JOURNAL OF ALGEBRA 125, 215–235 (1989)

Overgroups in $GL(nr, F)$ of Certain Subgroups of $SL(n, K)$, I

SHANGZHI LI

Department of Mathematics, University of Science and Technology of China, Hefei, Anhui, China

Communicated by Walter Feit

Received February 16, 1988

For any pair of division rings K and F with $K \supset F$ and $\dim_F K = r$ we determine all the overgroups of $SL(n, K)$ in $GL(nr, F)$, as well as the overgroups of $\mathrm{Sp}(n, K)$ in $GL(nr, F)$ (for commutative K and even n). The overgroups of $SU(n, K, f)$ and $\Omega(n, K, Q)$ in $GL(nr, F)$ will be determined in another paper, "Overgroups in $GL(nr, F)$ of certain subgroups of $SL(n, K)$, II".

Let K, F be division rings, with $K \supset F$, and $\dim_F K = r < \infty$ when we regard K as a left F-space. An n-dimensional left K-space $V(n, K)$ can be regarded as an nr-dimensional left space $V = V(nr, F)$ over F; thus the $GL(n, K)$ acting on $V(n, K)$ is embedded in the $GL(nr, F)$ acting on $V(nr, F)$. The purpose of this paper is to determine the overgroups of $N = SL(n, K)$ in $G = GL(nr, F)$, and the overgroups of $N_1 = \mathrm{Sp}(n, K, f)$ in $G = GL(nr, F)$ for commutative K, even n, and any non-degenerate alternating K-form f. For finite fields K there has been some related work. Kantor [2] determined the overgroups of $GL(1, q^r)$ in $GL(r, q)$ and thus determined the overgroups of $GL(n, q^r)$ in $GL(nr, q)$ (since $GL(n, q^r) > GL(1, q^{nr})$). Dye [3], [4], [5] studied the maximality of the normalizers of $\mathrm{Sp}(n, q^r)$ in $\mathrm{Sp}(nr, q)$ for prime $r = 2$ or 3. In this paper we shall settle all the cases for all the division rings and all dimensions $n \geqslant 2$ (thus the only remaining task is to determine the overgroups of $SL(1, K)$ in $GL(r, F)$). Applying our main results to the special cases for finite K and prime r we obtain the maximality of the subgroups in Aschbacher's class C_3 defined in [1].

I am indebted to G. M. Seitz for his helpful suggestions.

The main results of this paper are as follows.

THEOREM 1. *Let* $N = SL(n, K) \leqslant X \leqslant G = GL(nr, F)$; *then one of the following holds:*

0021-8693/89 $3.00

(i) *$SL(nd, E) \trianglelefteq X \leqslant GL(nd, E) \rtimes \text{Aut}\, E/F$ for a division ring E intermediate between F and K, where $d = \dim_E K$.*

(ii) *$n = 2$, K is commutative, $N = SL(2, K) = \text{Sp}(2, K, f)$ for any nondegenerate alternating K-form f, and $X \trianglerighteq \text{Sp}(2d, E. f_E)$ for an intermediate field E ($F \subseteq E \subseteq K$ and $d = \dim_E K$) and an alternating E-form $f_E = \varphi_E f$ with $0 \neq \varphi_E \in \text{Hom}_E(K, E)$.*

(iii) *$N = SL(2, 4) \cong A_5$ and $G = GL(4, 2) \cong A_8$, $X = (\text{Sp}(4, 2))' \cong A_6$, or $X \cong A_7$.*

COROLLARY 1. *Let F be a maximal division subring of K, $GL(nr, F) \geqslant G^* \geqslant SL(nr, F)$, then the normalizer M of $SL(n, K)$ ($n \geqslant 2$) in G^* is a maximal subgroup of G^*, provided $G^* = M \cdot SL(nr, F)$, with an exception $SL(2, K) = \text{Sp}(2, K)$ and $\text{Norm}_{K/F}^{-1}(\det G^*) \subseteq F^*$ (where $\text{Norm}_{K/F}$ denotes the norm mapping of K into F). When $SL(2, K) = \text{Sp}(2, K)$ and $\text{Norm}_{K/F}^{-1}(\det G^*) \subseteq F^*$ we have $M \lneqq (\text{GSp}(2r, F) \rtimes \text{Aut}\, K/F) \cap G^* \lneqq G^*$ thus M is not maximal in G^*.*

THEOREM 2. *Let K and F be commutative fields with $K \supset F$ and $\dim_F K = r < \infty$, $n = 2v$ be even, $N_1 = \text{Sp}(n, K, f) \leqslant X \leqslant G = GL(nr, F)$, then we have an intermediate field E between F and K (with $\dim_E K = d$) such that (i) $SL(nd, E) \trianglelefteq X \leqslant GL(nd, E) \rtimes \text{Aut}\, E/F$, or (ii) $\text{Sp}(nd, E, f_E) \trianglelefteq X \leqslant \text{GSp}(nd, E, f_E) \rtimes \text{Aut}\, E/F$ relative to an $f_E = \varphi_E f$ for a $0 \neq \varphi_E \in \text{Hom}_E(K, E)$, except in the case (iii) $N_1 = SL(2, 4) \cong A_5$ and $G = SL(4, 2) \cong A_8$, $X = (\text{Sp}(4, 2))' \cong A_6$ or $X \cong A_7$.*

COROLLARY 2. *Let F be a maximal subfield of K, $G\,\text{Sp}(2vr, F) \geqslant G^* \geqslant \text{Sp}(2vr, F)$, then the normalizer M of $\text{Sp}(2v, K)$ in G^* is a maximal subgroup in G^*, provided $G^* = M \cdot \text{Sp}(2vr, F)$, except in the case $G^* = \text{Sp}(4, 2)$ and $M = \text{Sp}(2, 4)$.*

The proofs in this paper are mainly based on matrix techniques. We need the following notations for vector spaces and matrices. For any division ring E, E-spaces will mean left vector spaces over E. $\langle S \rangle_E$ denotes the left E-subspace spanned by a subset S of an E-space; specifically, we write $\langle S \rangle$ instead of $\langle S \rangle_F$. For a ring R, we denote by $\text{Mat}_{m \times n} R$ the set of all $m \times n$ matrices over R, and we write $\text{Mat}_m R$ instead of $\text{Mat}_{m \times m} R$. We write $A^{(m \times n)}$ (resp. $A^{(m)}$) to suggest that $A \in \text{Mat}_{m \times n} F$ (resp. $A \in \text{Mat}_m F$). $A' \in \text{Mat}_{n \times m} R$ denotes the transpose of $A \in \text{Mat}_{m \times n} R$. We denote by $\text{diag}(A_1, ..., A_k)$ the quasi-diagonal matrix with $A_1, ..., A_k$ as its diagonal blocks. I and 0 denote the identity matrix and zero matrix, respectively. $E_{ij} \in \text{Mat}_{m \times n} R$ denotes the matrix in $\text{Mat}_{m \times n} R$ having a single 1 as its (i, j)th entry and all other entries zero. When $k \neq l$ we denote by

$P_{kl} \in GL(n, R)$ the matrix $P_{kl} = I - E_{kk} - E_{ll} + E_{kl} - E_{lk}$ in $GL(n, R)$, and we define $P_{kk} = I$, thus $P_{kl} \in SL(n, R)$ anyway. For $k \neq l$ and $c \in R$, $T_{kl}(c) \in SL(n, R)$ denotes the matrix $T_{kl}(c) = I + cE_{kl}$ in $\mathrm{Mat}_n R$. When R is a division ring, $T_{kl}(c) \in SL(n, R)$ (with $c \neq 0$) is a transvection of $SL(n, R)$, and all $T_{kl}(c)$ ($k \neq l$, $0 \neq c \in R$) generate $SL(n, R)$. When we write $T_{kl}(A^{(r)})$ it suggests that $T_{kl}(A^{(r)}) \in \mathrm{Mat}_n(\mathrm{Mat}_r F)$, namely, $T_{kl}(A^{(r)}) = (C_{ij}^{(r)})_{n \times n}$ with the (k, l)th block $C_{kl} = A$, the blocks $C_{ii} = I$ $(1 \leqslant i \leqslant n)$, and all other blocks $C_{ij} = 0$. For each $A \in \mathrm{Mat}_{m \times n} F$, $\mathrm{Im}\, A$ and $\mathrm{Ker}\, A$ denote the image and the kernel, respectively, of the associated F-linear mapping $\mathrm{Mat}_{1 \times m} F \to \mathrm{Mat}_{1 \times n} F$: $x \mapsto xA$; we have $\dim_F \mathrm{Im}\, A = \mathrm{rank}\, A$ and $m = \mathrm{rank}\, A + \dim_F \mathrm{Ker}\, A$.

