ is that of Hua and Reiner [6], utilizing what has been termed the method of involutions and including $n = 2$. In that paper, the automorphisms of $PGL(n, \mathbf{Z})$ that carry $PSL(n, \mathbf{Z})$ to itself are determined. This is indeed the full automorphism group for $n \geq 3$ as asserted in [6], but is only a subgroup of index 2 in $\mathscr{A}(PGL(2, \mathbf{Z}))$. This error appears not to have been noted previously.

We next obtain:

THEOREM C. *If* $n \geq 3$, $\mathscr{A}(SL(n, \mathbf{Z}))$ *is complete.*

Since $GL(n, \mathbf{Z})$ decomposes as the direct sum of $SL(n, \mathbf{Z})$ and its center $\{I_n, -I_n\}$ for odd $n$, $PGL$, $GL$, and $SL$ have the same automorphism groups in this case. Thus for Theorem A, there remains the general linear group in even dimensions, and the exceptional $n = 2$. The table which follows summarizes the remainder of the computations. The entry in row $\mathscr{A}^i$, column $G$ is a pair of numbers: the first is the index of $I(\mathscr{A}^{i-1}(G))$ in $\mathscr{A}^i(G)$ and the second is the order of the center of $\mathscr{A}^i(G)$. We remark that these centers are all elementary abelian 2-groups. The last entry in a column is the first $\mathscr{A}^i(G)$ for which $I$: $\mathscr{A}^i(G) \to \mathscr{A}^{i+1}(G)$ is an isomorphism.

| | $GL(n, \mathbf{Z})$, $n \geq 4$ and even | $GL(2, \mathbf{Z})$ | $SL(2, \mathbf{Z})$ |
|---|---|---|---|
| $\mathscr{A}$ | $2^2; 2$ | $2^2; 2^2$ | $2^2; 2$ |
| $\mathscr{A}^2$ | $2^2; 2^2$ | $2^6 \cdot 3; 1$ | $2^3; 2$ |
| $\mathscr{A}^3$ | $2^5 \cdot 3; 1$ | $2; 1$ | $2^3; 2^2$ |
| $\mathscr{A}^4$ | $2 \cdot 3; 1$ | $2; 1$ | $2^7; 2^4$ |
| $\mathscr{A}^5$ | $2; 1$ | | $2^{29} \cdot 3^2 \cdot 5 \cdot 7; 2$ |
| $\mathscr{A}^6$ | | | $2^6 \cdot 3 \cdot 7; 2^3$ |
| $\mathscr{A}^7$ | | | $2^{12} \cdot 3 \cdot 7; 1$ |
| $\mathscr{A}^8$ | | | $2^6 \cdot 3 \cdot 7; 1$ |

The automorphism groups of $GL(n, R)$ and $SL(n, R)$, when $n \geq 3$ and $R$ is any integral domain, have been determined by O'Meara [12]; those of the projective groups by Solazzi [15]. I conjecture that the automorphism sequences of these groups are finite whenever the automorphism group of the ring $R$ has a finite automorphism sequence.

This paper is organized as follows: in Section 1, after establishing notation, we state some consequences of the decomposition of $\mathscr{A}(G)$ relative to a split short exact sequence

$$1 \to K \to G \to Y \to 1$$

where $K$ is a characteristic subgroup of $G$. These results serve to organize the subsequent computations. In Section 2 we obtain Theorems A, B, and C for $n \geq 3$. Section 3 deals with the automorphisms of $V$ and the group of derivations of $SL(n, \mathbf{Z}_2)$ in $V$, for certain $SL(n, \mathbf{Z}_2)$-modules $V$. These results are required in Section 4; the modules which arise are the (additive) group $M(n, m)$ of $n \times m$ matrices over $\mathbf{Z}_2$ on which $SL(n, \mathbf{Z}_2)$ acts by left matrix multiplication, and the (additive) groups $M(n)$, $M(n)/\{0_n, I_n\}$ on which $SL(n, \mathbf{Z}_2)$ acts by conjugation. The results of this section imply that $H^1(SL(n, \mathbf{Z}_2), V) = 0$ for $n \neq 3$ (see also [4], [9]). Proofs are direct and elementary, proceeding from the Steinberg presentation of $SL(n, \mathbf{Z}_2)$ (see [16], [10]). The fourth and final section is devoted to the case $n = 2$. We obtain $\mathscr{A}(PGL(2, \mathbf{Z}))$, correcting the error in [6], and then Theorems A and B. It may be of interest to note the correction required in [6]: in the notation of that paper, case b of Theorem 2 cannot be eliminated (the assertion "whence $(S_1 T_1^2)^3 = \pm I$" on p. 469, lines 1 and 2, is false). This case does arise, and leads to an exceptional automorphism defined in terms of the generators

$$S = \pm\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad T = \pm\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad B = \pm\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

of $PGL(2, \mathbf{Z})$ by

$$S \to SB, \quad T \to TB, \quad B \to B.$$

Thus Theorem 2, the Corollary to Theorem 3, and Theorem 4 need modification for $n = 2$. A computation along the lines indicated in the proof of Theorem 4 establishes that $\mathscr{A}(PGL(4, \mathbf{Z}))$ is as asserted and then the induction proceeds as given.

Results are numbered consecutively within a section, so that 2.3 refers to the third result of Section 2.

## 1. Extensions and complete groups

For any group $G$ and any $x, y \in G$ we write $[x, y] = xyx^{-1}y^{-1}$. $H < G$ means that $H$ is a subgroup of $G$, and $H \lhd G$ that $H$ is a normal subgroup of $G$. $[G: H]$ is the index of $H$ in $G$, and $|G|$ the order of $G$. For any subset $S \subset G$, $\langle S \rangle$ is the subgroup generated by $S$ and nm $\langle S \rangle$ is the normal subgroup generated by $S$. For subsets $S$, $T$ of $G$, $[S, T]$ is the *subgroup* generated by all $[s, t]$ with $s \in S$, $t \in T$. The commutator or derived subgroup of $G$ is $[G, G]$, also denoted by $G'$; $G$ is *perfect* if $G = G'$. $\mathscr{C}_G(S)$ is the centralizer of $S$ in $G$:

$$\mathscr{C}_G(S) = \{g \in G \,|\, [g, s] = 1 \text{ for all } s \in S\}.$$

The center of $G$ is $\mathscr{C}_G(G)$, which will be written as $\mathscr{C}(G)$. $\mathscr{N}_G(S)$ is the normalizer of $S$ in $G$.

For $g \in G$, $\alpha \in \mathscr{A}(G)$ write $[g, \alpha] = g\alpha(g)^{-1}$. Then for $T \subset G$ and $S \subset \mathscr{A}(G)$, $[T, S]$ is the subgroup of $G$ generated by all $[g, \alpha]$ with $g \in T$, $\alpha \in S$; and $\mathscr{C}_G(S)$

is the fixed point set of $S$ in $G$:

$$\mathscr{C}_G(S) = \{g \in G \mid [g, \alpha] = 1 \text{ for all } \alpha \in S\}.$$

For any set $S$, $G^S$ is the group of functions from $S$ to $G$, under pointwise multiplication $\#: G^S \times G^S \to G^S$, where $(f \# g)(s) = f(s) \cdot g(s)$.

The map $I: G \to \mathscr{A}(G)$ defined by

$$I(g)(x) = gxg^{-1} \quad (g, x \in G)$$

is a homomorphism of $G$ onto the group $I(G)$ of inner automorphisms of $G$. $I(G) \lhd \mathscr{A}(G)$, since

$$\alpha I(g)\alpha^{-1} = I(\alpha(g)) \quad (\alpha \in \mathscr{A}(G), g \in G).$$

$G$ is *complete* if $I: G \to \mathscr{A}(G)$ is an isomorphism. Thus the automorphism sequence obtained from $G$ stabilizes in finitely many steps if and only if $\mathscr{A}^r(G)$ is complete for some integer $r$. If $\mathscr{C}(G) = 1$, $I: G \to \mathscr{A}(G)$ is a monomorphism. In this situation, we will identify $g \in G$ with $I(g) \in \mathscr{A}(G)$ whenever convenient.

Let $K$ be a *characteristic* subgroup of $G$; that is, restriction to $K$ induces a homomorphism $\mathscr{A}(G) \to \mathscr{A}(K)$. Then the natural projection $G \to G/K$ induces a homomorphism $\mathscr{A}(G) \to \mathscr{A}(G/K)$. Throughout this paper, homomorphisms of the type $G \to \mathscr{A}(G)$, $\mathscr{A}(G) \to \mathscr{A}(K)$, and $\mathscr{A}(G) \to \mathscr{A}(G/K)$ will always be given by $I$ (viewed as an inclusion if $\mathscr{C}(G) = 1$), restriction to the characteristic subgroup $K$, and the map induced by the natural projection, respectively. Note that in general these homomorphisms are neither monomorphisms nor epimorphisms.

For $n \geq 3$, $SL(n, \mathbf{Z})$ is the commutator subgroup of $GL(n, \mathbf{Z})$ (see [5] or [11, p. 108]). Therefore $SL(n, \mathbf{Z})$ is a characteristic subgroup of $GL(n, \mathbf{Z})$, and Wan [17] has established:

THEOREM 1.1 [17]. *For $n \geq 3$, the restriction map induces an epimorphism*

$$\mathscr{A}(GL(n, \mathbf{Z})) \to \mathscr{A}(SL(n, \mathbf{Z})). \quad \blacksquare$$

This result also follows from O'Meara's more general determination of the automorphisms of linear groups over integral domains.

If $K$, $Y$ are groups and $\mu: Y \to \mathscr{A}(K)$ is a homomorphism, the *semidirect product* of $K$ by $Y$ with action $\mu$, written $K \times_\mu Y$, is defined to be the set $K \times Y$ with multiplication

$$(h, x)(k, y) = (h \cdot \mu_x(k), xy) \quad (h, k \in K, x, y \in Y).$$

Thus $K \times_\mu Y$ is a group, and we will identify $K$, $Y$ as subgroups of $K \times_\mu Y$. Note that $G \simeq K \times_\mu Y$ if and only if $G$ contains subgroups $K^*$, $Y^*$ isomorphic to $K$, $Y$ respectively, such that

$$K^* \lhd G, \quad K^* \cap Y^* = 1, \quad G = \langle K^* \cup Y^* \rangle,$$

and such that $\mu$ corresponds to $Y^*$ acting on $K^*$ by conjugation. In the case $Y < \mathscr{A}(K)$ and $\mu$ is inclusion, we write $K \times_* Y$. If Im $\mu = 1$, $K \times_\mu Y \simeq K \oplus Y$.

For groups $K$, $Y$ define $\mathscr{M}(K, Y)$ as a set by $\mathscr{M}(K, Y) = \mathscr{A}(K) \times K^Y \times \mathscr{A}(Y)$; under the product

$$(\alpha_1, \delta_1, \beta_1)(\alpha_2, \delta_2, \beta_2) = (\alpha_1\alpha_2, \alpha_1\delta_2 \ \# \ \delta_1\beta_2, \beta_1\beta_2),$$

$\mathscr{M}(K, Y)$ is a group. Moreover, if $G$ satisfies an exact sequence

$$E\colon 1 \longrightarrow K \xrightarrow{i} G \xrightarrow{j} Y \longrightarrow 1$$

and $i(K)$ is a characteristic subgroup of $G$, then there is a monomorphism $\Phi\colon \mathscr{A}(G) \to \mathscr{M}(K, Y)$. The elements of Im $\Phi$ may be characterized in terms of the data associated with the extension $E$ (cf. [19]). We will exploit this result in the special case $G = K \times_\mu Y$.

Proposition 1.2. *Let $G = K \times_\mu Y$, and assume that $K$ is characteristic in $G$. Define $\Phi\colon \mathscr{A}(G) \to \mathscr{M}(K, Y)$ by*

$$\Phi(\gamma) = (\alpha, \delta, \beta) \quad \text{if } \gamma(k, x) = (\alpha(k)\delta(x), \beta(x)).$$

*Then $(\alpha, \delta, \beta) \in$ Im $\Phi$ if and only if the two conditions below hold, for all $x$, $y \in Y$:*

(1) $\alpha\mu_x\alpha^{-1} = I(\delta(x))\mu_{\beta(x)}$ *(in $\mathscr{A}(K)$), and*
(2) $\delta(xy) = \delta(x) \cdot \mu_{\beta(x)}(\delta(y))$ *(in $K$).*

*Proof.* To each $\gamma \in \mathscr{A}(G)$ there corresponds a commutative diagram

$$\begin{tikzcd}
1 \arrow[r] & K \arrow[r] \arrow[d, "\alpha"] & K \times_\mu Y \arrow[r] \arrow[d] & Y \arrow[r] \arrow[d, "\beta"] & 1 \\
1 \arrow[r] & K \arrow[r] & K \times_\mu Y \arrow[r] & Y \arrow[r] & 1
\end{tikzcd}$$

so that $\gamma(k, x) = (\alpha(k)\delta(x), \beta(x))$, and $(\alpha, \delta, \beta) \in \mathscr{M}(K, Y)$ by the Five-lemma. Conversely, the Five-lemma implies $(\alpha, \delta, \beta) \in$ Im $\Phi$ if and only if $\gamma$, defined as above, respects products in $K \times_\mu Y$. By computation, we require

$$\alpha\mu_x(h) \cdot \delta(xy) = \delta(x) \cdot \mu_{\beta x}(\alpha(h)) \cdot \mu_{\beta x}(\delta(y))$$

for $h \in K$, $x$, $y \in Y$ which is equivalent to (1) and (2). ∎

The propositions which follow are fairly immediate consequences of 1.2 and will be used in the computations of automorphism groups in the rest of this paper. Henceforth, we identify $\mathscr{A}(G)$ and Im $\Phi$.

Proposition 1.3. *Let $G = K \times_\mu Y$, where $K$ is a characteristic subgroup of $G$.*

(1) *The kernel of the projection $\mathscr{A}(G) \to \mathscr{A}(K) \oplus \mathscr{A}(Y)$ is*

$$\operatorname{Der}_\mu(Y, \mathscr{C}(K)) = \{\delta\colon Y \to \mathscr{C}(K) \mid \delta(xy) = \delta(x) \cdot \mu_x\delta(y) \text{ for all } x, y \in Y\}.$$

(2) *If* $\mathscr{C}(K) = 1$, $\mathscr{A}(G)$ *is isomorphic to the subgroup of* $\mathscr{A}(K) \oplus \mathscr{A}(Y)$ *given by*

$$\{(\alpha, \beta) \mid \alpha\mu_x\alpha^{-1}\mu_{\beta x}^{-1} \in I(K)\}.$$

(3) *If* $\mathscr{C}(K) = 1$ *and* Ker $\mu = 1$, *then* $\mathscr{A}(G)$ *is isomorphic to the normalizer of* $\langle I(K) \cup \mathrm{Im}\,(\mu)\rangle$ *in* $\mathscr{A}(K)$.

(4) *If* $K$ *is complete,* $\mathscr{A}(G) \simeq K \oplus \mathscr{A}(Y)$.

(5) *If* $Y$ *is complete,* $\mathscr{A}(G) \simeq M \times_\rho Y$ *where*

$$M = \{(\alpha, \delta) \in \mathscr{A}(K) \times \mathrm{Der}_\mu\,(Y, K) \mid [\alpha, \mu_x] = I(\delta(x)) \text{ for all } x \in Y\}$$

*with product* $(\alpha_1, \delta_1)(\alpha_2, \delta_2) = (\alpha_1\alpha_2, \alpha_1\delta_2 \# \delta_1)$, *and*

$$\rho_y(\alpha, \delta) = (\mu_y\alpha\mu_y^{-1}, \mu_y\delta I(y)^{-1}).$$

*Proof.* By 1.2, if $\alpha = \beta = 1$ then $I(\delta(x)) = 1$ so $\delta(x) \in \mathscr{C}(K)$ and $\delta$ is a $\mu$-derivation. This is part (1); for part (2), $I\colon K \to \mathscr{A}(K)$ is a monomorphism when $\mathscr{C}(K) = 1$. Consequently, equation (2) of 1.2 may be derived from equation (1):

$$\begin{aligned} I(\delta(xy)) &= \alpha\mu_{xy}\alpha^{-1}\mu_{\beta(xy)}^{-1} \\ &= \alpha\mu_x\mu_y\alpha^{-1}\mu_{\beta y}^{-1}\mu_{\beta x}^{-1} \\ &= I(\delta(x))\mu_{\beta x}I(\delta(y))\mu_{\beta x}^{-1} \\ &= I(\delta(x)\mu_{\beta x}(\delta(y))). \end{aligned}$$

Therefore Im $(\mathscr{A}(G) \to \mathscr{A}(K) \oplus \mathscr{A}(Y))$ is as described in part (2) above. Part (3) follows from part (2) since $I(K) \lhd \mathscr{A}(K)$ and $\beta \in \mathscr{A}(Y)$ is determined by $\mu\beta\colon Y \to \mathscr{A}(K)$. Part (4) also follows from part (2), since $K$ is complete means that $\mathscr{C}(K) = 1$ and $I(K) = \mathscr{A}(K)$.

When $Y$ is complete, $Y \simeq \{(\mu_y, 1, I(y)) \mid y \in Y\} < \mathscr{A}(G)$. The remaining statements of part (5) follow from 1.2 and the computation

$$(\mu_y, 1, I(y))(\alpha, \delta, 1)(\mu_y^{-1}, 1, I(y^{-1})) = (\mu_y\alpha\mu_y^{-1}, \mu_y\delta I(y^{-1}), 1). \quad \blacksquare$$

We remark that 1.3(3) is Lemma 1.1 in J. S. Rose [14], and that (4) follows from Baer's observation that any complete normal subgroup $K$ of a group $G$ is a direct summand of $G$ [1].

We turn now to the case in which $K$ is also assumed to be abelian. Write the group operation in $K$ additively, and view $K$ as a left $Y$-module by means of $\mu\colon Y \to \mathscr{A}(K)$. Denote by $\mathscr{A}_\mu(K)$ the group of $Y$-module automorphisms of $K$ (that is, $\mathscr{A}_\mu(K) = \mathscr{C}_{\mathscr{A}(K)}(\mathrm{Im}\,\mu)$), and by $\mathrm{Der}_\mu\,(Y, K)$ the additive group of derivations of $Y$ in $K$:

$$\mathrm{Der}_\mu\,(Y, K) = \{\delta\colon Y \to K \mid \delta(xy) = \delta(x) + \mu_x\delta(y) \text{ for all } x, y \in Y\}.$$

The inner derivation determined by $k \in K$ is the derivation $\delta(x) = k - \mu_x(k)$ and corresponds to $I(k) \in \mathscr{A}(K \times_\mu Y)$.

PROPOSITION 1.4. *Let $G = K \times_\mu Y$, where $K$ is a characteristic abelian subgroup of $G$, and let $\pi\colon \mathscr{A}(G) \to \mathscr{A}(Y)$ be the projection.*

(1) $\operatorname{Ker} \pi \simeq \operatorname{Der}_\mu (Y, K) \times_\sigma \mathscr{A}_\mu(K)$, *where* $\sigma_\alpha(\delta) = \alpha\delta$.
(2) $\operatorname{Im} \pi = \{\beta \in \mathscr{A}(Y) \mid \mu, \mu \circ \beta\colon Y \to \mathscr{A}(K)$ *are equivalent representations*$\}$.
(3) *If $Y$ is complete,* $\mathscr{A}(G) \simeq \operatorname{Ker} \pi \times_\rho Y$ *where* $\rho_y(\delta, \alpha) = (\mu_y \delta I(y^{-1}), \alpha)$.
(4) *If $Y$ is complete and* $\operatorname{Im} \mu$ *is abelian,* $\mathscr{A}(G) \simeq \operatorname{Ker} \pi \times_\rho Y$ *where now* $\rho_y(\delta, \alpha) = (\delta I(y^{-1}), \alpha)$.

*Proof.* Since $I(K) = 1$, (1) and (2) are immediate consequences of 1.2. When $Y$ is complete, (3) follows from 1.3(5). To obtain (4), define $s\colon Y \to \mathscr{M}(K, Y)$ by $s(y) = (1, 0, y)$. Then $\operatorname{Im} s < \mathscr{A}(G)$ by 1.2, $\pi s = 1\colon Y \to Y$; and so

$$\mathscr{A}(G) \simeq \operatorname{Ker} \pi \times_I \operatorname{Im} s \simeq \operatorname{Ker} \pi \times_\rho Y. \quad \blacksquare$$

COROLLARY 1.5. *Continue with the hypotheses of* 1.4.

(1) *If* $\operatorname{Ker} \mu = 1$,

$$\mathscr{A}(G) \simeq \operatorname{Der}_* (\operatorname{Im} \mu, K) \times_\rho \mathscr{N}_{\mathscr{A}(K)}(\operatorname{Im} \mu)$$

*where* $\rho_\alpha(\delta) = \alpha \circ \delta \circ I(\alpha^{-1})$.
(2) *If* $\operatorname{Im} \pi = I(Y)$ *and* $\mathscr{C}(Y) = 1$,

$$\mathscr{A}(G) \simeq \operatorname{Der}_\mu (Y, K) \times_\rho (Y \oplus \mathscr{A}_\mu(K))$$

*where* $\rho_{(y,\alpha)}(\delta) = \mu_y \alpha\delta I(y^{-1})$. $\blacksquare$

See also [14] for results related to 1.4 and 1.5.

COROLLARY 1.6. *Let $G \simeq K \oplus Y$ where $K$ is a characteristic abelian subgroup of $G$.*

(1) $\mathscr{A}(G) \simeq \operatorname{Hom} (Y, K) \times_\rho (\mathscr{A}(K) \oplus \mathscr{A}(Y))$, *where* $\rho_{(\alpha,\beta)}(\delta) = \alpha\delta\beta^{-1}$.
(2) *If $Y$ is complete,* $\mathscr{A}(G) \simeq (\operatorname{Hom} (Y, K) \times_\sigma \mathscr{A}(K)) \oplus Y$, *where* $\sigma_\alpha(\delta) = \alpha\delta$. $\blacksquare$

The criterion which follows is due to Burnside; it may be deduced from Rose's 1.3(3).

THEOREM 1.7 [2, p. 95]. *If $\mathscr{C}(G) = 1$, then $\mathscr{A}(G)$ is complete if and only if $G$ is a characteristic subgroup of $\mathscr{A}(G)$.* $\blacksquare$

Finally, we quote Wielandt's rather sweeping sufficient condition for finite automorphism sequences:

THEOREM 1.8 [18]. *If $G$ is a finite group and $\mathscr{C}(G) = 1$, then the automorphism sequence of $G$ stabilizes in finitely many steps.* $\blacksquare$

## 2. Automorphism sequences, $n \geq 3$

The projective general linear group $PGL(n, \mathbf{Z})$ is the quotient of $GL(n, \mathbf{Z})$ by its center $\{I_n, -I_n\}$, and is a group with trivial center. In this section we first prove that the groups $\mathscr{A}(PGL(n, \mathbf{Z}))$ are complete, by applying Burnside's criterion to a special case of Solazzi's Theorem ([15]; see 2.1 below). Theorems A and C for odd $n$ are then obtained as a corollary. Next, we utilize Hua and Reiner's determination of $\mathscr{A}(GL(n, \mathbf{Z}))$ ([5]; see 2.4 below) to complete the proof of Theorem C and then to compute the automorphism sequences for the general linear groups (even $n \geq 4$). We begin by stating the Solazzi results we require.

THEOREM 2.1 [15]. *For $n \geq 3$,*

$$\mathscr{A}(PGL(n, \mathbf{Z})) \simeq \mathscr{A}(PSL(n, \mathbf{Z})) \simeq PSL(n, \mathbf{Z}) \times_\sigma (\mathbf{Z}_2 \oplus \mathbf{Z}_2),$$

*where*

$$\operatorname{Im} \sigma = \{1, \alpha\colon \pm X \to \pm AXA^{-1}, \beta\colon \pm X \to \pm X^{-t}, \alpha\beta\}$$

*and*

$$A = \operatorname{diag}(-1, 1, \ldots, 1) \in GL(n, \mathbf{Z}). \quad ∎$$

We have written $\pm X \in PSL(n, \mathbf{Z})$ for the image of $X \in SL(n, \mathbf{Z})$ under the natural projection, $X^t$ is the transpose of $X$, and $X^{-t} = (X^{-1})^t$.

COROLLARY 2.2. (1) *For $n \geq 3$, $\mathscr{A}(PGL(n, \mathbf{Z}))$ is complete.*
(2) *If $n$ is odd, $n \geq 3$,*

$$\mathscr{A}(PGL(n, \mathbf{Z})) \simeq \mathscr{A}(SL(n, \mathbf{Z})) \simeq \mathscr{A}(GL(n, \mathbf{Z})).$$

*Proof.* (1) Since $PSL(n, \mathbf{Z})$ is perfect when $n \geq 3$ (cf. [11, p. 108]), $PSL(n, \mathbf{Z})$ is the derived group of $\mathscr{A}(PSL(n, \mathbf{Z}))$. Hence $PSL(n, \mathbf{Z})$ is a centerless group, characteristic in its automorphism group; and so Burnside's criterion (1.7) yields (1).

(2) For odd $n \geq 3$, $GL(n, \mathbf{Z})$ has the direct sum decomposition

$$GL(n, \mathbf{Z}) \simeq \{I_n, -I_n\} \oplus SL(n, \mathbf{Z}).$$

Consequently $PGL(n, \mathbf{Z}) \simeq PSL(n, \mathbf{Z}) \simeq SL(n, \mathbf{Z})$ and (e.g., by 1.6(1))

$$\mathscr{A}(GL(n, \mathbf{Z})) \simeq \mathscr{A}(SL(n, \mathbf{Z})). \quad ∎$$

As stated in the introduction, the groups $\mathscr{A}(PGL(n, \mathbf{Z}))$ ($n \geq 3$) were first determined by Hua and Reiner [6]. Ying [20] subsequently established that all automorphisms of $PSL(n, \mathbf{Z})$ are induced by automorphisms of $PGL(n, \mathbf{Z})$ for even $n \geq 6$ (odd $n$ cause no difficulty for $PGL$ and $PSL$ coincide). The final case $n = 4$ is part of Solazzi's result, and we thank the referee for supplying the reference. This replaced a rather laborious though elementary computation which showed directly that Burnside's criterion applies to $\mathscr{A}(PGL(n, \mathbf{Z}))$ (even

$n \geq 4$) by proving that $PGL(n, \mathbf{Z})$ is not isomorphic to the other two subgroups of index two in $\mathscr{A}(PGL(n, \mathbf{Z}))$.

The next proposition paves the way for us to apply the results of Section 1 in the remainder of this section.

PROPOSITION 2.3. *Let $n \geq 3$, and let $K$ be an arbitrary finite group. If $G = K \times_\mu \mathscr{A}(PGL(n, \mathbf{Z}))$, then $K$ is a characteristic subgroup of $G$.*

*Proof.* Let $\pi$: $G \to \mathscr{A}(PGL(n, \mathbf{Z}))$ denote the natural projection. If $K$ is not characteristic in $G$, then $G$ contains a finite normal subgroup whose image under $\pi$ is a non-trivial finite normal subgroup of $\mathscr{A}(PGL(n, \mathbf{Z}))$. Thus it suffices to show that $\mathscr{A}(PGL(n, \mathbf{Z}))$ contains no element with finitely many conjugates, except 1. Let, then, $\alpha \in \mathscr{A}(PGL(n, \mathbf{Z}))$ and assume that $\alpha$ has finitely many conjugates. Then $\alpha$ centralizes a normal subgroup $L$, say, of finite index in $\mathscr{A}(PGL(n, \mathbf{Z}))$. We will prove that $\alpha$ centralizes $PGL(n, \mathbf{Z})$, which implies that $\alpha = 1$. Let $x \in PGL(n, \mathbf{Z})$, $y \in L$; then $x^{-1}yx \in L$ so

$$x^{-1}yx = \alpha(x^{-1}yx) = \alpha(x)^{-1}y\alpha(x).$$

Thus $[x, \alpha] \in PGL(n, \mathbf{Z})$ and centralizes $L$. But $\pm e_{ij}^d \in L$ for $d = |\mathscr{A}(PGL(n, \mathbf{Z}))/L|$, say, where $e_{ij}$ is the elementary matrix with 1's down the main diagonal and in position $i, j$; zeros elsewhere. But the only element of $PGL(n, \mathbf{Z})$ that commutes with all $\pm e_{ij}^d$ $(d \neq 0)$ is the identity. Hence $[x, \alpha] = 1$, or $\alpha = 1$ as required. ∎

We now establish that the automorphism sequence of $SL(n, \mathbf{Z})$ and $GL(n, \mathbf{Z})$ are finite for $n$ even, $n \geq 4$. We first quote Hua and Reiner's determination of $\mathscr{A}(GL(n, \mathbf{Z}))$ for the case under consideration.