Notations of group theory are as usual. For example, $\langle S \rangle$ denotes the subgroup generated by a subset S of a group; $X \rtimes Y$ denotes a semi-direct product of X by Y; X' denotes the commutator group of X, generated by all the commutators $[g_1, g_2] = g_1^{-1} g_2^{-1} g_1 g_2$ with $g_1, g_2 \in X$; for a group X acting on a set S, $X_{A, B, \ldots}$ consists of all the elements in X stabilizing each of the subsets $A, B, \ldots$ of S.

Now we come back to consider our F-space $V = V(nr, F)$ obtained from $V(n, K)$. Take a left K-basis $\{e_1, \ldots, e_n\}$ of $V(n, K)$ and a left F-basis $\{k_1, \ldots, k_r\}$ of K, then $\{e_{ij} = k_j e_i \mid 1 \leqslant i \leqslant n,\ 1 \leqslant j \leqslant r\}$ forms an F-basis of V. With respect to this basis we write all vectors in $V = V(nr, F)$ as nr-dimensional rows (i.e., write V as $\mathrm{Mat}_{1 \times nr} F$) and write each $g \in GL(nr, F)$ as a matrix in $\mathrm{Mat}_{nr} F$, sending each $x \in V$ to xg. When viewing K as a left F-space we write K as $\vec{K}$ and denote each $\vec{x} = c_1 \vec{k}_1 + \cdots + c_r \vec{k}_r \in \vec{K}$ (with all $c_i \in F$) by $\vec{x} = (c_1, \ldots, c_r) \in \mathrm{Mat}_{1 \times r} F$, thus identifying $\vec{K}$ with $\mathrm{Mat}_{1 \times r} F$. On the other hand, each $\theta \in K$ can be viewed as an F-linear transformation $\vec{x} \mapsto \overrightarrow{x\theta}$ on $\vec{K}$ identified with the matrix $\theta^{(r)} \in \mathrm{Mat}_r F$ of this transformation relative to the basis $\{k_1, \ldots, k_r\}$. In this point of view we have $\mathrm{Mat}_n K \subset \mathrm{Mat}_n(\mathrm{Mat}_r F) = \mathrm{Mat}_{nr} F$; thus each $A \in \mathrm{Mat}_n K \subset \mathrm{Mat}_{nr} F$ has a $\mathrm{rank}\, A$ over F and a $\mathrm{rank}_K A$ over K, and we have $\mathrm{rank}\, A = r \cdot \mathrm{rank}_K A$. We shall also write each $\sigma \in \mathrm{Aut}\, K/F = \{\sigma \in \mathrm{Aut}\, K \mid a^\sigma = a$ for all $a \in F\}$ as the matrix $\sigma^{(r)}$ of the F-linear transformation $\vec{x} \mapsto \vec{x}^\sigma$ on $\vec{K}$ relative to the basis $\{\vec{k}_1, \ldots, \vec{k}_r\}$. We point out that the normalizer of K^* in $GL(r, F)$ is $K^* \rtimes \mathrm{Aut}\, K/F$. We shall also regard $\mathrm{Aut}\, K/F$ as a subgroup of $GL(nr, F)$, each $\sigma \in \mathrm{Aut}\, K/F$ sending each $v = \theta_1 e_1 + \cdots + \theta_n e_n \in V(n, K)$ (with all $\theta_i \in K$) to $\theta_1^\sigma e_1 + \cdots + \theta_n^\sigma e_n$, having the matrix $\sigma^{(nr)} = \mathrm{diag}(\sigma^{(r)}, \ldots, \sigma^{(r)})$. One can see that the normalizer of $N = SL(n, K)$ in $G = GL(nr, F)$ is $\Gamma = GL(n, K) \rtimes \mathrm{Aut}\, K/F$.

For each intermediate division ring E (i.e., a division ring E with $F \subseteq E \subseteq K$), we can take a pair of bases $\{w_1, \ldots, w_d\}$, $\{\varepsilon_1, \ldots, \varepsilon_h\}$ of K/E, E/F, resp., to construct a basis $\{\varepsilon_j w_i \mid 1 \leqslant i \leqslant d,\ 1 \leqslant j \leqslant h\} = \{\varepsilon_1 w_1, \ldots, \varepsilon_h w_1,\ \varepsilon_1 w_2, \ldots, \varepsilon_h w_2, \ldots, \varepsilon_1 w_d, \ldots, \varepsilon_1 w_d, \ldots, \varepsilon_h w_d\}$ of K/F. With respect to this basis we have $K \subset \mathrm{Mat}_d E$, namely, each $\theta \in K$ has the form

$\theta=(\alpha_{ij})_{d\times d}$ with each $\alpha_{ij}^{(h)}\in E$, where we regard $E\subset \mathrm{Mat}_h F$ by identifying each $\alpha\in E$ with the matrix $\alpha^{(h)}$ of the F-linear transformation $x\mapsto x\alpha$ on E relative to the basis $\{\varepsilon_1, ..., \varepsilon_h\}$ of E/F.

For each group X between $N=SL(n, K)$ and $G=GL(nr, F)$ we hope to find an intermediate division ring E, such that $SL(nd, E)\trianglelefteq X\leqslant GL(nd. E)\rtimes \mathrm{Aut}\, K/E$ (where $d=\dim_E K$). The following lemma will be useful in finding such E.

LEMMA 1. *Let* $0\neq A\in \mathrm{Mat}_r F$, $\mathrm{rank}\, A<r$, *and for all* $\theta\in K$ *assume either* $A\theta A=0$ *or* $\mathrm{rank}(A\theta A)=\mathrm{rank}\, A$, *then* Ker A and Im A *are subspaces of* K *over the division subring* E *of* K *generated by all* $\beta\beta_1^{-1}$, *with* $\vec{\beta}\in \mathrm{Im}\, A$, *for any given* $0\neq\vec{\beta}_1\in \mathrm{Im}\, A$. *We can choose suitable bases* $\{w_1, ..., w_d\}$ *and* $\{\varepsilon_1, ..., \varepsilon_h\}$ *of* K/E *and* E/F, resp., *to construct a basis* $\{\varepsilon_j w_i \mid 1\leqslant i\leqslant d, 1\leqslant j\leqslant h\}$ *of* K/F *to replace* $\{k_1, ..., k_r\}$, *to reduce* A *to the form* $\begin{pmatrix}0^{(r-h)} & \\ & \delta\end{pmatrix}$ *(when* $A^2\neq 0$*) or* $\begin{pmatrix} & 0^{(r-h)} \\ \delta & \end{pmatrix}$ *(when* $A^2=0$*) with* $\delta\in GL(h, F)$.