THEOREM 2.4 [5]. *Let $n$ be even, $n \geq 4$. Then*

$$\mathscr{A}(GL(n, \mathbf{Z})) \simeq PGL(n, \mathbf{Z}) \times_\sigma (\mathbf{Z}_2 \oplus \mathbf{Z}_2),$$

*where* $\operatorname{Im} \sigma = \{1, \alpha\colon X \to X^{-t}, \beta\colon X \to (\det X)X, \alpha\beta\}$. ∎

COROLLARY 2.5. *For even $n \geq 4$,*

(1) $\mathscr{A}(GL(n, \mathbf{Z})) \simeq \mathscr{A}(PGL(n, \mathbf{Z})) \oplus \mathbf{Z}_2$, *and*
(2) $\mathscr{A}(SL(n, \mathbf{Z})) \simeq \mathscr{A}(PGL(n, \mathbf{Z}))$ *and is complete.*

*Proof.* By 2.1 and 2.4, $\mathscr{A}(PGL(n, \mathbf{Z}))$ is a subgroup of $\mathscr{A}(GL(n, \mathbf{Z}))$ of index 2, and the action induced by $\beta$ on $PGL(n, \mathbf{Z})$ is trivial. This is (1). Since $\beta$ acts trivially on $SL(n, \mathbf{Z})$ and $\mathscr{A}(GL(n, \mathbf{Z})) \to \mathscr{A}(SL(n, \mathbf{Z}))$ is onto (1.1), (2) follows. ∎

Denote the additive group of $2 \times 2$ matrices over $\mathbf{Z}_2$ by $M(2, 2)$ and set $M = M(2, 2) \times_\lambda SL(2, \mathbf{Z}_2)$ where $\lambda$ denotes the (left) action of $SL(2, \mathbf{Z}_2)$ on $M(2, 2)$ given by matrix multiplication:

$$\lambda_A\colon X \to AX \quad (X \in M(2, 2),\ A \in SL(2, \mathbf{Z}_2)).$$

Note that $SL(2, \mathbf{Z}_2) \simeq S_3$, the symmetric group on 3 symbols (they are nonabelian groups of order 6), and that $S_3$ is complete.

THEOREM 2.6. *Let $n \geq 4$ be even.*

(1) $\mathscr{A}^2(GL(n, \mathbf{Z})) \simeq \mathbf{Z}_2 \oplus \mathbf{Z}_2 \oplus \mathscr{A}(PGL(n, \mathbf{Z}))$.
(2) $\mathscr{A}^3(GL(n, \mathbf{Z})) \simeq M \oplus \mathscr{A}(PGL(n, \mathbf{Z}))$ *and* $\mathscr{C}(M) = 1$.
(3) $\mathscr{A}^{3+k}(GL(n, \mathbf{Z})) \simeq \mathscr{A}^k(M) \oplus \mathscr{A}(PGL(n, \mathbf{Z}))$ *and the automorphism sequence of* $GL(n, \mathbf{Z})$ *stabilizes infinitely many steps.*

*Proof.* Put $Y = \mathscr{A}(PGL(n, \mathbf{Z}))$ and write $\mathscr{A}^r$ for $\mathscr{A}^r(GL(n, \mathbf{Z}))$. It follows from (2.1) that $Y/Y' \simeq \mathbf{Z}_2 \oplus \mathbf{Z}_2$.

Since $\mathscr{A}^1 \simeq \mathbf{Z}_2 \oplus Y$, $\mathbf{Z}_2$ is characteristic in $\mathscr{A}^1$ and 1.6(2) yields

$$\begin{aligned}\mathscr{A}^2 &\simeq \{\mathrm{Hom}\,(Y, \mathbf{Z}_2) \times_\sigma \mathscr{A}(\mathbf{Z}_2)\} \oplus Y \\ &\simeq \mathrm{Hom}\,(Y/Y', \mathbf{Z}_2) \oplus Y \\ &\simeq \mathbf{Z}_2 \oplus \mathbf{Z}_2 \oplus Y.\end{aligned}$$

This is (1), and we may again apply 1.6(2) to obtain

$$\begin{aligned}\mathscr{A}^3 &\simeq \{\mathrm{Hom}\,(Y, \mathbf{Z}_2 \oplus \mathbf{Z}_2) \times_\sigma \mathscr{A}(\mathbf{Z}_2 \oplus \mathbf{Z}_2)\} \oplus Y \\ &\simeq \{\mathrm{Hom}\,(\mathbf{Z}_2 \oplus \mathbf{Z}_2, \mathbf{Z}_2 \oplus \mathbf{Z}_2) \times_\sigma \mathscr{A}(\mathbf{Z}_2 \oplus \mathbf{Z}_2)\} \oplus Y \\ &\simeq M \oplus Y.\end{aligned}$$

Since $\mathscr{C}(S_3) = 1$, $\mathscr{C}(M) = \mathscr{C}_{M(2,2)}(SL(2, \mathbf{Z}_2))$, which is the fixed point set of $SL(2, \mathbf{Z}_2)$ acting on $M(2, 2)$ by left multiplication $\lambda$. Consequently $\mathscr{C}(M) = 1$.

Finally, part (3) follows from Wielandt's Theorem (1.8) once we prove that $\mathscr{A}(K \oplus Y) = \mathscr{A}(K) \oplus Y$ for any finite centerless group $K$. By 2.3, $K$ is characteristic in $K \oplus Y$ so 1.3(2) implies

$$\mathscr{A}(K \oplus Y) \simeq \mathscr{A}(K) \oplus \mathscr{A}(Y) \simeq \mathscr{A}(K) \oplus Y. \quad \blacksquare$$

As a consequence of 2.6(3), $\mathscr{A}^{3+k}(GL(n, \mathbf{Z}))$ is complete if and only if $\mathscr{A}^k(M)$ is complete; we conclude this section by establishing that $\mathscr{A}^2(M)$ is a complete group. One preliminary lemma is required; the group $M$ and these computations appear again in Section 4 in connection with the automorphism sequence of $GL(2, \mathbf{Z})$.

LEMMA 2.7. *View $M(2, 2)$ as a left $SL(2, \mathbf{Z}_2)$-module under left multiplication $\lambda$.*

(1) $\mathrm{Der}_\lambda\,(SL(2, \mathbf{Z}_2), M(2, 2)) \simeq M(2, 2)$, *where* $X \in M(2, 2)$ *corresponds to the (inner) derivation* $A \to (I - A)X$.
(2) $\mathscr{A}_\lambda(M(2, 2)) \simeq SL(2, \mathbf{Z}_2)$, *where* $B \in SL(2, \mathbf{Z}_2)$ *corresponds to the automorphism* $X \to XB^{-1}$.

*Proof.* Let $\delta$ be any derivation, and put

$$C = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \in SL(2, \mathbf{Z}_2).$$

Then $C^2 + C + I = 0$, so that we may define $X \in M(2, 2)$ by

$$\delta(C) = (I - C)X.$$

Note that $\delta(C^{-1}) = (I - C^{-1})X$, and that $ACA^{-1} = C^{\pm 1}$ for all $A \in SL(2, \mathbf{Z}_2)$. Consequently,

$$\delta(C^{\pm 1}) = \delta(ACA^{-1}) = (I - C^{\pm 1})\delta(A) + A\delta(C)$$

which implies that $\delta(A) = (I - A)X$ as required.

Now let $\alpha \in \mathscr{A}_\lambda(M(2, 2))$; we claim $\alpha(X) = XB^{-1}$ for some $B \in SL(2, \mathbf{Z}_2)$. Let $m_{ij} \in M(2, 2)$ denote the matrix with 1 in position $i, j$ and zeros elsewhere. The fixed point set of $\langle e_{12} \rangle$ in $M(2, 2)$ is $\langle m_{11}, m_{12} \rangle$, hence $\alpha$ restricted to $\langle m_{11}, m_{12} \rangle$ is an automorphism. Consequently there is a (unique) $B \in SL(2, \mathbf{Z}_2)$ defined by $\alpha(m_{1j}) = m_{1j}B^{-1}$ for $j = 1, 2$. Then

$$\alpha(m_{2j}) = \alpha(e_{21}m_{1j} + m_{1j}) = e_{21}\alpha(m_{1j}) + \alpha(m_{1j}) = m_{2j}B^{-1}.$$

Thus $\alpha X = XB^{-1}$, since $\langle m_{ij} \mid 1 \le i, j \le 2 \rangle = M(2, 2)$. ∎

**Proposition 2.8.** *Let $M = M(2, 2) \times_\lambda SL(2, \mathbf{Z}_2)$. The automorphism sequence of $M$ is*

$$M \lhd \mathscr{A}(M) \lhd \mathscr{A}^2(M) = \mathscr{A}^3(M) = \cdots$$

*where the factor groups are $SL(2, \mathbf{Z}_2)$, $\mathbf{Z}_2$, 1, 1, ....*

*Proof.* Since $M' = M(2, 2) \times_\lambda \langle C \rangle$, where

$$C = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix},$$

and $M'' = M(2, 2)$, $M(2, 2)$ is characteristic in $M$. Moreover $SL(2, \mathbf{Z}_2)$ is complete, so 1.5(2) yields

$$\mathscr{A}(M) \simeq \text{Der}\,(SL(2, Z_2), M(2, 2)) \times_\rho \{SL(2, \mathbf{Z}_2) \oplus \mathscr{A}_\lambda(M(2, 2))\}$$

where $\rho_{(A,\alpha)}(\delta) = \mu_A \alpha\delta I(A^{-1})$. Using the isomorphisms of 2.7 we may write

$$\mathscr{A}(M) = M(2, 2) \times_\rho \{SL(2, \mathbf{Z}_2) \oplus SL(2, \mathbf{Z}_2)\}$$

where now $\rho_{(A,B)}(X) = AXB^{-1}$. Here,

$$I(M) = M(2, 2) \times_\rho \{SL(2, \mathbf{Z}_2) \oplus 1\}.$$

Since $\mathscr{A}(M)'' = M(2, 2)$, $M(2, 2)$ is characteristic in $\mathscr{A}(M)$ and we may apply 1.4(1), (2) where

$$\pi\colon \mathscr{A}^2(M) \to \mathscr{A}(SL(2, \mathbf{Z}_2) \oplus SL(2, \mathbf{Z}_2)).$$

A computation shows that

$$\mathscr{A}(S_3 \oplus S_3) \simeq (S_3 \oplus S_3) \times_* \langle \tau \mid \tau^2 = 1 \rangle,$$

where $\tau(x, y) = (y, x)$. (See Rose, Lemma 1.4 [14]. For a direct proof, the elements of order 3 in $S_3 \oplus S_3$ with centralizers of order $2 \cdot 3^2$ are $(a, 1)$ or $(1, a)$ where $a^3 = 1$. Use an element of $\langle I(S_3 \oplus S_3), \tau \rangle$ to assume these are fixed. Then so are their centralizers, and now fix an element of order 2 in each centralizer by applying $I(\mathbf{Z}_3 \oplus \mathbf{Z}_3)$.) Recall that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{\mathrm{adj}} = \begin{pmatrix} d & b \\ c & a \end{pmatrix} \quad \text{for} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M(2, 2).$$

Thus $\mathrm{adj} \in \mathscr{A}(M(2, 2))$ is an automorphism of $M(2, 2)$ as an abelian group, and is an anti-automorphism of the matrix ring $M(2, 2)$ such that $X^{\mathrm{adj}}X = (\det X)I_2$. It follows that for $(A, B) \in SL(2, \mathbf{Z}_2) \oplus SL(2, \mathbf{Z}_2)$,

$$\mathrm{adj} \circ \rho_{(A,B)} = \rho_{\tau(A,B)} \circ \mathrm{adj} \in \mathscr{A}(M(2, 2)).$$

By 1.4(2), $\tau \in \operatorname{Im} \pi$. In fact, $\tau = \pi(\tilde{\tau})$ where $\tilde{\tau}(X, (A, B)) = (X^{\mathrm{adj}}, (B, A))$. Therefore $\pi$ is onto, and

$$\operatorname{Im} \pi \simeq \langle \pi(I(\mathscr{A}(M))), \pi(\tilde{\tau}) \rangle \simeq \mathscr{A}(M)/I(M(2, 2)) \times_* \langle \tau \rangle.$$

We prove next that $\operatorname{Ker} \pi = I(M(2, 2))$; it follows that $\mathscr{A}^2(M) \simeq \mathscr{A}(M) \times_* \langle \tilde{\tau} \rangle$. By 1.4(1), we have

$$\operatorname{Ker} \pi \simeq \operatorname{Der}_\rho (SL(2, \mathbf{Z}_2) \oplus SL(2, \mathbf{Z}_2), M(2, 2)) \times \mathscr{A}_\rho(M(2, 2)).$$

Thus $\operatorname{Ker} \pi = I(M(2, 2))$ if and only if $\mathscr{A}_\rho(M(2, 2)) = 1$ and every derivation is inner. The first statement follows from 2.7(2) and the fact that $\mathscr{C}(SL(2, \mathbf{Z}_2)) = 1$, and the second follows from the argument of 2.7(1).

Finally, we prove that $\mathscr{A}^2(M)$ is complete. We will do this by exhibiting a characteristic centerless subgroup $L$ of $\mathscr{A}^2(M)$ such that $L < \mathscr{A}(M)$ and such that the induced homomorphism $\mathscr{A}^2(M) \to \mathscr{A}(L)$ is an isomorphism. The fact that $\mathscr{A}^2(M)$ is complete is then a consequence of Burnside's Theorem (1.7). We claim that $L = \mathscr{A}^2(M)''$ has the required properties; note that $L < \mathscr{A}(M)$ since $\mathscr{A}^2(M)/\mathscr{A}(M)$ is abelian. We have

$$L \simeq M(2, 2) \times_\tau (\mathbf{Z}_3 \oplus \mathbf{Z}_3),$$

where $\tau_{(e,f)} X = C^e X C^f$ with

$$C = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

A computation shows that $\mathscr{C}(L) = 1$. $L$ is clearly characteristic in $\mathscr{A}^2(M)$, and it remains to prove that $\mathscr{A}^2(M) \to \mathscr{A}(L)$ is an isomorphism. Since $\mathscr{A}^2(M)/L$ acts faithfully on $L/L' \simeq \mathbf{Z}_3 \oplus \mathbf{Z}_3$, $\mathscr{A}^2(M) \to \mathscr{A}(L)$ is an injection.

We now compute $\mathscr{A}(L)$. Since $\operatorname{Ker} \tau = 1$ and $M(2, 2)$ is characteristic in $L$ (it is the 2-Sylow subgroup of $L$), by 1.5(1)

$$\mathscr{A}(L) \simeq \mathrm{Der}_*\ (\mathrm{Im}\ \tau,\ M(2,2)) \times_\rho \mathscr{N}_{\mathscr{A}(M(2,2))}(\mathrm{Im}\ \tau).$$

The argument of 2.7(1) shows that every derivation is inner, whence

$$|\mathrm{Der}_*\ (\mathrm{Im}\ \tau,\ M(2,2))| = |M(2,2)| = 2^4.$$

Next, we determine the order of the normalizer of Im $\tau$ in $\mathscr{A}(M(2, 2))$. Since $|\mathrm{Im}\ \tau| = 3^2$ and $\mathscr{A}(M(2,2)) \simeq SL(4, \mathbf{Z}_2)$, Im $\tau$ is a Sylow-3-subgroup of $SL(4, \mathbf{Z}_2)$. It therefore suffices to determine the order of the normalizer of

$$S = \left\{ \begin{pmatrix} C^e & 0 \\ 0 & C^f \end{pmatrix} \,\middle|\, e, f \in \mathbf{Z}_3 \right\}$$

in $SL(4, \mathbf{Z}_2)$. A computation yields

$$\mathscr{N}_{SL(4,\mathbf{Z}_2)}(S) = \left\{ \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} \,\middle|\, X_i \in SL(2, \mathbf{Z}_2) \right\} \cup \left\{ \begin{pmatrix} 0 & Y_1 \\ Y_2 & 0 \end{pmatrix} \,\middle|\, Y_i \in SL(2, \mathbf{Z}_2) \right\}$$

which has order $2^3 \cdot 3^2$. Hence $|\mathscr{A}(L)| = |M(2,2)|2^3 \cdot 3^2 = 2^7 \cdot 3^2$. But $|\mathscr{A}^2(M)| = 2|\mathscr{A}(M)| = 2^7 \cdot 3^2$ and $\mathscr{A}^2(M) \to \mathscr{A}(L)$ is a monomorphism. Thus $\mathscr{A}^2(M) \simeq \mathscr{A}(L)$ as required. ∎

### 3. Some $SL(n, \mathbf{Z}_2)$-modules

The results of this section (and Lemma 2.7) are not new (see [4] or [9, p. 25]). However, we require only some rather specific computations for Section 4, so the proofs below are correspondingly elementary. Write $M(n, m)$, $M(n)$ for the additive groups of $n \times m$, $n \times n$ matrices over $\mathbf{Z}_2$ respectively. These groups are left $SL(n, \mathbf{Z}_2)$-modules, where the action $\lambda$ on $M(n, m)$ is by matrix multiplication, and $\kappa$ on $M(n)$ is by conjugation, thus

$$\lambda_A \colon X \to AX, \quad \kappa_A \colon Y \to AYA^{-1} \quad (A \in SL(n, \mathbf{Z}_2),\ X \in M(n, m),\ Y \in M(n)).$$

Since $\{0_n, I_n\}$ is a $\kappa$-submodule of $M(n)$, the quotient inherits a module structure also denoted by $\kappa$, and the natural projection

$$*\colon M(n) \to PM(n) = M(n)/\{0_n, I_n\}$$

is of course a $\kappa$-homomorphism.

Let $m_{ij} \in M(n, m)$ denote the matrix whose sole nonzero entry is a 1 in position $i, j$. For matrices of appropriate shape,

$$m_{ij} m_{kl} = \begin{cases} 0 & \text{if } j \neq k \\ m_{il} & \text{if } j = k. \end{cases}$$

Consequently, $m_{ij} X m_{kl}$ is zero if and only if $X$ has a zero in position $j, k$; we write $X(j, k)$ for this entry.

We view $SL(n, \mathbf{Z}_2)$ $(n \geq 3)$ as the group presented in Steinberg form on generators $e_{ij}$ $(i \neq j,\ 1 \leq i,\ j \leq n)$, where $e_{ij}$ corresponds to the elementary matrix $I_n + m_{ij} \in SL(n, \mathbf{Z}_2)$, subject to the defining relations

$$\begin{aligned} e_{ij}^2 &= I_n && \text{for } i \neq j, \\ [e_{ij}, e_{jk}] &= e_{ik} && \text{for distinct } i, j, k, \\ [e_{ij}, e_{kl}] &= I_n && \text{if } j \neq k, i \text{ and } l \neq k, i. \end{aligned}$$

(See [16], p. 72, or [10] for a full discussion of the Steinberg groups and further references.)

PROPOSITION 3.1. *If $n \geq 4$, every derivation of $SL(n, \mathbf{Z}_2)$ in $M(n, m)$, $M(n)$, or $PM(n)$ is inner. That is,*

(1) $\mathrm{Der}_\lambda\,(SL(n, \mathbf{Z}_2), M(n, m)) \simeq M(n, m)$,
(2) $\mathrm{Der}_\kappa\,(SL(n, \mathbf{Z}_2), M(n)) \simeq PM(n)$, *and*
(3) $\mathrm{Der}_\kappa\,(SL(n, \mathbf{Z}_2), PM(n)) \simeq PM(n)$.

*Proof.* Let $V$ denote any $SL(n, \mathbf{Z}_2)$-module. Since $e_{i,i+1}$ ($i = 1, \ldots, n$ and the subscripts are taken modulo $n$) generate $SL(n, \mathbf{Z}_2)$, a derivation $\delta \in \mathrm{Der}_*\,(SL(n, \mathbf{Z}_2), V)$ is inner if there exists $v \in V$ such that

(1) $$\delta(e_{i,i+1}) = v - e_{i,i+1} \cdot v = (1 - e_{i,i+1}) \cdot v,$$

and then $v$ is determined modulo the fixed point set of $SL(n, \mathbf{Z}_2)$ in $V$. Moreover, the $\delta(e_{ij}) \in V$ are subject only to the conditions

(2) $$(1 + e_{ij}) \cdot \delta(e_{ij}) = 0 \quad (i \neq j),$$

(3) $$\delta(e_{ik}) + \delta(e_{ij}) + e_{ik} e_{jk} \cdot \delta(e_{ij}) + e_{ij} \cdot \delta(e_{jk}) + e_{ik} \cdot \delta(e_{jk}) = 0 \quad (i, j, k \text{ distinct}),$$

(4) $$(1 + e_{kl}) \cdot \delta(e_{ij}) + (1 + e_{ij}) \cdot \delta(e_{kl}) = 0 \quad (j \neq i, k;\ l \neq i, k).$$

Consider first $\delta \in \mathrm{Der}_\lambda\,(SL(n, \mathbf{Z}_2), M(n, m))$; we seek $X \in M(n, m)$ such that

$$\delta(e_{i,i+1}) = m_{i,i+1} X \quad (i = 1, \ldots, n).$$

The matrix $m_{i,i+1} X$ has all rows zero except perhaps its $i$th row, which is the $(i + 1)$st row of $X$. Hence we can solve for a unique $X$ provided equations (2), (3), and (4) imply $m_{rk}\delta(e_{i,i+1}) = 0$ for some $r$ and all $k \neq i$. From equation (2), $m_{ij}\delta(e_{ij}) = 0$ and from equation (4) applied with $i, j, k, l$ distinct ($n \geq 4$),

$$m_{il}\delta(e_{ij}) = m_{ik} m_{kl} \delta(e_{ij}) = m_{ik} m_{ij} \delta(e_{kl}) = 0.$$

Next, let $\delta \in \mathrm{Der}_\kappa\,(SL(n, \mathbf{Z}_2), M(n))$. Equation (1) now reads

$$\delta(e_{i,i+1}) = m_{i,i+1} X + X m_{i,i+1} + m_{i,i+1} X m_{i,i+1}.$$

The nonzero entries of the matrix on the right occur only in row $i$ or column $i + 1$. They determine the off-diagonal entries in row $i + 1$, column $i$ of $X$ and the sum $X(i, i) + X(i + 1, i + 1)$. Hence solutions of equation (1) determine an element of $PM(n)$ (reflecting the fact that $\{0_n, I_n\}$ is the fixed point set of

$SL(n, \mathbf{Z}_2)$ in $M(n)$); and solutions exist if and only if (2), (3), and (4) imply

$$m_{rl}\delta(e_{i,i+1})m_{ks} = 0 \quad \text{for some } r, s \text{ and all } l \neq i,\ k \neq i+1.$$

Note first that equation (2) implies that $\delta(e_{ij})$ commutes with $m_{ij}$, and equation (4) yields

$$e_{ij}e_{kl}(\delta(e_{ij}) + \delta(e_{kl})) = (\delta(e_{ij}) + \delta(e_{kl}))e_{ij}e_{kl}$$

for $j \neq i, k$ and $l \neq i, k$. Since $m_{kl}e_{ij}e_{kl} = m_{kl}$,

$$m_{kl}(\delta(e_{ij}) + \delta(e_{kl})) = m_{kl}(\delta(e_{ij}) + \delta(e_{kl}))e_{ij}e_{kl}.$$

But $m_{kl}$ commutes with $\delta(e_{kl})$, whence $0 = m_{kl}\delta(e_{ij})m_{kl}$ ($l \neq i,\ k,\ j \neq i,\ k$). It remains to show that $m_{rl}\delta(e_{i,i+1})m_{ls} = 0$ for $l \neq i, i+1$ and some $r, s$. For this, we use equation (3): pick $l \neq i, j, k$ ($n \geq 4$) and multiply (3) on the left and right by $m_{ll}$. This yields $m_{ll}\delta(e_{ij})m_{ll} = 0$ for all distinct $i, j, l$ and completes the proof of part (2).

Finally, let $\delta \in \mathrm{Der}_K\,(SL(n, \mathbf{Z}_2), PM(n))$. Let $D_{ij} \in M(n)$ satisfy

$$D_{ij}^* = \delta(e_{ij}) \in PM(n), \quad D_{ij}(r, r) = 0 \text{ for some } r = r(i, j) \neq i, j.$$

Then (2), (3), (4), in terms of the $D_{ij}$, read

$$D_{ij} + e_{ij}D_{ij}e_{ij} = a(i, j)I_n, \tag{2$'$}$$

$$D_{ik} + D_{ij} + e_{ik}e_{jk}D_{ij}e_{jk}e_{ik} + e_{ij}D_{jk}e_{ij} + e_{ik}D_{jk}e_{ik} = a(i, j, k)I_n, \tag{3$'$}$$

$$D_{ij} + e_{kl}D_{ij}e_{kl} + D_{kl} + e_{ij}D_{kl}e_{ij} = a(i, j, k, l)I_n \tag{4$'$}$$

where $i \neq j$ in (2′); $i, j, k$ are distinct in (3′); $j \neq i, k$; $l \neq i, k$ in (4′); and $a(i, j)$, $a(i, j, k)$, $a(i, j, k, l) \in \mathbf{Z}_2$. We claim that all $a$'s are zero, whence part (3) follows from part (2).

Pick $k \neq i, j$ and multiply (2′) left and right by $m_{kk}$ to obtain $a(i, j) = 0$. Consequently $m_{ij}$ commutes with $D_{ij}$ so

$$m_{ij}D_{ij}m_{ji} = m_{ii}m_{ij}D_{ij}m_{ji}m_{ii} = m_{ii}D_{ij}m_{ii}.$$

Therefore the trace of the matrices appearing in (3′) is

$$n \cdot a(i, j, k) = \mathrm{tr}\,(D_{ik}) = \sum_{s \neq i,k} D_{ik}(s, s).$$

However for any $s \neq i, j, k$, when (3′) is multiplied left and right by $m_{ss}$ we obtain

$$m_{ss}D_{ik}m_{ss} = a(i, j, k)m_{ss}.$$

Consequently $D_{ik}(s, s) = a(i, j, k)$ for $s \neq i, j, k$. If the choice of $D_{ik} \in \delta(e_{ik})$ was such that $r(i, k) \neq i,\ j,\ k$ then $a(i, j, k) = D_{ik}(r(i, k), r(i, k)) = 0$. Otherwise, $r(i, k) = j$ and the trace relation now yields

$$a(i, j, k) = D_{ik}(j, j) = 0.$$

Finally, we establish that all $a(i, j, k, l)$ are zero. For $n \geq 5$, or for $n = 4$ and $i = k$ or $j = l$, there is an $s \neq i, j, k, l$ with $s \in \{1, 2, \ldots, n\}$. Multiply (4′) left and right by $m_{ss}$ to obtain $a(i, j, k, l) = 0$. In the remaining case, $n = 4$ and $i, j, k, l$ are distinct. Multiply (3′) on the left by $m_{kl}$ and on the right by $m_{jk}$ to obtain

$$0 = m_{kl} D_{ik} m_{jk} + m_{kl} D_{ij} m_{ik} = m_{kl} D_{ik} m_{jk} + m_{kl} D_{ij} m_{ij} m_{jk};$$

and so $0 = m_{kl} D_{ik} m_{jk}$ by (2′), valid for all distinct $i, j, k, l$. Now (4′) multiplied left and right by $m_{ii}$ yields $a(i, j, k, l) = 0$. ∎

COROLLARY 3.2. *If at least one of $n$, $m$ is $\neq 3$,*

$$\mathrm{Der}_\rho\,(SL(n, \mathbf{Z}_2) \oplus SL(m, \mathbf{Z}_2), M(n, m)) \simeq M(n, m),$$

*where* $\rho_{(A,B)}(X) = AXB^{-1}$.

*Proof.* Apply transpose if necessary to assume $n \neq 3$. If

$$\delta\colon SL(n, \mathbf{Z}_2) \oplus SL(m, \mathbf{Z}_2) \to M(n, m)$$

is a derivation, by 3.1(1) or the argument of 2.7(1) there is an $X \in M(n, m)$ such that

$$\delta(A, I_m) = (I_n - A)X \quad \text{for all } A \in SL(n, \mathbf{Z}_2).$$

Since $(A, I_m)$ and $(I_n, B)$ commute, $(I_n - A)(\delta(I_n, B) - X(I_m - B^{-1})) = 0$. In particular

$$m_{i,i+1}(\delta(I_n, B) - X(I_m - B^{-1})) = 0 \quad \text{for } i = 1, \ldots, n$$

or

$$\delta(I_n, B) = X(I_m - B^{-1}).$$

Consequently $\delta(A, B) = X - AXB^{-1}$ as required. ∎

PROPOSITION 3.3. (1) $\mathrm{Der}_\lambda\,(SL(3, \mathbf{Z}_2), M(3, m)) \simeq M(4, m)$.
(2) $\mathrm{Der}_\kappa\,(SL(3, \mathbf{Z}_2), M(3)) \simeq \mathrm{Der}_\kappa\,(SL(3, \mathbf{Z}_2), PM(3)) \simeq PM(3)$.

*Proof.* We establish (2) first: as in 3.1(2), (3), $e_{ij} D_{ij} e_{ij} + D_{ij} = I_3$ is impossible, so $D_{ij}$ commutes with $e_{ij}$. Therefore the $(i, i)$ and $(j, j)$ entries of $D_{ij}$ are equal and the other entries in row $j$, column $i$ are zero. For $k \neq i, j$, the $(k, k)$ entry of $D_{ij}$ is zero either as a consequence of equation (4) or by the choice of $D_{ij} \in D_{ij}^*$ made as in 3.1. This yields part (2).

For part (1), equations (2), (3), and (4) reduce to

$$m_{ij}\delta(e_{ij}) = 0 \qquad (i \neq j),$$

$$\delta(e_{ik}) = m_{jk}\delta(e_{ij}) \quad (i, j, k \text{ distinct}).$$

Consequently $m_{1k}\delta(e_{ij}) = m_{11}\delta(e_{23})$, and there is a unique $X \in M(3, m)$ such that

$$\delta(e_{ij}) = (1 - e_{ij})X + m_{k1}\delta(e_{23}) \quad (k \neq i, j).$$

Thus $\delta$ corresponds to the pair $(X, m_{11}\delta(e_{23}))$. Since $m_{11}\delta(e_{23})$ can be any matrix in $M(3, m)$ with rows 2 and 3 zero, we obtain part (1). ∎

PROPOSITION 3.4. *Let $n \geq 3$.*

(1) $\mathscr{A}_\rho(M(n, m) \oplus M(n)) \simeq SL(m, \mathbf{Z}_2) \oplus \mathbf{Z}_2$, *where the isomorphism is given by*

$$(B, b) \to \beta\colon (X, Y) \to (XB^{-1}, Y + b \operatorname{Tr}(Y) I_n)$$

*and $SL(n, \mathbf{Z}_2)$ acts on $M(n, m) \oplus M(n)$ by $\rho_A(X, Y) = (AX, AYA^{-1})$.*

(2) $\mathscr{A}_\rho(M(n, m) \oplus PM(n)) \simeq SL(m, \mathbf{Z}_2)$, *where the action of $SL(n, \mathbf{Z}_2)$ is as in* (1).