Proof. Denote $\vec{U}=\mathrm{Ker}\, A$, then $0<\dim_F U<r$. For each $\theta=\beta_1^{-1}u\in K$ with $u\in U$, we have $\vec{\beta}_1\theta=\vec{u}\in\vec{U}$ thus $\vec{\beta}_1\theta A=0$, which implies that the F-linear mapping $\varphi\colon \mathrm{Im}\, A\to\vec{K}$ defined by $\varphi(\vec{x})=\vec{x}\theta A$ has $\mathrm{Ker}\,\varphi\ni\vec{\beta}_1$, thus $\mathrm{Ker}\,\varphi\neq 0$, $\mathrm{rank}(A\theta A)=\dim_F(\mathrm{Im}\,\varphi)<\dim_F(\mathrm{Im}\, A)=\mathrm{rank}\, A$. By our assumption we must have $A\theta A=0$, $\mathrm{Im}\,\varphi=0$, $\overrightarrow{\beta\beta_1^{-1}u}=\vec{\beta}\theta\in \mathrm{Ker}\, A$ for all $\vec{\beta}\in\mathrm{Im}\, A$. This shows that $\beta\beta_1^{-1}U\subseteq U$ for all $\vec{\beta}\in\mathrm{Im}\, A$, $EU\subseteq U$ for the ring E generated by all $\beta\beta_1^{-1}$ ($\vec{\beta}\in\mathrm{Im}\, A$). Since $F\subseteq E\subseteq K$ and $\dim_F=r<\infty$ we know E is a division ring. $\vec{U}=\mathrm{Ker}\, A$ is a left E-subspace of K, $\dim_E U=(1/h)\dim_F U=(1/h)(r-\dim_F(\mathrm{Im}\, A))=d-(1/h)\dim_F(\mathrm{Im}\, A)$, where $h=\dim_F E$ and $d=\dim_E K$. But $0\neq\mathrm{Im}\, A\subseteq E\beta_1$, thus $0<(1/h)\dim_F(\mathrm{Im} A)\leqslant(1/h)\dim_F(E\beta_1)=(1/h)\cdot h=1$; since $(1/h)\dim_F(\mathrm{Im}\, A)=d-\dim_E U$ should be an integer we must have $(1/h)\dim_F(\mathrm{Im}\, A)=1$. $\mathrm{Im}\, A=\overrightarrow{E\beta_1}$ a 1-dimensional E-space, and $\dim_E(\mathrm{Ker}\, A)=d-1$. Take an E-basis $\{w_1, ..., w_d\}$ of K, with $Ew_1\oplus\cdots\oplus Ew_{d-1}=\mathrm{Ker}\, A$ and $\mathrm{Im}\, A=Ew_d$ (when $A^2\neq 0$; thus $\mathrm{Im}\, A\nsubseteq\mathrm{Ker}\, A$) or $\mathrm{Im}\, A=Ew_1$ (when $A^2=0$; thus $\mathrm{Im}\, A\subseteq\mathrm{Ker}\, A$), and take any F-basis $\{\varepsilon_1, ..., \varepsilon_h\}$ of E; then replacing $\{k_1, ..., k_r\}$ by $\{\varepsilon_j w_i \mid 1\leqslant i\leqslant d, 1\leqslant j\leqslant h\}$ we reduce A to the needed form.

We shall also need the following lemma.

LEMMA 2. *Let* T_0 *be a transvection of* $SL(n, K)$, $n\geqslant 3$, *and* $T_1=g_1T_0g_1^{-1}\in\Gamma_{Ke_1}$ *for a* $g_1\in GL(nr, F)$; *then* T_1 *is a transvection of* $SL(n, K)$.

Proof. $T_1\in\Gamma_{Ke_1}$, thus $T_1=(\theta_{ij}\sigma)_{n\times n}$ for a $\sigma^{(r)}\in\mathrm{Aut}\, K/F$ and a $(\theta_{ij}^{(r)})_{n\times n}\in GL(n, K)$ with all $\theta_{1j}=0$ $(2\leqslant j\leqslant n)$. Since T_1 is a conjugate of T_0 in $GL(nr, F)$, from $(T_0-I)^2=0$ and $\mathrm{rank}(T_0-I)=r$ we know $(T_1-I)^2=0$ and $\mathrm{rank}(T_1-I)=r$. If we can show $\sigma=1$, then $T_1\in GL(n, K)$; from $(T_1-I)^2=0$ and $\mathrm{rank}_K(T_1-I)=$

$(1/r)\,\mathrm{rank}(T_1 - I) = 1$ we know T_1 is a transvection of $SL(n, K)$. If $\theta_{ij} \neq 0$ for some $i \neq j$ and $j \geqslant 2$, we have $r = \mathrm{rank}(T_1 - I) \geqslant \mathrm{rank}(\theta_{11}\sigma - I) + \mathrm{rank}(\theta_{ij}\sigma) = \mathrm{rank}(\theta_{11}\sigma - I) + r$, thus $\mathrm{rank}(\theta_{11}\sigma - I) = 0$, $\theta_{11}\sigma = I$, $\sigma = 1$ as desired. Suppose $\theta_{ij} = 0$ for all $i \neq j$ and $j \geqslant 2$, then we have $r = \mathrm{rank}(T_1 - I) \geqslant \sum_{i=1}^{n} \mathrm{rank}(\theta_{ii}\sigma - I)$. We cannot have all $\mathrm{rank}(\theta_{ii}\sigma - I) \geqslant r/2$, otherwise $\sum_{i=1}^{n} \mathrm{rank}(\theta_{ii}\sigma - I) \geqslant nr/2 > r$, a contradiction. So we have $\mathrm{rank}(\theta\sigma - I) < r/2$ for $\theta = \theta_{ii} \in K^*$ for some i. The solution space $U = \{x \in K \mid \vec{x}(\theta\sigma - I) = \vec{0}\}$ has $\dim_F U = r - \mathrm{rank}(\theta\sigma - I) > r/2$, thus $U \cap Ub \neq 0$ for any $0 \neq b \in U$, $ab \in U$ for some non-zero $a, b \in U$. Note that for each $x \in U \subseteq K$ we have $\vec{x}(\theta\sigma - I) = \vec{0}$, $(x\theta)^{\sigma} = x$, $x\theta = x^{\sigma^{-1}}$, especially $ab\theta = (ab)^{\sigma^{-1}} = a^{\sigma^{-1}} b^{\sigma^{-1}} = a\theta b\theta$, $\theta = 1$, thus $U = \{x^{\sigma} = x \mid x \in K\}$ is a division subring of K with $\dim_U K = \dim_F K / \dim_F U < 2$, thus $\dim_U K = 1$, $U = K$, $\sigma = 1$, as desired.

1. Overgroups of $SL(n, K)$ in $GL(nr, F)$

We state and prove the following lemma and Lemma 5 in Section 2 in a general way so that we can also use them in some other papers.

LEMMA 3. *Let R be a ring with 1, D a division ring lying in R as a subring. Let $n \geqslant 3$, Γ be the normalizer of $GL(n, D)$ in $GL(n, R)$, $g_1 = (a_{ij})_{n \times n} \in GL(n, R) \backslash \Gamma$ with all $a_{ij} = 0$ $(2 \leqslant j \leqslant n)$, $Y = \langle SL(n, D), g_1 SL(n, D) g_1^{-1} \rangle$. Then Y contains $T_{ij}(c)$ for all $i \neq j$ and all c in a subring L of R containing D properly.*

Proof. Let $L = \{c \in R \mid T_{n1}(c) \in Y\}$. Considering the conjugates of all $T_{n1}(c) \in Y$ $(c \in L)$ by all $P_{ij} \in SL(n, D) < Y$ we see all $T_{ij}(c) \in Y$ $(i \neq j, c \in L)$.

For any $a, b \in L$, we have $T_{n1}(a \pm b) = T_{n1}(a)\, T_{n1}(b)^{\pm 1} \in Y$, thus $a \pm b \in L$. And since $n \geqslant 3$ we have $T_{n1}(ab) = [T_{n2}(a), T_{21}(b)] \in Y$, thus $ab \in L$. These entail that L is a subring of R.