(3) $\mathscr{A}_\sigma(M(n, m) \oplus PM(n)) = 1$, *where $SL(n, \mathbf{Z}_2) \oplus SL(m, \mathbf{Z}_2)$ acts by*

$$\sigma_{(A,B)}(X, Y) = (AXB^{-1}, AYA^{-1}).$$

(4) $\mathscr{A}_\lambda(M(n, m)) \simeq SL(m, \mathbf{Z}_2)$.

*Proof.* Note that (3), (4) follow from (1), (2).

Let $\beta \in \mathscr{A}_\rho(M(n, m) \oplus M(n))$, where $\rho_A(X, Y) = (AX, AYA^{-1})$ for $A \in SL(n, \mathbf{Z}_2)$. Let $S_1$, $S_2$ be the subgroups of $SL(n, \mathbf{Z}_2)$ defined by

$$S_1 = \left\{ \begin{pmatrix} 1 & * \cdots * \\ 0 & \\ \vdots & A \\ 0 & \end{pmatrix} \middle| A \in SL(n-1, \mathbf{Z}_2) \right\},$$

$$S_2 = \left\{ \begin{pmatrix} 1 & 0 \cdots 0 \\ 0 & \\ \vdots & A \\ 0 & \end{pmatrix} \middle| A \in SL(n-1, \mathbf{Z}_2) \right\}.$$

The fixed point sets of $\rho(S_1)$, $\rho(S_2)$ in $M(n, m) \oplus M(n)$ are

$$T_1 = \left\{ \left( \sum_j a_j m_{1j}, aI_n \right) \right\}, \quad T_2 = \left\{ \left( \sum_j a_j m_{1j}, bI_n + m_{11} \right) \right\},$$

respectively, and Im $\rho$ fixes $\langle (0, I_n) \rangle$. Since the restrictions of $\beta$ to these fixed point sets are isomorphisms, we have first a (unique) $B \in SL(m, \mathbf{Z}_2)$ such that

$$(1) \qquad \beta(m_{1j}, 0) = (m_{1j}B^{-1}, a(j)I_n)$$

and second that

$$(2) \qquad \beta(0, m_{11}) = \left( \sum_j b_j m_{1j}, m_{11} + bI_n \right).$$

Follow $\beta$ by $(X, Y) \to (XB, Y)$ so as to assume $B = I_m$ in equation (1). Then $\beta(e_{k1} m_{1j}, 0) = \rho(e_{k1})\beta(m_{1j}, 0)$ implies

$$\beta(m_{kj}, 0) = (m_{kj}, 0), \quad k \neq 1$$

and the action of $e_{1k}$ applied to the equation above yields $\beta(X, 0) = (X, 0)$. Next, apply $e_{k1}$ to equation (2) to obtain

$$\beta(0, m_{k1}) = \left(\sum_j b_j m_{kj}, m_{k1}\right), \quad k \neq 1,$$

and then apply $e_{1k}$ to (2) to obtain $\beta(0, m_{1k}) = (0, m_{1k})$, $k \neq 1$. But $m_{k1} = m_{1k}^t$ so these matrices are conjugate, which implies $b_j = 0$. Now follow $\beta$ by

$$(X, Y) \to (X, Y + b \operatorname{Tr}(Y) I_n)$$

so as to obtain $b = 0$ in equation (2). Since $m_{11}$ and $m_{ii}$ are conjugate, $\beta(0, m_{11}) = (0, m_{11})$ implies $\beta(0, m_{ii}) = (0, m_{ii})$.

The argument that establishes part (2) is similar; the fixed point set of $S_1, S_2$ are the images of $T_1$, $T_2$ in $M(n, m) \oplus PM(n)$. ▮

## 4. The case $n = 2$

In this final section, we derive Theorem B, Proposition 2.3, and then Theorem A for the groups $PGL(2, \mathbf{Z})$, $GL(2, \mathbf{Z})$, and $SL(2, \mathbf{Z})$. It will be convenient to relate these groups to the free product structure of $PSL(2, \mathbf{Z})$. We will view $SL(2, \mathbf{Z})$ as a split extension of the free product $L = \mathbf{Z}_3 * \mathbf{Z}_3$ by $\mathbf{Z}_4$, and $GL(2, \mathbf{Z})$ as a split extension of $L$ by the dihedral group of order 8. Since the index of $L$ is in each case a power of 2 whereas $L$ is generated by elements of order 3, it follows that $L$ is characteristic. We first summarize the properties of free products (with amalgamation) that will be used below; proofs may be found in Magnus, Karrass and Solitar [8, Chapter 4, Section 2], for example. We write $H *_A K$ for the free product of $H$ and $K$ with amalgamated subgroup $A$; we assume given and fixed monomorphisms $A \to H$, $A \to K$ and identify $A$ as the intersection $H \cap K$ in $H *_A K$. The groups $H$ and $K$ are termed the *factors* of the free product $H *_A K$. A *right transversal* of $A$ in $H$ denotes a subset $\mathscr{H} \subset H$ such that $1 \in \mathscr{H}$ and $H$ is the disjoint union $\bigcup_{h \in \mathscr{H}} Ah$.

THEOREM 4.1. (1) *Let $\mathscr{H}$, $\mathscr{K}$ be right transversals of $A$ in $H$, $K$ respectively. Then each element of $H *_A K$ has a unique expression in the form $as_1 s_2 \cdots s_m$ where $a \in A$, $s_i \in \mathscr{H} \cup \mathscr{K}$, and $s_i$, $s_{i+1}$ lie in different factors.*

(2) *Any element of finite order in $H *_A K$ is conjugate to an element of finite order in one of the factors.*

(3) $\mathscr{C}(H *_A K) = A \cap \mathscr{C}(H) \cap \mathscr{C}(K)$. ▮

Let $D$ be the dihedral group of order 8; $D \simeq \mathbf{Z}_4 \times_* \mathscr{A}(\mathbf{Z}_4)$. Fix the presentation

$$D = \langle a, b \mid a^4 = b^2 = (ba)^2 = 1 \rangle.$$

Note that every nontrivial normal subgroup of $D$ contains $\mathscr{C}(D) = \langle a^2 \rangle$, and that the only elements of order 4 in $D$ are $a^{\pm 1}$. Consequently $\mathbf{Z}_4$ is characteristic in $\mathbf{Z}_4 \times_* \mathscr{A}(Z_4) \simeq \mathbf{Z}_4 \times_* \mathbf{Z}_2$, so 1.4(1) and (2) yield

$$\mathscr{A}(D) \simeq \mathbf{Z}_4 \times_* \mathscr{A}(\mathbf{Z}_4) = \langle \sigma, \tau \mid \sigma^4 = \tau^2 = (\sigma\tau)^2 = 1 \rangle$$

where $\sigma: b \to ab, a \to a$ and $\tau: b \to b, a \to a^{-1}$. Furthermore $I(a) = \sigma^2, I(b) = \tau$.

Let $L = \mathbf{Z}_3 * \mathbf{Z}_3$ have the presentation $L = \langle c_1, c_2 \mid c_1^3 = c_2^3 = 1 \rangle$.

PROPOSITION 4.2. (1) *$\mathscr{A}(L) \simeq L \times_\rho D$, where $\rho: D \to \mathscr{A}(L)$ is defined by*

$$\rho_a: c_1 \to c_2^{-1}, c_2 \to c_1; \qquad \rho_b: c_1 \to c_2, c_2 \to c_1.$$

(2) *$\mathscr{A}(L)$ is complete.*

(3) *If $G = K \times_\mu \mathscr{A}(L)$ and $K$ is finite, then $K$ is characteristic in $G$.*

*Proof.* (1) By direct computation, $\rho_{a^2} \neq 1$ so $\rho: D \to \mathscr{A}(L)$ is an injective homomorphism. Moreover $\rho(D) \cap L = 1$ (where we have identified $L$ with $I(L)$), since $L$ has no 2-torsion (4.1(2), (3)). We claim $\mathscr{A}(L) = \langle L, \rho(D) \rangle$, and this is part (1). Let $\gamma \in \mathscr{A}(L)$; $\gamma$ is determined by $\gamma(c_1)$ and $\gamma(c_2)$. By 4.1(2), the conjugacy classes of elements of order 3 in $\mathscr{A}(L)$ are represented by $\{c_1, c_1^{-1}, c_2, c_2^{-1}\}$. Since $\rho_a$ acts transitively on this set, we may follow $\gamma$ by a suitable element of $\langle L, \rho_a \rangle$ to obtain $\gamma(c_2) = c_2$. Now $\gamma(c_1)$ is $L$-conjugate to $c_1$ or to $c_1^{-1}$. Since $\gamma(c_2) = c_2$ and $\gamma(c_1)$ generate $L$, 4.1(1) implies that $\gamma(c_1) = c_2^{-e} c_1^{\pm 1} c_2^e$ whence $\gamma \in \langle I(c_2), \rho_{ba} \rangle$.

(2) We have $\mathscr{C}(L) = 1$ (4.1(3)), and $L$ is generated by all elements of order 3 in $\mathscr{A}(L)$. Hence $L$ is characteristic in $\mathscr{A}(L)$; and so $\mathscr{A}(L)$ is complete by Burnside's criterion.

(3) As in the proof of 2.3, we must show that $\mathscr{A}(L)$ has no nontrivial finite normal subgroups, which is equivalent to showing that the centralizer of any normal subgroup of finite index in $\mathscr{A}(L)$ is trivial. By 4.1(2), $L$ has this property. Now suppose $H \lhd \mathscr{A}(L)$ and $|\mathscr{A}(L)/H|$ is finite. If $\alpha \in \mathscr{C}_{\mathscr{A}(L)}(H)$, then for all $x \in L$, $y \in H \cap L$,

$$x^{-1}yx = \alpha(x^{-1}yx) = \alpha(x)^{-1}y\alpha(x).$$

Consequently $\alpha(x)x^{-1} \in \mathscr{C}_L(H \cap L) = 1$, or $\alpha = 1$. ∎

Fix the presentation $\mathbf{Z}_4 = \langle a \mid a^4 = 1 \rangle$ together with the monomorphism $\mathbf{Z}_4 \to D$ defined by $a \to a$.

PROPOSITION 4.3. *There is a commutative diagram*

$$\begin{tikzcd}
1 \arrow[r] & L \arrow[r] \arrow[d, "1"] & SL(2, \mathbf{Z}) \arrow[r] \arrow[d] & \mathbf{Z}_4 \arrow[r] \arrow[d] & 1 \\
1 \arrow[r] & L \arrow[r] & GL(2, \mathbf{Z}) \arrow[r] & D \arrow[r] & 1
\end{tikzcd}$$

*where the middle vertical arrow is inclusion, and both rows are split exact (that is, the middle group is a semidirect product of the end groups). The action $\mu: D \to \mathscr{A}(L)$ is given by $\mu_a = \rho_b$, $\mu_b = \rho_{a^2}$. Moreover, $L$ is a characteristic subgroup of $SL(2, \mathbf{Z})$ and of $GL(2, \mathbf{Z})$, and $SL(2, \mathbf{Z})$ is a characteristic subgroup of $GL(2, \mathbf{Z})$.*

*Proof.* The group $SL(2, \mathbf{Z})$ has the well-known description as $\mathbf{Z}_4 *_{\mathbf{Z}_2} \mathbf{Z}_6$ (see [11, p. 139] or [8, p. 47]); thus

$$SL(2, \mathbf{Z}) \simeq \langle x, y \mid x^2 = y^3, x^4 = 1 \rangle,$$

where

$$x \to \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad y \to \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}$$

defines the requisite isomorphism. By 4.1(1), each element of $SL(2, \mathbf{Z})$ may be written uniquely in the form

$$as_1 s_2 \cdots s_m$$

where $a \in \{1, x^2\}$, $s_i \in \{x\} \cup \{y, y^{-1}\}$, and $s_i$, $s_{i+1}$ lie in different factors $\langle x \rangle$, $\langle y \rangle$. Define the homomorphism $L \to SL(2, \mathbf{Z})$ by

$$c_1 \to y^2 = x^2 y^{-1}, \qquad c_2 \to xy^2x^{-1} = xy^{-1}x.$$

The uniqueness of the normal forms 4.1(1) implies that $L \to SL(2, \mathbf{Z})$ is a monomorphism. Since $I(y) = I(y^{-2})$ and $I(x^2) = 1$ in $\mathscr{A}(SL(2, \mathbf{Z}))$, a computation shows that $L \lhd SL(2, \mathbf{Z})$. By 4.1(2), $L$ is the group generated by all elements of order 3 in $SL(2, \mathbf{Z})$; and so is a characteristic subgroup. The map $SL(2, \mathbf{Z}) \to \mathbf{Z}_4$ defined by $x \to a$, $y \to a^2$ is an epimorphism whose kernel contains $L$ and whose restriction to $\langle x \rangle$ is an isomorphism. Since $SL(2, \mathbf{Z}) = \langle L, x \rangle$ and $L \cap \langle x \rangle = 1$, the first row is split exact.

To obtain the exactness of the second row, note that the sequence

$$1 \to SL(2, \mathbf{Z}) \to GL(2, \mathbf{Z}) \to \mathbf{Z}_2 = \langle b \mid b^2 = 1 \rangle \to 1$$

is split exact, where we select the splitting map $\mathbf{Z}_2 \to GL(2, \mathbf{Z})$ to be

$$b \to \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then the matrices

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

generate a subgroup of $GL(2, \mathbf{Z})$ isomorphic to $D$, and the argument now proceeds as above.

Finally, since $L$ is characteristic in $GL(2, \mathbf{Z})$ and $D$ contains a unique subgroup isomorphic to $\mathbf{Z}_4$, it follows that $SL(2, \mathbf{Z})$ is characteristic in $GL(2, \mathbf{Z})$. ∎

COROLLARY 4.4. (1) *$PSL(2, \mathbf{Z}) \simeq L \times_\rho \langle b \rangle$, $PGL(2, \mathbf{Z}) \simeq L \times_\rho \langle a^2, b \rangle$.*

(2) *$\mathscr{A}(PSL(2, \mathbf{Z})) \simeq PGL(2, \mathbf{Z})$.*

(3) *$\mathscr{A}(PGL(2, \mathbf{Z})) \simeq L \times_\rho D \simeq \mathscr{A}(L)$ and is complete.*

*Proof.* By 4.3, $PSL(2, \mathbf{Z}) \simeq L \times_{\mu} (\mathbf{Z}_4/\text{Ker } \mu)$ and $PGL(2, \mathbf{Z}) \simeq L \times_{\mu} (D/\text{Ker } \mu)$. These correspond to the descriptions given in (1).