$L \supseteq D$ trivially. We need to show that $L \neq D$, namely, to show the existence of some $T_{n1}(c) \in Y$ with $c \notin D$. Since $g_1 = (a_{ij})_{n \times n} \in GL(n, R) \backslash \Gamma$ with all $a_{1j} = 0$ $(j \geqslant 2)$, we have $g_1^{-1} = (\tilde{a}_{ij})_{n \times n} \in GL(n, R) \backslash \Gamma$ with $\tilde{a}_{11} = a_{11}^{-1}$ and all $\tilde{a}_{1j} = 0$ $(j \geqslant 2)$. For each $2 \leqslant l \leqslant n$ and $\theta \in D^*$, consider

$$g_2 = g_1 T_{l1}(\theta)\, g_1^{-1} = \begin{pmatrix} 1 & & & \\ b_2 & 1 & & \\ \vdots & & \ddots & \\ b_n & & & 1 \end{pmatrix} \in Y,$$

with $b_i = a_{il}\theta a_{11}^{-1}$ for $2 \leqslant i \leqslant n$. If we can choose l and θ to make some $b_k \notin D$, then replacing g_2 by $P_{2k}\, g_2 P_{2k}^{-1} \in Y$ (for $P_{2k} \in SL(n, D) < Y$) we may assume that $b_2 \notin D$, $T_{n2}(1)\, g_2 T_{n2}(-1)\, g_2^{-1} = T_{n1}(b_2) \in Y$ is just what we need. Suppose $b_i = a_{ij}\theta a_{11}^{-1} \in D$ for all chosen $i, j \geq 2$ and $\theta \in D$. Putting $\theta = 1$ we see $\theta_{ij} = a_{ij}a_{11}^{-1} \in D$ for all $i, j \geq 2$. Since g_1 is invertible we must have some $a_{k2} \neq 0$ $(k \geq 2)$, thus $\theta_{k2} = a_{k2}a_{11}^{-1} \in D^*$, $a_{11}\theta a_{11}^{-1} = \theta_{k2}^{-1}a_{k2}\theta a_{11}^{-1} \in D$ for all $\theta \in D$. Now $a_{11} SL(n, D) a_{11}^{-1} = SL(n, D)$, $g_1 SL(n, D) g_1^{-1} = (g_1 a_{11}^{-1}) SL(n, D)(g_1 a_{11}^{-1})^{-1}$, we can replace g_1 by $g_1 a_{11}^{-1}$ to reduce to the case $a_{11} = 1$ without changing Y. Now we have all $a_{ij} = a_{ij}a_{11}^{-1} \in D$ $(i, j \geq 2)$. Take $B = (a_{ij})_{2 \leq i,j \leq n} \in GL(n-1, D)$, and take $z = \begin{pmatrix} \lambda & 0 \\ 0 & B \end{pmatrix} \in SL(n, D)$ for a suitable $\lambda \in D^*$. Replacing g_1 by $z^{-1}g_1$, we reduce to the case

$$g_1 = \begin{pmatrix} \lambda^{-1} & & & \\ a_2 & 1 & & \\ \vdots & & \ddots & \\ a_n & & & 1 \end{pmatrix},$$

with some $a_k \notin D$ $(k \geq 2)$. Replacing such g_1 by $P_{2k}\, g_1 P_{2k}^{-1}$ we may assume $a_2 \notin D$, thus $T_{n1}(1)\, g_1 T_{n1}(-1)\, g_1^{-1} = T_{n1}(a_2\lambda) \in Y$, with $a_2\lambda \notin D$, as desired.

Proof of Theorem 1 (*for the case* $n \geq 3$). Let $SL(n, K) \leq X \leq GL(nr, F)$. Choose a minimal intermediate division ring E (between F and K) such that $X \geq SL(nd, E)$ (where $d = \dim_E K$). If $E = F$, the theorem holds trivially. Suppose $E \supsetneq F$; then we can replace K by E. Namely, we may assume that there is no $SL(nd, E) \leq X$ with $E \subsetneq K$. It suffices to prove $SL(n, K) \trianglelefteq X \leq \Gamma = GL(n, K) \rtimes \operatorname{Aut} K/F$. Suppose $X \nleq \Gamma$; we try to find an $SL(nd, E) \leq X$ with $E \subsetneq K$, thus obtaining a contradiction. To do this we try to find an E-transvection T (i.e., a transvection T of $SL(nd, E)$) in X, from which it may be seen that all the E-transvections lie in X, leading to $SL(nd, E) \leq X$ as desired.

Take a $g_1 = (A_{ij}^{(r)})_{n \times n} \in X_{\langle e_{11}, \ldots, e_{1k} \rangle} \backslash \Gamma$ with maximal $k \leq r$. We prove that $k = r$. Suppose $k \leq r-1$; we try to obtain a contradiction. Denote by u_{ij} the $((i-1)r + j)$th row of g_1 (i.e., the jth row of $(A_{i1} A_{i2} \cdots A_{in})$). We can take $z \in SL(n, K)$, sending u_{11} to e_{11}; thus $g_1 z \in X_{e_{11}} \backslash \Gamma$, which says that $k \geq 1$. Now we can take $z \in SL(n, K)$, fixing e_1 (thus fixing $u_{11}, \ldots, u_{1k}$ lying in Ke_1) when sending $u_{1,k+1}$ into $Ke_1 \oplus Ke_2$. In $g_1 z = (B_{ij}^{(r)})_{n \times n}$ the block B_{1n} has the first $k+1$ rows zero, thus $g_2 = (g_1 z)\, T_{n1}(I^{(r)})(g_1 z)^{-1} \in X_{\langle e_{11}, \ldots, e_{1,k+1} \rangle}$. By the maximality of k we must have $g_2 \in \Gamma$, $g_2 = (\theta_{ij}\sigma)_{n \times n}$ for a $\sigma^{(r)} \in \operatorname{Aut} K/F$ and a $(\theta_{ij})_{n \times n} \in GL(n, K)$. Since all the blocks $\theta_{1j}\sigma$ $(2 \leq j \leq n)$ have the first $k+1$ rows zero, thus are singular, we must have all $\theta_{1j} = 0$ $(2 \leq j \leq n)$, $g_2 \in \Gamma_{Ke_1}$. By Lemma 2 we know g_2 is a transvection of $SL(n, K)$, $z_1 g_2 z_1^{-1} = T_{n1}(I^{(r)})$ for a $z_1 \in SL(n, K)$, $T_{n1}(I^{(r)}) = z_1 g_2 z_1^{-1} = \tilde{g}_1 T_{n1}(I^{(r)})\, \tilde{g}_1^{-1}$ for $\tilde{g}_1 = z_1 g_1 z \in X \backslash \Gamma$, and we can see $\tilde{g}_1 \in X_{Ke_1}$ from $(T_{n1}(I^{(r)}) - I)\, \tilde{g}_1 = \tilde{g}_1(T_{n1}(I^{(r)}) - I)$, $k = r$ as desired.

So we have $g_1 = (A_{ij}^{(r)})_{n \times n} \in X_{Ke_1} \backslash \Gamma$. Applying Lemma 3 to the case

$R = \mathrm{Mat}_r F$ and $D = K$ we know that $T_{n1}(A^{(r)}) \in \langle N, g_1 N g_1^{-1} \rangle \leqslant X$ for all A in a ring $L \subseteq \mathrm{Mat}_r F$ with $L \supsetneq K$. For any $A \in L \backslash K$ we can take $\theta \in K$ having the same first row as A, thus $A - \theta \in L$ has the first row zero, $0 < \mathrm{rank}(A - \theta) < r$. Choose an $A \in L$ with smallest rank $A = h > 0$, then $h < r$. We can choose an $\alpha \in K^*$ sending a non-zero row $\vec{\beta}_1$ of A to $\vec{\beta}_1 \alpha \notin \mathrm{Ker}\, A$, thus $\vec{\beta}_1 \alpha A \neq \vec{0}$ $A\alpha A \neq 0$, $(A\alpha)^2 \neq 0$, and $\mathrm{rank}(A\alpha) = \mathrm{rank}\, A$; replacing A by such $A\alpha$ we may assume $A^2 \neq 0$. For all $\theta \in K$ we have $A\theta A \in L$ and $\mathrm{rank}(A\theta A) \leqslant \mathrm{rank}\, A$; by the minimality of rank A we have either $\mathrm{rank}(A\theta A) = \mathrm{rank}\, A$ or $A\theta A = 0$. By Lemma 1 we know that Ker A and Im A are subspaces of K over the division ring E generated by all $\beta\beta_1^{-1}$ with $\vec{\beta} \in \mathrm{Im}\, A$ (for a non-zero $\vec{\beta}_1 \in \mathrm{Im}\, A$), with $\dim_E(\mathrm{Ker}\, A) = \dim_E K - 1$ and $\dim_E(\mathrm{Im}\, A) = 1$. And we can take a pair of bases $\{w_1, ..., w_d\}$ and $\{\varepsilon_1, ..., \varepsilon_h\}$ of K/E and E/F, resp., with $Ew_1 \oplus \cdots \oplus Ew_{d-1} = \mathrm{Ker}\, A$ and $Ew_d = \mathrm{Im}\, A$, to construct a basis $(\varepsilon_j w_i \mid 1 \leqslant i \leqslant d,\ 1 \leqslant j \leqslant h\}$ of K/F to replace $\{k_1, ..., k_r\}$ and thus to reduce A to the form $(^0\ {}_\delta)$. We can choose an $\alpha^{(h)} \in E^*$ having the same first row as $\delta^{(h)}$, and can choose a $\theta^{(r)} = (\theta_{ij})_{d \times d} \in K \subset \mathrm{Mat}_d E$ with $\theta_{dd} = \alpha$ and all $\theta_{dj} = 0$ $(1 \leqslant j \leqslant d-1)$. Now $A^2 - A\theta = (^0\ {}_{\delta_1}) \in L$ with $\delta_1 = \delta(\delta - \alpha)$ and $\mathrm{rank}(A^2 - A\theta) = \mathrm{rank}(\delta - \alpha) < h$. By the minimality of rank $A = h$ we must have $\delta_1 = 0$, $\delta = \alpha \in E^*$. Now $T_{n1}(A) = T_{n1}((^0\ {}_\alpha)) \in X$ is a transvection in $SL(nd, E)$. For each $1 \leqslant p,\ q \leqslant d$ and $s \in E^*$ we can choose $\theta_1 = (\alpha_{ij})_{d \times d}$ and $\theta_2 = (\beta_{ij})_{d \times d}$ in $K \subset \mathrm{Mat}_d E$, with all $\alpha_{id} = \beta_{dj} = 0$ $(i \neq p,\ j \neq q)$, $\alpha_{p_d} = s$, and $\beta_{dq} = \alpha^{-1}$; thus $\theta_1 A \theta_2 = E_{pq}(s) \in L$, where we denote $E_{pq}(s) = sE_{pq} \in \mathrm{Mat}_d E$. X contains all the E-transvections $T_{kl}(E_{pq}(s))$ $(k \neq l)$, and it contains $[T_{kl}(E_{pq}(s)), T_{lk}(E_{qq}(I))] = \mathrm{diag}(D_1, ..., D_n)$ when $p \neq q$, with $D_k^{(r)} = T_{pq}(s^{(h)})$ and all other $D_i^{(r)} = I$. Now X contains enough E-transvections to generate $SL(nd, E)$, $X \geqslant SL(nd, E)$, but $E \subsetneq K$, contradicting our assumption, as desired.

Proof of Theorem 1 (*for the case* $n = 2$). The case $K = F_4$ (i.e., $GL(nr, F) = SL(4, 2)$) can be settled by considering the isomorphism $SL(4, 2) \cong A_8$; thus it will be excluded in the following discussion.

We still suppose X contains no $SL(2d, E)$ $(d = \dim_E K)$ with $E \subsetneq K$, and suppose $X \lneqq \Gamma = GL(2, K) \rtimes \mathrm{Aut}\, K/F$. When K is commutative we have $SL(2, K) = \mathrm{Sp}(2, K, f)$ for any nondegenerate alternating K-form f (we choose $f(e_1, e_2) = 1$), and we have $\mathrm{Sp}(2, K, f) \leqslant \mathrm{Sp}(2d, E, f_E)$ for each intermediate field E $(F \subseteq E \subseteq K,\ d = \dim_E K)$ and each alternating E-form $f_E = \varphi_E f$ with $0 \neq \varphi_E \in \mathrm{Hom}_E(K, E)$. Those X containing an $\mathrm{Sp}(2d, E, f_E)$ with $d \geqslant 2$ will be determined in the proof of Theorem 2 in Section 2 of this paper, which should normalize an $\mathrm{Sp}(2d_1, E_1, f_{E_1})$ or an $SL(2d_1, E_1)$ for a field E_1 between E and K, with $\dim_{E_1} K = d_1$. So we need only consider those X containing no $\mathrm{Sp}(2d, E, f_E)$ with $E \subsetneq K$. Take any $g_1 \in X \backslash \Gamma$, then g_1 does not stabilize the K-structure $V(2, K)$; namely, $(Ku) g_1 \neq K(ug_1)$ for some $u \neq 0$. Since $N = SL(2, K)$ is transitive among non-zero vectors in $V(2, K)$, we can find $z_1, z_2 \in N$ with $u = e_{11} z_1$ and $u g_1 z_2 = e_{11}$; replacing g_1 by $z_1 g_1 z_2 \in X \backslash \Gamma$, we may assume $e_{11} g_1 = e_{11}$ and $(K e_{11}) g_1 \neq K e_{11}$, namely, $g_1 = (A_{ij}^{(r)})_{2 \times 2} \in X_{e_{11}}$ with $A_{12} \neq 0$ and rank $A_{12} < r$ (since A_{12} has the first row zero). So we can choose $g_1 = (A_{ij}^{(r)})_{2 \times 2} \in X$ with smallest rank $A_{12} = h > 0$ and with $h < r$. Also, $g_1^{-1} = (\tilde{A}_{ij}^{(r)})_{2 \times 2} \in X$ has the block $\tilde{A}_{12} \neq 0$. Consider $g_2 = g_2(\theta) = g_1 \left(\begin{smallmatrix} I & 0 \\ \theta & I \end{smallmatrix}\right) g_1^{-1} = (B_{ij}^{(r)})_{2 \times 2} \in X$ with $\theta^{(r)} \in K$; we can choose $\theta = \theta_0$ to make $B_{12} = A_{12} \theta_0 \tilde{A}_{12} \neq 0$. Since $0 < \operatorname{rank} B_{12} \leqslant \operatorname{rank} A_{12} = h$ we know rank $B_{12} = h$ by the minimality of rank $A_{12} = h$, and we can replace g_1 by $g_2(\theta_2)$ to reduce to the case $g_1^{-1} = 2I - g_1$ (since $g_2(\theta_0)^{-1} = g_2(-\theta_0) = 2I - g_2(\theta_0)$), and especially $\tilde{A}_{12} = -A_{12}$. Now for such g_1 we have $B_{12} = -A_{12} \theta A_{12}$ in $g_2(\theta)$; by the minimality of h we must have either $\operatorname{rank}(A_{12} \theta A_{12}) = \operatorname{rank} A_{12}$ or $A_{12} \theta A_{12} = 0$, thus Lemma 1 applies, so we can reduce to the case $A_{12} = \Delta_1 = \left(\begin{smallmatrix} & 0^{(r-h)} \\ \delta & \end{smallmatrix}\right)$ or $A = \Delta_2 = \left(\begin{smallmatrix} 0^{(r-h)} & \\ & \delta \end{smallmatrix}\right)$ for a $\delta \in GL(h, F)$ by replacing the basis $\{k_1, \ldots, k_r\}$ of K/F by the $\{\varepsilon_j w_i \mid 1 \leqslant i \leqslant d,\ 1 \leqslant j \leqslant h\}$ obtained from a pair of bases $\{w_1, \ldots, w_d\}$, $\{\varepsilon_1, \ldots, \varepsilon_h\}$ of K/E, E/F, resp., for the division subring E generated by all $\beta \beta_1^{-1}$ with $\vec{\beta} \in \operatorname{Im} A$ (for a given $0 \neq \vec{\beta}_1 \in \operatorname{Im} A$). For each $\left(\begin{smallmatrix} I & 0 \\ C^{(r)} & I \end{smallmatrix}\right) \in X$ we have $g_2(C) = g_1 \left(\begin{smallmatrix} I & 0 \\ C & I \end{smallmatrix}\right) g_1^{-1} = \left(\begin{smallmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{smallmatrix}\right) = \left(\begin{smallmatrix} I + A_{12} C (2I - A_{11}) & -A_{12} C A_{12} \\ A_{22} C (2I - A_{11}) & I - A_{22} C A_{12} \end{smallmatrix}\right) \in X$. Specifically, we have $g_2(\theta) \in X$ for all $\theta \in K$. We can choose $\theta = \theta_0 \in K$ with $A_{12} \theta_0 A_{12} \neq 0$; replacing g_1 by $g_2(\theta_0)$ we reduce to the cases $A_{11} = \left(\begin{smallmatrix} I^{(r-h)} & 0 \\ * & * \end{smallmatrix}\right)$ and $A_{22} = \left(\begin{smallmatrix} * & 0 \\ * & I^{(r-h)} \end{smallmatrix}\right)$ (when $A_{12} = \Delta_1$) or $A_{22} = \left(\begin{smallmatrix} I^{(r-h)} & * \\ 0 & * \end{smallmatrix}\right)$ (when $A_{12} = \Delta_2$).