$L$ is characteristic in both $PSL(2, \mathbf{Z})$ and $PGL(2, \mathbf{Z})$, $\mathscr{C}(L) = 1$, and $\text{Ker } \rho = 1$. Therefore 1.3(3) implies that

$$\mathscr{A}(PSL(2, \mathbf{Z}))/I(PSL(2, \mathbf{Z})) \simeq \mathscr{N}_D(\langle b \rangle) = \langle a^2, b \rangle,$$

and $\mathscr{A}(PGL(2, \mathbf{Z})) = \mathscr{A}(L)$. ∎

This completes the proof of Theorem B and establishes 2.3 for the case $n = 2$.

THEOREM 4.5. (1) $\mathscr{A}(SL(2, \mathbf{Z})) \simeq \mathbf{Z}_2 \oplus PGL(2, \mathbf{Z})$.
(2) $\mathscr{A}(GL(2, \mathbf{Z})) \simeq \mathbf{Z}_2 \oplus \mathbf{Z}_2 \oplus PGL(2, \mathbf{Z})$.

*Proof.* We have $SL(2, \mathbf{Z}) \simeq L \times_{\mu} \mathbf{Z}_4$; and so 1.3(2) implies that $\mathscr{A}(SL(2, \mathbf{Z}))$ is isomorphic to the subgroup of $\mathscr{A}(L) \oplus \mathscr{A}(Z_4)$ given by

$$\{(\alpha, \beta) \mid \alpha \mu_a \alpha^{-1} \mu_{\beta a}^{-1} \in I(L)\}.$$

Since $\beta(a) = a^{\pm 1}$ and $a^2 \in \text{Ker } \mu$, $\mu_{\beta a} = \mu_a$. Hence

$$\begin{aligned}\mathscr{A}(SL(2, \mathbf{Z})) &\simeq \{(\alpha, \beta) \mid [\alpha, \rho_b] \in I(L)\} \simeq L \times_{\rho} \mathscr{C}_D(b) \oplus \mathscr{A}(\mathbf{Z}_4) \\ &\simeq L \times_{\rho} \langle a^2, b \rangle \oplus \mathbf{Z}_2 \simeq PGL(2, \mathbf{Z}) \oplus \mathbf{Z}_2.\end{aligned}$$

Again by 4.3, we have $GL(2, \mathbf{Z}) \simeq L \times_{\mu} D$, and now 1.3(2) yields

$$\begin{aligned}&\mathscr{A}(GL(2, \mathbf{Z})) \\ &\quad \simeq \{(\alpha, \gamma) \in \mathscr{A}(L) \oplus \mathscr{A}(D) \mid \alpha \mu_a \alpha^{-1} \mu_{\gamma a}^{-1} \in I(L) \text{ and } \alpha \mu_b \alpha^{-1} \mu_{\gamma b}^{-1} \in I(L)\}.\end{aligned}$$

As above, $\gamma a = a^{\pm 1}$ so the first condition on $(\alpha, \gamma)$ reads $\alpha \in PGL(2, \mathbf{Z})$. Since $\mu_b = \rho_{a^2}$ and $a^2 \in \mathscr{C}(D)$, the second condition is that $b^{-1}\gamma(b) \in \text{Ker } \mu$. Thus

$$\begin{aligned}\mathscr{A}(GL(2, \mathbf{Z})) &\simeq PGL(2, \mathbf{Z}) \oplus \{\gamma \in \mathscr{A}(D) \mid b^{-1}\gamma(b) \in \langle a^2 \rangle\} \\ &= PGL(2, \mathbf{Z}) \oplus I(D) \simeq PGL(2, \mathbf{Z}) \oplus \mathbf{Z}_2 \oplus \mathbf{Z}_2.\end{aligned}$$ ∎

Recall that $M = M(2, 2) \times_{\lambda} SL(2, \mathbf{Z}_2)$, where $M(2, 2)$ is the additive group of $2 \times 2$ matrices over $\mathbf{Z}_2$ viewed as a left $SL(2, \mathbf{Z}_2)$-module by $\lambda_A \colon X \to AX$.

THEOREM 4.6. (1) $\mathscr{A}^2(GL(2, \mathbf{Z})) \simeq M \times_{\sigma} \mathscr{A}(PGL(2, \mathbf{Z}))$.
(2) $\mathscr{A}^3(GL(2, \mathbf{Z})) \simeq (M \times_{\tau} \mathbf{Z}_2) \oplus \mathscr{A}(PGL(2, \mathbf{Z}))$.
(3) $\mathscr{A}^4(GL(2, \mathbf{Z})) \simeq ((M \times_{\tau} \mathbf{Z}_2) \times_{\phi} \mathbf{Z}_2) \oplus \mathscr{A}(PGL(2, \mathbf{Z}))$ *and is a complete group.*

*Proof.* Put $Y = L \times_{\rho} D \simeq \mathscr{A}(L)$, and recall that $PGL(2, \mathbf{Z}) \simeq L \times_{\rho} \langle a^2, b \rangle$ and $\mathscr{A}(PGL(2, \mathbf{Z})) \simeq Y$.

We have $\mathscr{A}(GL(2, \mathbf{Z})) \simeq (\mathbf{Z}_2 \oplus \mathbf{Z}_2) \oplus L \times_{\rho} \langle a^2, b \rangle$. By 1.6(1),

$$\mathscr{A}^2(GL(2, \mathbf{Z})) \simeq \text{Hom}\,(L \times_{\rho} \langle a^2, b \rangle, \mathbf{Z}_2 \oplus \mathbf{Z}_2) \times_{\sigma} (\mathscr{A}(\mathbf{Z}_2 \oplus \mathbf{Z}_2) \oplus Y)$$

where $\sigma_{(\alpha,y)}(\delta) = \alpha \cdot \delta \cdot I(y^{-1})$. Since $(L \times_\rho \langle a^2, b\rangle)' = L$,

$$\text{Hom}\,(L \times_\rho \langle a^2, b\rangle, \mathbf{Z}_2 \oplus \mathbf{Z}_2) \simeq \text{Hom}\,(\langle a^2, b\rangle, M(2, 1)) \simeq M(2, 2)$$

where the second isomorphism assigns the matrix whose first column is obtained by evaluation at $a^2$ and second column by evaluation at $b$. Then

$$\mathscr{A}(\mathbf{Z}_2 \oplus \mathbf{Z}_2) \simeq SL(2, \mathbf{Z}_2)$$

and we choose this isomorphism so that $SL(2, \mathbf{Z}_2)$ acts on $M(2, 2)$ by $X \to AX$. For $y \in Y$, the action is that induced by $I(y^{-1})$ on $\langle a^2, b\rangle$. Consequently, $L \times_\rho \langle a^2, b\rangle$ acts trivially and $I(a^{-1})$ induces the map $a^2 \to a^2$, $b \to a^2 b$ which corresponds to right matrix multiplication by $e_{12} \in SL(2, \mathbf{Z}_2)$. Therefore

$$\mathscr{A}^2(GL(2, \mathbf{Z})) \simeq M(2, 2) \times_\sigma \{SL(2, \mathbf{Z}_2) \oplus Y\} \simeq M \times_\sigma Y \simeq M \times_\sigma (L \times_\rho D)$$

where $L \times_\rho \langle a^2, b\rangle = \text{Ker}\, \sigma$ and $\sigma_a(X, A) = (Xe_{12}, A)$.

Since $M$ is finite, it is characteristic in $\mathscr{A}^2(GL(2, \mathbf{Z}))$ by 4.2(3). Moreover $\mathscr{C}(M) = 1$ so 1.3(2) and 4.2(2) yield

$$\mathscr{A}^3(GL(2, \mathbf{Z})) \simeq \{\alpha, y) \in \mathscr{A}(M) \oplus Y \mid \alpha\sigma_x\alpha^{-1}\sigma^{-1}_{yxy^{-1}} \in I(M) \text{ for all } x \in Y\}.$$

Since $\text{Im}\, \sigma \simeq \mathbf{Z}_2$, $\sigma_{yxy^{-1}} = \sigma_x$; and so

$$\mathscr{A}^3(GL(2, \mathbf{Z})) \simeq \{\alpha \in \mathscr{A}(M) \mid [\alpha, \sigma_a] \in I(M)\} \oplus Y.$$

By Proposition 2.8, $\mathscr{A}(M)/I(M) \simeq SL(2, \mathbf{Z}_2)$, and the centralizer of $\sigma_a \cdot I(M)$ in $\mathscr{A}(M)/I(M)$ is $\langle \sigma_a \rangle \cdot I(M)$. Thus we have part (2), where

$$M \times_\tau Z_2 \simeq M(2, 2) \times_\mu \{SL(2, \mathbf{Z}_2) \oplus \langle e_{12}\rangle\}$$

and $\mu_{(A,B)}(X) = AXB^{-1}$. This group is characteristic in $\mathscr{A}^3(GL(2, \mathbf{Z}))$. Iterating 1.3(2) and 4.2(3) yields

$$\mathscr{A}^{3+k}(GL(2, \mathbf{Z})) \simeq \mathscr{A}^k(M \times_\tau \mathbf{Z}_2) \oplus Y.$$

It remains to show that $\mathscr{A}(M \times_\tau \mathbf{Z}_2) \simeq (M \times_\tau \mathbf{Z}_2) \times_\rho \mathbf{Z}_2$ and that this group is complete.

Since $M(2, 2) = (M \times_\tau \mathbf{Z}_2)''$ we may apply 1.5(1) to obtain

$$\mathscr{A}(M \times_\tau \mathbf{Z}_2) \simeq \text{Der}\,(\text{Im}\, \mu, M(2, 2)) \times_\rho \mathscr{N}_{\mathscr{A}(M(2,2))}(\text{Im}\, \mu).$$

The argument of 2.7(1) shows that every derivation is inner; and so

$$\mathscr{A}(M \times_\tau \mathbf{Z}_2)/I(M(2, 2)) \simeq \mathscr{N}_{\mathscr{A}(M(2,2))}(\text{Im}\, \mu).$$

Note that $\text{Im}\, \mu$ is a dihedral group of order 12, for $\text{Im}\, \mu \simeq SL(2, \mathbf{Z}_2) \oplus \langle e_{12}\rangle$ in which

$$z = \left(\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, e_{12}\right)$$

has order 6 and is inverted by $w = (e_{12}, 1)$ which has order 2. Take the basis $\{m_{11}, m_{21}, m_{12}, m_{22}\}$ for $M(2, 2)$ over $\mathbf{Z}_2$, and identity $\mathscr{A}(M(2, 2))$ with

$SL(4, \mathbf{Z}_2)$ acting from the left on $M(4, 1) \simeq M(2, 2)$. A computation shows that the matrices corresponding to $z$, $w$ are

$$Z = \begin{pmatrix} C & C \\ 0 & C \end{pmatrix} \quad \text{and} \quad W = \begin{pmatrix} e_{21} & 0 \\ 0 & e_{21} \end{pmatrix} \quad \text{where } C = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

We assert that

$$\mathcal{N}_{SL(4,\mathbf{Z}_2)}(\langle Z, W \rangle) = \langle Z, W \rangle \cdot \langle T \rangle \quad \text{where } T = \begin{pmatrix} I & C \\ 0 & I \end{pmatrix}.$$

Since the center of $\langle Z, W \rangle$ is generated by

$$Z^3 = \begin{pmatrix} I & I \\ 0 & I \end{pmatrix}$$

and has order 2, any matrix $X$ which normalizes $\langle Z, W \rangle$ commutes with $\mathbf{Z}^3$ and therefore has block upper triangular form. Since $\langle C, e_{21} \rangle = SL(2, \mathbf{Z}_2)$ we may multiply $X$ from the left by an element of $\langle Z, W \rangle$ to obtain

$$X = \begin{pmatrix} I & Y \\ 0 & I \end{pmatrix} \quad \text{where } Y \in M(2, 2). \tag{1}$$

Note that the matrices in $\langle Z, W \rangle$ of this form are $Z^3$ and $I_4$, and that $X^2 = I_4$. Since $\langle Z^2 \rangle = \langle Z, W \rangle'$, $XZ^2X = Z^{\pm 2}$, which implies that $YC = C^{\pm 1}Y$. Moreover, $[X, Z] \in \langle Z, W \rangle$ implies that

$$CYC^{-1} + Y = 0_2 \text{ or } I_2.$$

Since $CYC^{-1} + Y = (C^{1 \pm 1} + I)Y$, it follows that $CYC^{-1} + Y = 0$ so $Y = aI + bC$. Therefore $X$ or $Z^3X$ is the matrix $T$. But $TZT^{-1} = Z$, $TWT^{-1} = Z^3W$, which shows that $\langle Z, W \rangle \cdot \langle T \rangle$ is the normalizer of $\langle Z, W \rangle$. Therefore

$$\begin{aligned} \mathcal{A}(M \times_\tau \mathbf{Z}_2) &\simeq M(4, 1) \times_\lambda \langle Z, W, T \rangle \\ &\simeq (M(2, 2) \times_\mu \langle z, w \rangle) \times_\phi \mathbf{Z}_2 \\ &\simeq (M \times_\tau \mathbf{Z}_2) \times_\phi \mathbf{Z}_2. \end{aligned}$$

A computation establishes that $M(2, 2) \simeq \mathcal{A}(M \times_\tau \mathbf{Z}_2)''$, so 1.5(1) applied as above yields

$$\mathcal{A}^2(M \times_\tau \mathbf{Z}_2) \simeq M(4, 1) \times_\lambda \mathcal{N}_{SL(4,\mathbf{Z}_2)}(\langle Z, W, T \rangle).$$

We obtain part (3) once we show that $\langle Z, W, T \rangle$ is a self-normalizing subgroup of $SL(4, \mathbf{Z}_2)$. To this end, since $\langle Z, W, T \rangle' = \langle Z \rangle$, any matrix $X$ which normalizes $\langle Z, W, T \rangle$ may be selected modulo $\langle Z, W \rangle$ to be block upper triangular of form (1) and to centralize $Z$. But then the upper right hand $2 \times 2$ corner of $X$ commutes with $C$. Thus $X = Z^{3a}T^b$, as required. ∎

We now turn to the determination of the automorphism sequence of $SL(2, \mathbf{Z})$. We shall first show that $\mathscr{A}(PGL(2, \mathbf{Z}))$ is a direct summand of $\mathscr{A}^4(SL(2, \mathbf{Z}))$.