For each $\alpha \in E$ we shall denote

$$\lambda(\alpha) = w_1^{-1} \alpha w_1 = \begin{pmatrix} \alpha & 0 & \cdots & 0 \\ * & & \cdots & * \\ \vdots & & \cdots & \vdots \\ * & & \cdots & * \end{pmatrix} \in K,$$

$$\Lambda(\alpha) = w_d^{-1} \alpha w_d = \begin{pmatrix} * & & \cdots & * \\ \vdots & & \cdots & \vdots \\ * & & \cdots & * \\ 0 & \cdots & 0 & \alpha \end{pmatrix} \in K,$$

$$\xi(\alpha) \begin{pmatrix} * & \cdots & * & \alpha \\ & & & 0 \\ \vdots & \cdots & \vdots & \vdots \\ * & \cdots & * & 0 \end{pmatrix} \in K, \quad \text{and} \quad \eta(\alpha) = w_d^{-1} \alpha w_1 = \begin{pmatrix} * & & \cdots & * \\ \vdots & & \cdots & \vdots \\ * & & \cdots & * \\ \alpha & 0 & \cdots & 0 \end{pmatrix} \in K.$$

We prove that when $E \neq F_2$ we can always find an $\left(\begin{smallmatrix} I & B^{(r)} \\ 0 & I \end{smallmatrix}\right) \in X$ with $0 < \operatorname{rank} B \leqslant h$ (thus rank $B = h$, by the minimality of h) to replace g_1, to reduce to the case $g_1 = \left(\begin{smallmatrix} I & A_{12} \\ 0 & I \end{smallmatrix}\right)$. First we consider the case $A_{12} = \left(\begin{smallmatrix} & 0 \\ \delta & \end{smallmatrix}\right)$. Take a $g_2 = g_1 \left(\begin{smallmatrix} I & 0 \\ \lambda(\alpha) & I \end{smallmatrix}\right) g_1^{-1} = \left(\begin{smallmatrix} B_{11} & 0 \\ B_{21} & B_{22} \end{smallmatrix}\right) \in X$ for each

$$\lambda(\alpha) = \begin{pmatrix} \alpha & 0 & \cdots & 0 \\ * & \cdots & * & \lambda_2 \\ \vdots & \cdots & \vdots & \vdots \\ * & \cdots & * & \lambda_d \end{pmatrix} \in K,$$

with

$$B_{11}^{(r)} = \begin{pmatrix} I & & \\ & \ddots & \\ \delta\alpha & & I \end{pmatrix} \quad \text{and} \quad B_{22}^{(r)} = \begin{pmatrix} I & & & \\ -\lambda_2\delta & I & & \\ \vdots & & \ddots & \\ -\lambda_d\delta & & & I \end{pmatrix},$$

and take $g_3 = [\left(\begin{smallmatrix} I & 0 \\ I^{(r)} & I \end{smallmatrix}\right), g_2] = \left(\begin{smallmatrix} I & 0 \\ B & I \end{smallmatrix}\right) \in X$ with $B = B_{22}^{-1} B_{11} - I = B_{11} - B_{22}$ of rank $\leqslant h$. If we can choose α to make $B \neq 0$, then $\left(\begin{smallmatrix} I & B \\ 0 & I \end{smallmatrix}\right) = \left(\begin{smallmatrix} & I \\ I & \end{smallmatrix}\right)\left(\begin{smallmatrix} I & 0 \\ B & I \end{smallmatrix}\right)^{-1}\left(\begin{smallmatrix} 0 & -I \\ I & 0 \end{smallmatrix}\right) \in X$ can replace g_1, as desired. When char $F \neq 2$ we just take $\alpha = 1$, $\lambda(\alpha) = I$, $B = \left(\begin{smallmatrix} & 0 \\ 2\delta & \end{smallmatrix}\right) \neq 0$. Suppose char $F = 2$ and suppose we cannot choose $B \neq 0$, then $B = 0$ for all chosen $\alpha \in E$, thus

$$B_{22} = B_{11} = \begin{pmatrix} I & & \\ & \ddots & \\ \delta\alpha & & I \end{pmatrix},$$

all $\lambda_j = 0$ $(2 \leqslant j \leqslant d-1)$ and $\lambda_d = \delta\alpha\delta^{-1}$, $\delta\alpha\delta^{-1} \in E$ for all $\alpha \in E$. Now we take $[\left(\begin{smallmatrix} I & 0 \\ \xi(\beta) & I \end{smallmatrix}\right), g_2] = \left(\begin{smallmatrix} I & 0 \\ C & I \end{smallmatrix}\right) \in X$ for each

$$\xi(\beta) = \begin{pmatrix} \xi_1(\beta) & \cdots & \xi_{d-1}(\beta) & \beta \\ * & \cdots & * & 0 \\ \vdots & \cdots & \vdots & \vdots \\ * & \cdots & * & 0 \end{pmatrix} \in K \subset \operatorname{Mat}_d E,$$

(where we denote by $\xi_j(\beta)$ $(1 \leqslant j \leqslant d-1)$ the $(1, j)$th entry of $\xi(\beta) \in \operatorname{Mat}_d E$, which is a function of β), with

$$C = C(\alpha, \beta) = B_{22}^{-1}\xi(\beta)\, B_{11} - \xi(\beta) = \begin{pmatrix} \beta\delta\alpha & & \\ 0 & & \\ \vdots & & 0 \\ 0 & & \\ \delta\alpha\beta\,\delta\alpha + \delta\alpha\xi_1(\beta) & \cdots & \delta\alpha\beta \end{pmatrix}$$

dependent on $\alpha, \beta \in E$. For each $\beta, s \in E^*$ we choose $\alpha = \delta^{-1}\beta^{-1}\delta s$ (note that $\delta^{-1}\beta^{-1}\delta \in E^*$, thus $\delta^{-1}\beta^{-1}\delta s \in E^*$), thus $\beta\,\delta\alpha = \delta s$ is independent of the choice of β; we have $\begin{pmatrix} I & 0 \\ \Delta & I \end{pmatrix} = \begin{pmatrix} I & 0 \\ C(\delta^{-1}\beta^{-1}\delta s, \beta) & I \end{pmatrix}\begin{pmatrix} I & 0 \\ C(s, 1) & I \end{pmatrix} \in X$, with $\Delta = \Delta(s, \beta) = C(\delta^{-1}\beta^{-1}\delta s, \beta) + C(s, 1)$ dependent on s and β, and rank $\Delta \leqslant h$. If we can choose $\Delta \neq 0$, $\begin{pmatrix} I & \Delta \\ 0 & I \end{pmatrix} \in X$ can replace g_1. Suppose we cannot do this; then for all chosen s, β we have $\Delta = 0$, $C(\delta^{-1}\beta^{-1}\delta s, \beta) = C(s, 1)$. Specifically, $\delta \cdot \delta^{-1}\beta^{-1}\delta s \cdot \beta = \delta \cdot s \cdot 1$, $s\beta s^{-1} = \delta^{-1}\beta\delta$. This also holds for $s = 1$, thus $\delta^{-1}\beta\delta = s\beta s^{-1} = \beta$, E is a commutative field, and δ centralizes E and thus lies in E^*. From $C(\beta^{-1}s, \beta) = C(\delta^{-1}\beta^{-1}\delta s, \beta) = C(s, 1)$ we also obtain $\delta^2 s^2 \beta^{-1} + \delta s \beta^{-1}\xi_1(\beta) = \delta^2 s^2 + \delta s \xi_1(1)$, $\delta s(\beta^{-1} + 1) = \beta^{-1}\xi_1(\beta) + \xi_1(1) = \delta(\beta^{-1} + 1)$, but when $E \neq F_2$ we can choose $s \neq 1$ and $\beta \neq 1$, thus $\delta s(\beta^{-1} + 1) \neq \delta(\beta^{-1} + 1)$, a contradiction as desired. Now consider the case $A_{12} = \begin{pmatrix} 0 & \\ & \delta \end{pmatrix}$. We take $g_2 = g_1 \begin{pmatrix} I & 0 \\ \eta(\alpha) & I \end{pmatrix} g_1^{-1} = \begin{pmatrix} B_{11} & 0 \\ B_{21} & B_{22} \end{pmatrix} \in X$ for each