PROPOSITION 4.7. (1) $\mathscr{A}^2(SL(2, \mathbf{Z})) \simeq (\mathbf{Z}_2 \oplus \mathbf{Z}_2) \times_\sigma \mathscr{A}(PGL(2, \mathbf{Z}))$.
(2) $\mathscr{A}^3(SL(2, \mathbf{Z})) \simeq (\mathbf{Z}_2 \oplus D) \times_\mu \mathscr{A}(PGL(2, \mathbf{Z}))$.
(3) $\mathscr{A}^4(SL(2, \mathbf{Z})) \simeq (M(8, 1) \times_\lambda \langle Z \rangle) \oplus \mathscr{A}(PGL(2, \mathbf{Z}))$, *where*

$$Z \in SL(8, \mathbf{Z}_2)$$

*is the block lower triangular matrix*

$$\begin{pmatrix} I_4 & 0 \\ I_4 & I_4 \end{pmatrix}.$$

*Moreover, the centers of these three groups are elementary abelian groups whose orders are* 2, $2^2$, *and* $2^4$, *respectively.*

*Proof.* In 4.5(1), we obtained $\mathscr{A}(SL(2, \mathbf{Z})) \simeq \mathbf{Z}_2 \oplus PGL(2, \mathbf{Z})$. Thus by 1.6(1) and 4.4,

$$\mathscr{A}^2(SL(2, \mathbf{Z})) \simeq M(1, 2) \times_\sigma Y$$

where $Y = L \times_\rho D \simeq \mathscr{A}(PGL(2, \mathbf{Z}))$, and $\sigma\colon Y \to \mathscr{A}(M(1, 2))$ is defined by

$$L \times_\rho \langle a^2, b \rangle = \operatorname{Ker} \sigma, \qquad \sigma_a(X) = Xe_{12}.$$

Next, by 4.2(3) and 1.4(4),

$$\mathscr{A}^3(SL(2, \mathbf{Z})) \simeq \{\operatorname{Der}_\sigma (Y, M(1, 2)) \times_{\mu_1} \mathscr{A}_\sigma(M(1, 2))\} \times_{\mu_2} Y$$

where $\mu_{1,\alpha}(\delta) = \alpha\delta$ and $\mu_{2,y}(\delta, \alpha) = (\delta \circ I(y^{-1}), \alpha)$. First,

$$\mathscr{A}_\sigma(M(1, 2)) \simeq \mathscr{C}_{SL(2,\mathbf{Z}_2)}(\operatorname{Im} \sigma) = \operatorname{Im} \sigma.$$

Next, let $\delta\colon Y \to M(1, 2)$ be a $\sigma$-derivation. Since $L < \operatorname{Ker} \sigma$ and the restriction of $\delta$ to $\operatorname{Ker} \sigma$ is a homomorphism, $\delta(L) = 0$ and $\delta$ is determined by $\delta(a)$, $\delta(b) \in M(1, 2)$. These must satisfy $\delta(a^4) = \delta(b^2) = \delta(abab) = 0$, which reduce to $(\delta(a) + \delta(b))m_{12} = 0$. Therefore the first entries of $\delta(a)$, $\delta(b)$ coincide and we may write $\delta(a) = (e, f)$, $\delta(b) = (e, g)$. Define the map

$$\operatorname{Der}_\sigma (Y, M(1, 2)) \times_{\mu_1} \langle e_{12} \rangle \to \mathbf{Z}_2 \oplus D$$

by $(\delta, e_{12}^h) \to (c^{f+g}, a^{2g+e}b^h)$ (where $\delta$ is given above). By direct computation, this map is an isomorphism, and the map $\mu_2$ corresponds to $\mu_2\colon Y \to \mathscr{A}(\mathbf{Z}_2 \oplus D)$ given by

$$L < \operatorname{Ker} \mu_2, \; \mu_{2,a}(c^r, a^s b^t) = (c^{r+s}, a^{-s}b^t) \quad \text{and} \quad \mu_{2,b}(c^r, a^s b^t) = (c^{r+s}, a^s b^t).$$

Adjust the presentation of $\mathscr{A}^3(SL(2, \mathbf{Z}))$, replacing $((1, 1), a)$ by $((1, b), a)$, to obtain

$$\mathscr{A}^3(SL(2, \mathbf{Z})) \simeq (\mathbf{Z}_2 \oplus D) \times_\mu Y$$

where $\mu|_L = 1$ and $\mu_a = \mu_b\colon (c^r, a^s b^t) \to (c^{r+s}, a^s b^t)$.

Since $\mathbf{Z}_2 \oplus D$ is characteristic in $\mathscr{A}^3(SL(2, \mathbf{Z}))$, we apply 1.3(5) to obtain

$$\mathscr{A}^4(SL(2, \mathbf{Z})) \simeq N \times_\rho Y,$$

where

$$N = \{\alpha, \delta) \in \mathscr{A}(\mathbf{Z}_2 \oplus D) \times \mathrm{Der}_\mu (Y, \mathbf{Z}_2 \oplus D) \,|\, [\alpha, \mu_x] = I(\delta(x)) \text{ for all } x \in Y\}$$

and $\rho_y(\alpha, \delta) = (\mu_y \alpha \mu_y^{-1}, \mu_y \delta I(y^{-1}))$. Since $L < \operatorname{Ker} \mu$, $L < \operatorname{Ker} \delta$; and so any derivation is determined by $\delta(a), \delta(b) \in \mathbf{Z}_2 \oplus D$ which must satisfy

$$1 = (\delta(a) \;\#\; \mu_a \delta(a))^2 = \delta(b) \;\#\; \mu_b \delta(b) = \delta(ab) \;\#\; \delta(ab).$$

These equations are equivalent to $\delta(a) = (c^{r_1}, a^{s_1} b^{k_1})$, $\delta(b) = (c^{r_2}, a^{2g_2} b^{k_2})$ and

$$(*) \qquad (a^{s_1} b^{k_1 + k_2})^2 = 1.$$

Now let $\alpha \in \mathscr{A}(\mathbf{Z}_2 \oplus D)$. The derived group of $\mathbf{Z}_2 \oplus D$ is $1 \oplus \langle a^2 \rangle$, and

$$\mathscr{C}(\mathbf{Z}_2 \oplus D) = \langle c \rangle \oplus \langle a^2 \rangle.$$

Hence

$$\alpha(1, a) = (c^{e_1}, a^{1+2f_1}), \quad \alpha(1, b) = (c^{e_2}, a^h b), \quad \alpha(c, 1) = (c, a^{2k}).$$

The equalities $\alpha\mu_b = I(\delta(b))\mu_b \alpha$ and $\alpha\mu_a = I(\delta(a))\mu_a \alpha$ evaluated at $(1, a)$ together with $(*)$ above impose the conditions $k = k_1 = k_2, s_1 \equiv 0 \bmod 2$. Then evaluate on $(1, b)$ to obtain $h \equiv 0 \bmod 2$. These exhaust the restrictions on $\alpha, \delta$ so we have now $(\alpha, \delta) \in N$ if and only if

$$\begin{aligned}
&\delta(a) = (c^{r_1}, a^{2g_1} b^k), \quad \delta(b) = (c^{r_2}, a^{2g_2} b^k) \\
(**) \qquad &\alpha(1, a) = (c^{e_1}, a^{1+2f_1}), \quad \alpha(1, b) = (c^{e_2}, a^{2f_2} b), \\
&\alpha(c, 1) = (c, a^{2k}),
\end{aligned}$$

where all exponents are taken mod 2. A computation shows that $\rho\colon Y \to \mathscr{A}(N)$ is given by $\rho\,|_L = 1$ and

$$\rho_a = \rho_b \colon (\alpha, \delta) \to (I(1, b^k)) \circ \alpha, \delta).$$

Let $(A, 1) \in N$ be defined by

$$A\colon (1, a) \to (c, a), (1, b) \to (1, b), (c, 1) \to (c, 1).$$

Then $\rho_a(A, 1) = (A, 1)$ and for all $(\alpha, \delta) \in N$,

$$\rho_a(\alpha, \delta) = (A, 1) \cdot (\alpha, \delta) \cdot (A, 1)^{-1}.$$

We may therefore adjust the presentation of $\mathscr{A}^4(SL(2, \mathbf{Z}))$, replacing $a, b \in Y$ by $((A, 1), a)$, $((A, 1), b)$, to obtain $\mathscr{A}^4(SL(2, \mathbf{Z})) \simeq N \oplus Y$. Finally, $N \simeq M(8, 1) \times_\lambda \langle Z \rangle$; to see this, to $(\alpha, \delta)$ given by $(**)$ associate $(X, Z^k)$ where

$$X^t = (e_1, e_2, r_1 + (k+1)e_2 + f_2, r_2 + (k+1)e_2 + f_2, f_1, f_2, g_1 + kf_2, g_2 + kf_2).$$

This map is a homomorphism (using the product in $N$ given in 1.3(5)), and is clearly injective. That the centers of these automorphism groups are as described above follows by computing $\mathscr{C}_{M(1,2)}(\mathrm{Im}\,\sigma)$, $\mathscr{C}_{\mathbf{Z}_2 \oplus D}(\mathrm{Im}\,\mu)$, and $\mathscr{C}(N) = (I_8 + Z)M(8, 1)$. ∎

COROLLARY 4.8. *For all* $k \geq 0$,

$$\mathscr{A}^{4+k}(SL(2, \mathbf{Z})) \simeq H_k \oplus \mathscr{A}(PGL(2, \mathbf{Z})),$$

*where* $H_0 = M(8, 1) \times_\lambda \langle Z \rangle$ *and for* $k \geq 1$, $H_{k+1} = \mathrm{Hom}\,(\mathbf{Z}_2 \oplus \mathbf{Z}_2, \mathscr{C}(H_k)) \times_* \mathscr{A}(H_k)$.

*Proof.* Since $H_0$ is finite, so are all $H_k$. We have $\mathscr{A}(PGL(2, \mathbf{Z})) \simeq L \times_\rho D$, so $\mathscr{A}(PGL(2, \mathbf{Z}))/\mathscr{A}(PGL(2, \mathbf{Z}))' \simeq \mathbf{Z}_2 \oplus \mathbf{Z}_2$. Therefore 4.2(3) together with 1.3(5) yields the result above. ∎

Note that, if $\mathscr{C}(H_k) = 1$, then $H_{k+r} = \mathscr{A}^r(H_k)$ and the finiteness of the automorphism sequence of $SL(2, \mathbf{Z})$ is a consequence of Wielandt's Theorem. The groups $H_k$, and their centers, will be determined explicitly in the following propositions. The notation is that of Section 3.

PROPOSITION 4.9. $H_1 \simeq (M(4, 3) \oplus M(4)) \times_\rho SL(4, \mathbf{Z}_2)$ *and* $\mathscr{C}(H_1) \simeq \mathbf{Z}_2$, *with action*

$$(X, Y) \to (AX, AYA^{-1}).$$

*Proof.* We have $H_0 = M(8, 1) \times_\lambda \langle Z \rangle$, and for any $X \in M(8, 1)$,

$$\mathscr{C}_{H_0}\{(X, Z)\} \subset \langle \mathscr{C}(H_0), Z \rangle, \mathscr{C}_{H_0}\{(X, 1)\} \supset M(8, 1).$$

Therefore $M(8, 1)$ is characteristic in $H_0$; and so 1.5(1) implies

$$\mathscr{A}(H_0) \simeq \mathrm{Der}_\lambda\,(\langle Z \rangle, M(8, 1)) \times_\mu \mathscr{C}_{SL(8,\mathbf{Z}_2)}\{Z\},$$

where $\mu_\alpha(\delta) = \alpha\delta$. A computation shows that

$$\mathscr{C}_{SL(8,\mathbf{Z}_2)}\{Z\} = \left\{ \begin{pmatrix} A & 0 \\ Y & A \end{pmatrix} \mid A \in SL(4, \mathbf{Z}_2),\ Y \in M(4) \right\}.$$

Any derivation is determined by $\delta(Z)$ which is subject to the one restriction $(I_8 + Z)\delta(Z) = 0$. Therefore $\delta(Z) \in (I + Z)M(8, 1) = \mathscr{C}(H_0)$ and is otherwise arbitrary. View $\mathscr{C}(H_0)$ as $M(4, 1)$ imbedded in $M(8, 1)$ naturally as $(I + Z)M(8, 1)$. Then $\mu_\alpha$ corresponds to left multiplication by $A$, where

$$\alpha = \begin{pmatrix} A & 0 \\ Y & A \end{pmatrix}.$$

Thus

$$\mathscr{A}(H_0) \simeq \{M(4, 1) \oplus M(4)\} \times_\rho SL(4, \mathbf{Z}_2),$$

where $(\delta, \alpha) \in \mathscr{A}(H_0)$ corresponds to $((\delta(Z), Y), A)$.

The action of $\mathscr{A}(H_0)$ on Hom $(\mathbf{Z}_2 \oplus \mathbf{Z}_2, \mathscr{C}(H_0)) \simeq \mathscr{C}(H_0) \oplus \mathscr{C}(H_0)$ is given by the restriction of $\mathscr{A}(H_0)$ to $\mathscr{C}(H_0) \simeq M(4, 1)$; thus the action corresponds to left multiplication by $A$, whence

$$H_1 \simeq (M(4, 3) \oplus M(4)) \times_\rho SL(4, \mathbf{Z}_2).$$

Since $\mathscr{C}(SL(4, \mathbf{Z}_2)) = 1$, $\mathscr{C}(H_1)$ is the fixed point set of $SL(4, \mathbf{Z}_2)$ in $M(4, 3) \oplus M(4)$. Hence $\mathscr{C}(H_1) = \langle (0, I_4) \rangle \simeq \mathbf{Z}_2$ (see 3.1(1), (2)). ∎

Proposition 4.10. *$H_2 \simeq M(3, 1) \oplus K$, where*

$$K = (M(4, 3) \oplus PM(4)) \times_\sigma (SL(4, \mathbf{Z}_2) \oplus SL(3, \mathbf{Z}_2))$$

*and* $\sigma_{(A,B)}(X, Y) = (AXB^{-1}, AYA^{-1})$. *Moreover, $K$ is complete and* $\mathscr{C}(H_2) \simeq M(3, 1)$.