$$\eta(\alpha) = \begin{pmatrix} * & \cdots & * & \eta_1(\alpha) \\ \vdots & \cdots & \vdots & \vdots \\ * & \cdots & * & \eta_{d-1}(\alpha) \\ \alpha & 0 & \cdots & 0 \end{pmatrix} \in K \subset \mathrm{Mat}_d E,$$

with

$$B_{11} = \begin{pmatrix} I & & \\ & \ddots & \\ \delta\alpha & & I \end{pmatrix} \quad \text{and} \quad B_{22} = \begin{pmatrix} I & & -\eta_1(\alpha)\delta \\ & \ddots & \vdots \\ & & -\eta_{d-1}(\alpha)\delta \\ & & I \end{pmatrix};$$

then take $g_3 = \left[\begin{pmatrix} I & 0 \\ \eta(1) & I \end{pmatrix}, g_2\right] = \begin{pmatrix} I & 0 \\ B & I \end{pmatrix} \in X$ with

$$B = B_{22}^{-1}\eta(1)\, B_{11} - \eta(1) = \begin{pmatrix} \eta_1(\alpha)\,\delta + \eta_1(1)\,\delta\alpha & & \\ \vdots & & 0 \\ \eta_{d-1}(\alpha)\,\delta + \eta_{d-1}(1)\,\delta\alpha & & \\ 0 & \cdots & 0 \end{pmatrix}$$

of rank $\leqslant h$, and we need only to choose $B \neq 0$. Since $\eta(1) \in K^*$ is invertible we must have $\eta_l(1) \neq 0$ for an l. When char $F \neq 2$ we take $\alpha = 1$, thus

$$B = \begin{pmatrix} 2\eta_1(1) & & \\ \vdots & & 0 \\ 2\eta_{d-1}(1) & & \\ 0 & \cdots & 0 \end{pmatrix} \neq 0$$

(since $2\eta_l(1)\,\delta \neq 0$). Suppose char $F = 2$ and suppose $B = 0$, then all

$\eta_j(\alpha)\delta=\eta_j(1)\delta\alpha \quad (1\leqslant j\leqslant d-1)$ for all $\alpha\in E$, and we see $\delta\alpha\delta^{-1}=\eta_l(1)^{-1}\eta_l(\alpha)\in E$ for all $\alpha\in E$. Now we take $[\begin{pmatrix} I & 0\\ A(\beta) & I\end{pmatrix}, g_2]=\begin{pmatrix} I & 0\\ C & I\end{pmatrix}\in X$ for each

$$A(\beta)=\begin{pmatrix} * & \cdots & A_1(\beta)\\ \vdots & \cdots & \vdots\\ * & \cdots & A_{d-1}(\beta)\\ 0 & \cdots & \beta\end{pmatrix}\in K\subset \operatorname{Mat}_d E,$$

with

$$\begin{aligned} C=C(\alpha,\beta)&=B_{22}^{-1}A(\beta)B_{11}-A(\beta)\\ &=\begin{pmatrix} (A_1(\beta)+\eta_1(1)\,\delta\alpha\beta)\,\delta\alpha & & \eta_1(1)\,\delta\alpha\beta\\ \vdots & 0 & \vdots\\ (A_{d-1}(\beta)+\eta_{d-1}(1)\,\delta\alpha\beta)\,\delta\alpha & & \eta_{d-1}(1)\,\delta\alpha\beta\\ \beta\,\delta\alpha & 0 \quad \cdots & 0\end{pmatrix}. \end{aligned}$$

For each $\beta, s\in E^*$ we choose $\alpha=s\beta^{-1}$, thus all $\eta_j(1)\,\delta\alpha\beta=\eta_j(1)\,\delta s$ $(1\leqslant j\leqslant d-1)$; as we have $\begin{pmatrix} I & 0\\ \Delta & I\end{pmatrix}\in X$ for the $\Delta=\Delta(s,\beta)=C(s\beta^{-1},\beta)+C(s,1)$ dependent on s and β and rank $\Delta\leqslant h$, we need only to choose $\Delta\neq 0$. Suppose all $\Delta(s,\beta)=0$, $C(s\beta^{-1},\beta)=C(s,1)$, specifically $\beta\,\delta s\beta^{-1}=\delta s$, $\delta^{-1}\beta\delta=s\beta s^{-1}=1\beta 1^{-1}=\beta$, E is commutative, and δ centralizes E and thus lies in E^*. From $C(s\beta^{-1},\beta)=C(s,1)$ we also obtain $(A_l(\beta)+\eta_l(1)\,\delta s)\,\delta s\beta^{-1}=(A_l(1)+\eta_l(1)\,\delta s)\,\delta s$, thus $A_l(\beta)\,\beta^{-1}+A_l(1)=\eta_l(1)\,\delta s(\beta^{-1}+1)=\eta_l(1)\,\delta 1(\beta^{-1}+1)$, but when $E\neq F_2$ we can choose $\beta\neq 1$ and $s\neq 1$ to make $\eta_l(1)\,\delta s(\beta^{-1}+1)\neq\eta_l(1)\,\delta(\beta^{-1}+1)$, a contradiction.

So when $E\neq F_2$ we can always reduce to the case $g_1=\begin{pmatrix} I & A_{12}\\ 0 & I\end{pmatrix}$. In this case we also have $\begin{pmatrix} I & 0\\ A_{12} & I\end{pmatrix}\in X$, thus $\begin{pmatrix} \theta^{-1} & \\ & \theta\end{pmatrix}\begin{pmatrix} I & 0\\ A_{12} & I\end{pmatrix}\begin{pmatrix} \theta & \\ & \theta^{-1}\end{pmatrix}\begin{pmatrix} I & 0\\ -\theta_1 & I\end{pmatrix}=\begin{pmatrix} I & 0\\ C & I\end{pmatrix}\in X$ for all $\theta, \theta_1\in K^*$, with $C=\theta A_{12}\theta-\theta_1$, and $g_2=g_1\begin{pmatrix} I & 0\\ C & I\end{pmatrix}g_1^{-1}=\begin{pmatrix} B_{11} & B_{12}\\ B_{21} & B_{22}\end{pmatrix}\in X$ with $B_{12}=-A_{12}CA_{12}=-A_{12}(\theta A_{12}\theta-\theta_1)\,A_{12}$. We can choose an $\alpha^{(h)}\in E$ having the same first row as $\delta^{(h)}$, thus $0\leqslant \operatorname{rank}(\delta-\alpha)<h$. When $A_{12}=\begin{pmatrix} & 0\\ \delta & \end{pmatrix}$ we take

$$\theta=\xi(1)=\begin{pmatrix} * & \cdots & * & I^{(h)}\\ & & & 0\\ \vdots & \cdots & \vdots & \vdots\\ * & \cdots & * & 0\end{pmatrix}\in K^* \qquad\text{thus}\qquad \theta A_{12}\theta=\begin{pmatrix} * & \cdots & * & \delta\\ & & & 0\\ & & \cdot^{\cdot^{\cdot}} & \\ 0 & & & \end{pmatrix},$$