*Proof.* By 4.8 and 4.9,

$$H_2 = \text{Hom}\ (\mathbf{Z}_2 \oplus \mathbf{Z}_2, \mathbf{Z}_2) \times_* \mathscr{A}(H_1)$$
$$\simeq M(1, 2) \oplus \mathscr{A}(H_1) \simeq M(2, 1) \oplus \mathscr{A}(H_1).$$

Since $SL(4, \mathbf{Z}_2)$ is simple, $M(4, 3) \oplus M(4)$ is characteristic in $H_1$ so that we may apply 1.4(2) to $\pi\colon \mathscr{A}(H_1) \to \mathscr{A}(SL(4, \mathbf{Z}_2))$. It is a classical theorem that

$$\mathscr{A}(SL(4, \mathbf{Z}_2)) = I(SL(4, \mathbf{Z}_2)) \times_\gamma \mathbf{Z}_2,$$

where the nontrivial element in Im $\gamma$ is the inverse-transpose map [3]. Let $S_1 < SL(4, \mathbf{Z}_2)$ be the subgroup described in 3.4:

$$S_1 = \left\{ \begin{pmatrix} 1 & * & * & * \\ 0 & & & \\ 0 & & A & \\ 0 & & & \end{pmatrix} \middle| A \in SL(3, \mathbf{Z}_2) \right\}.$$

The fixed point set of $S_1$ in $M(4, 3) \oplus M(4)$ has order $2^4$, while its image under inverse-transpose has fixed point set of order 2. Consequently Im $\pi = I(SL(4, \mathbf{Z}_2))$ and we may now apply 1.5(2) to obtain $\mathscr{A}(H_1) \simeq D \times_\mu F$ where

$$D = \text{Der}_\rho\ (SL(4, \mathbf{Z}_2), M(4, 3) \oplus M(4)) \simeq M(4, 3) \oplus PM(4),$$

and

$$F = \mathscr{A}_\rho(M(4, 3) \oplus M(4)) \oplus SL(4, \mathbf{Z}_2) \simeq (SL(3, \mathbf{Z}_2) \oplus \mathbf{Z}_2) \oplus SL(4, \mathbf{Z}_2)$$

(using 3.1(1), (2) and 3.4(1)). Since $\mu_{(\alpha, A)}(X, Y) = \alpha(AX, AYA^{-1})$, it follows that $\mathscr{A}(H_1) \simeq K \oplus \mathbf{Z}_2$; and so $H_2 \simeq M(1, 3) \oplus K$. Therefore $\mathscr{C}(H_2) \simeq M(3, 1)$, and it remains to prove that $I\colon K \to \mathscr{A}(K)$ is an epimorphism.

$M(4, 3) \oplus PM(4)$ is the maximal normal nilpotent subgroup of $K$, since $SL(3, \mathbf{Z}_2)$ and $SL(4, \mathbf{Z}_2)$ are simple. Furthermore, up to an inner automorphism, any $\alpha \in \mathscr{A}(SL(3, \mathbf{Z}_2) \oplus SL(4, \mathbf{Z}_2))$ is of the form $\alpha_1 \oplus \alpha_2$ where each $\alpha_i$

is either 1 or inverse-transpose. As above, only $1 \oplus 1$ extends to an automorphism of $K$, so that 1.5(2) again applies. We obtain $\mathscr{A}(K) \simeq D \times_* F$ where now (using 3.1, 3.2 and the additivity of Der),

$$\begin{aligned} D &= \operatorname{Der}_\sigma (SL(3, \mathbf{Z}_2) \oplus SL(4, \mathbf{Z}_2), M(4, 3) \oplus PM(4)) \\ &\simeq \operatorname{Der}_\sigma (SL(3, \mathbf{Z}_2) \oplus SL(4, \mathbf{Z}_2), M(4, 3)) \\ &\qquad\qquad \oplus \operatorname{Der}_\sigma (SL(3, \mathbf{Z}_2) \oplus SL(4, \mathbf{Z}_2), PM(4)) \\ &\simeq M(4, 3) \oplus \operatorname{Der}_\kappa (SL(4, \mathbf{Z}_2), PM(4)) \oplus \operatorname{Hom} (SL(3, \mathbf{Z}_2), PM(4)) \\ &\simeq M(4, 3) \oplus PM(4) \oplus 1, \end{aligned}$$

and

$$\begin{aligned} F &= (SL(3, \mathbf{Z}_2) \oplus SL(4, \mathbf{Z}_2)) \oplus \mathscr{A}_\sigma(M(4, 3) \oplus PM(4)) \\ &\simeq SL(3, \mathbf{Z}_2) \oplus SL(4, \mathbf{Z}_2) \oplus 1. \end{aligned}$$

Consequently $|\mathscr{A}(K)| = |D| \cdot |F| = |I(K)|$ so $I(K) = \mathscr{A}(K)$. ∎

Proposition 4.11. *$H_3 \simeq \{M(3, 3) \times_\lambda SL(3, \mathbf{Z}_3)\} \oplus K$ where $K$ is the group described in* 4.10, *and $\mathscr{C}(H_3) = 1$.*

*Proof.* We have

$$K = (M(4, 3) \oplus PM(4)) \times_\sigma (SL(4, \mathbf{Z}_2) \oplus SL(3, \mathbf{Z}_2)),$$

so the map $\phi$: $K \to \mathbf{Z}_2$ defined by $((X, Y), (A, B)) \to \operatorname{Tr} Y$ is a homomorphism whose kernel contains $K'$. We claim $K' = \operatorname{Ker} \phi$. Since $SL(4, \mathbf{Z}_2) \oplus SL(3, \mathbf{Z}_2)$ is perfect and

$$\{(A - I)X \mid X \in M(4, 3), A \in SL(4, \mathbf{Z}_2)\} = M(4, 3),$$

we need only show that $m_{ij}^*$ and $m_{ii}^* + m_{jj}^*$ $(i \neq j)$ belong to $K'$; this follows from the equations

$$\begin{aligned} m_{ij}^* &= m_{kj}^* + e_{ik} m_{kj}^* e_{ik} \quad (i, j, k \text{ distinct}), \\ m_{ii}^* + m_{jj}^* &= m_{ij}^* + e_{ji} m_{ij}^* e_{ji} + m_{ji}^*. \end{aligned}$$

By 4.8, we have

$$H_3 = \operatorname{Hom} (\mathbf{Z}_2 \oplus \mathbf{Z}_2, \mathscr{C}(H_2)) \times_* \mathscr{A}(H_2) \simeq M(3, 2) \times_* \mathscr{A}(H_2).$$

Since $H_2 \simeq M(3, 1) \oplus K$, we apply 1.6(2) to obtain

$$\mathscr{A}(H_2) \simeq \operatorname{Hom} (K, M(3, 1)) \times_* \mathscr{A}(M(3, 1)) \oplus K.$$

But $K/K' \simeq \mathbf{Z}_2$, so $\mathscr{A}(H_2) \simeq M(3, 1) \times_\lambda SL(3, \mathbf{Z}_2) \oplus K$. Note that $\mathscr{A}(H_2) \to \mathscr{A}(\mathscr{C}(H_2))$ is given by

$$M(3, 1) \times_\lambda SL(3, \mathbf{Z}_2) \oplus K \to (0) \times_* SL(3, \mathbf{Z}_2) \oplus 1 \simeq SL(3, \mathbf{Z}_2).$$

Consequently $H_3 \simeq M(3, 3) \times_\lambda SL(3, \mathbf{Z}_2) \oplus K$, and $\mathscr{C}(H_3) = 1$ as required. ∎

PROPOSITION 4.12. *$H_4 \simeq \{M(4, 3) \times_\tau (SL(3, \mathbf{Z}_2) \oplus SL(3, \mathbf{Z}_2))\} \oplus K$, and $H_4$ is complete.*

*Proof.* We have $H_3 \simeq \{M(3, 3) \times_\lambda SL(3, \mathbf{Z}_2)\} \oplus K$, and $H_4 = \mathscr{A}(H_3)$. Now

$$V = M(3, 3) \oplus M(4, 3) \oplus PM(4)$$

is a normal subgroup of $H_3$ with $H_3/V \simeq SL(3, \mathbf{Z}_2) \oplus SL(3, \mathbf{Z}_2) \oplus SL(4, \mathbf{Z}_2)$. This extension splits, with action

$$\mu_{(B_1, B_2, A)}(X, Y, Z) = (B_1 X, AYB_2^{-1}, AZA^{-1}).$$

Now $V$ is the maximal normal nilpotent subgroup of $H_3$, hence characteristic, while $SL(4, \mathbf{Z}_2)$ is characteristic in $H/V$ (if not, there is a nontrivial homomorphism $SL(4, \mathbf{Z}_2) \to SL(3, \mathbf{Z}_2)$. See also [14, Lemma 1.4]). Thus

$$V \times_\mu (1 \oplus 1 \oplus SL(4, \mathbf{Z}_2))$$

is characteristic in $H_3$, as is its center $M(3, 3)$. But $\mathscr{C}_{H_3}(M(3, 3)) \simeq M(3, 3) \oplus K$, which has derived group $K'$. Finally $\mathscr{C}_{H_3}(K') = M(3, 3) \times_\lambda SL(3, \mathbf{Z}_2)$ and is characteristic in $H_3$, as is its centralizer $K$. Consequently

$$H_4 = \mathscr{A}(H_3) \simeq \mathscr{A}(M(3, 3) \times_\lambda SL(3, \mathbf{Z}_2)) \oplus K.$$

As in 4.10, using 1.4(2) and 1.5(2),

$$\mathscr{A}(M(3, 3) \times_\lambda SL(3, \mathbf{Z}_2))$$
$$\simeq \mathrm{Der}_\lambda (SL(3, \mathbf{Z}_2), M(3, 3)) \times_\tau \{SL(3, \mathbf{Z}_2) \oplus \mathscr{A}_\lambda(M(3, 3))\}$$

with $\tau_{(A,\alpha)}(\delta) = \alpha(A \cdot \delta \cdot I(A^{-1}))$. By 3.3, the group of derivations is isomorphic to $M(4, 3)$, and by 3.4, $\mathscr{A}_\lambda(M(3, 3)) \simeq SL(3, \mathbf{Z}_2)$ where $B \in SL(3, \mathbf{Z}_2)$ acts by $X \to XB^{-1}$. The action of $SL(3, \mathbf{Z}_2)$ on $\mathrm{Der}_\lambda$ is also by $Y \to YB^{-1}$.

To show $H_4$ complete, we prove that $H_3$ is characteristic in $H_4$ and apply Burnside's criterion. The argument of the first paragraph applies to

$$V = M(4, 3) \oplus (M(4, 3) \oplus PM(4)) \lhd H_4$$

where now

$$H_4/V \simeq SL(3, \mathbf{Z}_2) \oplus SL(3, \mathbf{Z}_2) \oplus SL(3, \mathbf{Z}_2) \oplus SL(4, \mathbf{Z}_2).$$

We conclude that $K$ and $\mathscr{A}(M(3, 3) \times_\lambda SL(3, \mathbf{Z}_2))$ are characteristic in $H_4$, so it suffices to prove that $L = M(3, 3) \times_\lambda SL(3, \mathbf{Z}_2)$ is a characteristic subgroup of

$$\mathscr{A}(L) \simeq M(4, 3) \times_\tau \{SL(3, \mathbf{Z}_2) \oplus SL(3, \mathbf{Z}_2)\}.$$

We have $M(4, 3) \simeq M(3, 3) \oplus M(1, 3)$ (where the first summand corresponds to the group of inner derivations) characteristic in $\mathscr{A}(L)$. But $M(3, 3)$ is an irreducible $\mathscr{A}(L)/M(4, 3)$-module, for $\tau_{(A,B)}(X) = AXB^{-1}$, $X \in M(3, 3)$. Moreover, if $\delta$ is any derivation and $A \in SL(3, \mathbf{Z}_2)$,

$$\tau_{(A,I)}(\delta) = A\delta I(A^{-1}) = \delta + I(\delta(A)),$$

where $I(\delta(A))$ is the inner derivation determined by $\delta(A)$. Consequently either $I(\delta(A)) = 0$ so that $\delta = 0$, or $M(3, 3)$ is contained in the submodule generated by $\delta$. Thus $M(3, 3)$ is characteristic in $\mathscr{A}(L)$, and

$$\begin{aligned}\mathscr{A}(L)/I(M(3, 3)) &\simeq M(1, 3) \times_\tau \{I(SL(3, \mathbf{Z}_2)) \oplus SL(3, \mathbf{Z}_2)\} \\ &\simeq I(SL(3, \mathbf{Z}_2)) \oplus \{M(3, 1) \times_\lambda SL(3, \mathbf{Z}_2)\}.\end{aligned}$$

Here $M(3, 1)$ is characteristic, and its centralizer is $I(SL(3, \mathbf{Z}_2)) \oplus M(3, 1)$, with derived group $I(SL(3, \mathbf{Z}_2))$. Thus $L = I(L)$ and is characteristic in $\mathscr{A}(L)$, as required. ▌

## REFERENCES


1. R. BAER, *Absolute retracts in group theory*, Bull. Amer. Math. Soc., vol. 52 (1946), pp. 501–506.
2. W. BURNSIDE, *Theory of groups of finite order*, Dover, New York, 1955.
3. J. DIEUDONNE, *On the automorphisms of the classical groups*, Mem. Amer. Math. Soc., no. 2, 1951.
4. D. HERTZIG, *Cohomology of algebraic groups*, J. Algebra, vol. 6 (1967), pp. 317–334.
5. L.-K. HUA AND I. REINER, *Automorphisms of the unimodular group*, Trans. Amer. Math. Soc., vol. 71 (1951), pp. 331–348.
6. ———, *Automorphisms of the projective unimodular group*, Trans. Amer. Math. Soc., vol. 72 (1952), pp. 467–473.
7. J. A. HULSE, *Automorphism towers of polycyclic groups*, J. Algebra, vol. 16 (1970), pp. 347–398.
8. W. MAGNUS, A. KARRASS, AND D. SOLITAR, *Combinatorial group theory*, Wiley, New York, 1966.
9. W. H. MILLS, *The automorphisms of the holomorph of a finite abelian group*, Trans. Amer. Math. Soc., vol. 85 (1957), pp. 1–34.
10. J. MILNOR, *Introduction to algebraic K-theory*, Princeton University Press, Princeton, N.J., 1971.
11. M. NEWMAN, *Integral matrices*, Academic Press, New York, 1972.
12. O. T. O'MEARA, *The automorphisms of the linear groups over any integral domain*, J. Reine Angew. Math., vol. 223 (1966), pp. 56–100.
13. ———, *Lectures on linear groups*, Conference Board of the Mathematical Sciences, Regional Conference Series in Mathematics, no. 22, 1973.
14. J. S. ROSE, *Automorphism groups of groups with trivial centre*, Proc. London Math. Soc., series 3, vol. 31 (1975), pp. 167–193.
15. R. SOLAZZI, *The automorphisms of certain subgroups of $PGL_n(V)$*, Illinois J. Math., vol. 16 (1972), pp. 330–348.
16. R. STEINBERG, *Lectures on Chevalley groups*, mimeographed, Yale, 1967.
17. C.-H. WAN, *Automorphisms of the modular group* (Chinese, English summary), Advances in Math., vol. 3 (1957), pp. 216–233.
18. H. WIELANDT, *Eine Verallgemeinerung der invarianten Untergruppen*, Math. Zeitschr., vol. 45 (1939), pp. 209–244.
19. K.-W. YANG, *Isomorphisms of group extensions*, Pacific J. Math., vol. 50 (1974), pp. 299–304.
20. M.-C. YING, *Automorphisms of the projective unimodular group* (Chinese, English summary), Acta Math. Sinica, vol. 10 (1960), pp. 55–65.


LEHMAN COLLEGE AND THE GRADUATE CENTER, CITY UNIVERSITY OF NEW YORK
NEW YORK