and we choose $\theta_1=\xi(\alpha)$, thus $B_{12}=\begin{pmatrix} & 0\\ \delta_1 & \end{pmatrix}$ with $\delta_1=-\delta(\delta-\alpha)\delta$, rank $B_{12}=\operatorname{rank}(\delta-\alpha)<h$. By the minimality of h we must have $B_{12}=0$, $\delta=\alpha\in E^*$. When $A_{12}=\begin{pmatrix} 0 & \\ & \delta\end{pmatrix}$ we take $\theta=I^{(r)}$ and choose any $\theta_1=(\alpha_{ij}^{(h)})_{d\times d}\in K^*$ with $\alpha_{dd}=\alpha$; then $B_{12}=\begin{pmatrix} 0 & \\ & \delta_1\end{pmatrix}$ with $\delta_1=-\delta(\delta-\alpha)\delta$,

rank $B_{12} = \operatorname{rank}(\delta - \alpha) < h$, leading to $\delta = \alpha \in E^*$ again. So in $g_1 = \left(\begin{smallmatrix} I & A_{12} \\ 0 & I \end{smallmatrix}\right)$ with $A_{12} = \left(\begin{smallmatrix} & 0 \\ \delta & \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix} 0 & \\ & \delta \end{smallmatrix}\right)$ we must have $\delta^{(h)} \in E^*$, g_1 a transvection of $SL(nd, E)$. When K is commutative g_1 is the symplectic transvection $\rho_{w_l e_2, 1}$: $x \mapsto x + f_E(x, w_l e_2)\, w_l e_2$ of the $\mathrm{Sp}(2d, E, f_E) > SL(2, K) = \mathrm{Sp}(2, K, f)$, relative to the alternating E-form $f_E = \varphi_E f$ with $\varphi_E \in \mathrm{Hom}_E(K, E)$ defined by $\varphi_E(w_l w_d) = \delta$ and all $\varphi_E(w_l w_i) = 0$ $(1 \leqslant i \leqslant d-1)$, where $l = 1$ when $A_{12} = \left(\begin{smallmatrix} & 0 \\ \delta & \end{smallmatrix}\right)$ or $l = d$ when $A_{12} = \left(\begin{smallmatrix} 0 & \\ & \delta \end{smallmatrix}\right)$. All the $g^{-1} \rho_{w_l e_2, 1}\, g = \rho_{w_l e_2 g, 1} \in X$ $(g \in \mathrm{Sp}(2, K, f))$ exhaust the conjugates of $\rho_{w_l e_2, 1}$ in $\mathrm{Sp}(2d, E, f_E)$ (since $w_l e_2\, g$ ranges over non-zero vectors); thus they generate $\mathrm{Sp}(2d, E, f_E)$, $X \geqslant \mathrm{Sp}(2d, E, f_E)$, a contradiction (since we assume X contains no $\mathrm{Sp}(2d, E, f_E)$ with $E \subsetneq K$). When K is non-commutative, for each $(\beta_1, ..., \beta_{d-1}) \in \mathrm{Mat}_{1 \times (d-1)} E$ we take a $\theta = (a_{ij})_{d \times d} \in K \subset \mathrm{Mat}_d E$, with $(\alpha_{11}, ..., \alpha_{1d}) = (\delta^{-1}\beta_1, ..., \delta^{-1}\beta_{d-1}, 0)$ when $A_{12} = \left(\begin{smallmatrix} & 0 \\ \delta & \end{smallmatrix}\right)$ or $(\alpha_{d1}, ..., \alpha_{dd}) = (\delta^{-1}\beta_1, ..., \delta^{-1}\beta_{d-1}, 0)$ when $A_{12} = \left(\begin{smallmatrix} 0 & \\ & \delta \end{smallmatrix}\right)$; then we have $g_2 = g_1 \left(\begin{smallmatrix} I & 0 \\ \theta & I \end{smallmatrix}\right) g_1^{-1} = \left(\begin{smallmatrix} B_{11} & 0 \\ B_{21} & B_{22} \end{smallmatrix}\right) \in X$ with

$$B_{11} = I + A_{12}\theta = \begin{pmatrix} I & & \\ & \ddots & \\ \beta_1 \cdots \beta_{d-1} & & I \end{pmatrix},$$

which says that for each E-transvection $T = I + v'u \in SL(d, E)$ with

$$v' = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ I \end{pmatrix} \in \mathrm{Mat}_{d \times 1} E$$

and $u = (\beta_1, ..., \beta_{d-1}, 0) \in \mathrm{Mat}_{1 \times d} E$ with $uv' = 0$ we have a $\left(\begin{smallmatrix} T & 0 \\ * & * \end{smallmatrix}\right) \in X$ and $\left(\begin{smallmatrix} \lambda & \\ & \lambda^{-1} \end{smallmatrix}\right)\left(\begin{smallmatrix} T & 0 \\ * & * \end{smallmatrix}\right)\left(\begin{smallmatrix} \lambda^{-1} & \\ & \lambda \end{smallmatrix}\right) = \left(\begin{smallmatrix} \lambda T \lambda^{-1} & 0 \\ * & * \end{smallmatrix}\right) \in X$ for all $\lambda \in K^*$, with $\lambda T \lambda^{-1} = I + (\lambda v')(u\lambda^{-1})$ ranging over all E-transvections in $SL(d, E)$ (since $\lambda v'$ ranges over non-zero columns in $\mathrm{Mat}_{d \times 1} E$). The group $\{A \in GL(d, E) \mid \text{ some } \left(\begin{smallmatrix} A & 0 \\ * & * \end{smallmatrix}\right) \in X\}$ contains all the E-transvections in $SL(d, E)$, thus contains $SL(d, E)$. Since K is non-commutative we can take a commutator γ of K^* not centralizing K^*, $\left(\begin{smallmatrix} \gamma & \\ & I \end{smallmatrix}\right) \in SL(2, K) < X$, thus the subgroup $\langle g \left(\begin{smallmatrix} \gamma & \\ & I \end{smallmatrix}\right) g^{-1} \mid g = \left(\begin{smallmatrix} A & 0 \\ * & * \end{smallmatrix}\right) \in X_{Ke_1} \rangle$ of X contains a $\left(\begin{smallmatrix} P & 0 \\ * & I \end{smallmatrix}\right)$ for each $P \in \langle A\gamma A^{-1} \mid A \in SL(d, E) \rangle = SL(d, E)$ and X contains $[\left(\begin{smallmatrix} I & 0 \\ I & I \end{smallmatrix}\right), \left(\begin{smallmatrix} P & 0 \\ * & I \end{smallmatrix}\right)] = \left(\begin{smallmatrix} I & 0 \\ P-I & I \end{smallmatrix}\right)$, thus X contains all $\left(\begin{smallmatrix} I & 0 \\ B & I \end{smallmatrix}\right)$, with B lying in the additive group generated by all $P - I$ with $P \in SL(d, E)$. These B can range over all $sE_{ij} \in \mathrm{Mat}_d E$ with $s \in E^*$ and $1 \leqslant i, j \leqslant d$; we can see $X \geqslant SL(2d, E)$ as in treating the case $n \geqslant 3$, a contradition as desired.

Now consider the remaining case, $E = F = F_2$. Excluding the settled case

$K=F_4$ we have $r=d\geqslant 3$ and $\delta=1$. When $g_1=\begin{pmatrix}A_{11}&A_{12}\\A_{21}&A_{22}\end{pmatrix}$ with $A_{12}=\begin{pmatrix}&0^{(r-1)}\\1&\end{pmatrix}$ we take $g_2=g_1\begin{pmatrix}I&0\\I&I\end{pmatrix}g_1^{-1}=\begin{pmatrix}B_{11}&0\\B_{21}&B_{22}\end{pmatrix}\in X$ with

$$B_{11}=B_{22}=\begin{pmatrix}1&&\\&\ddots&\\1&&1\end{pmatrix}.$$

$\begin{pmatrix}B_{11}&\\&B_{22}\end{pmatrix}$ is just the 2-transvection $t_{w_1e_1,w_1e_2}$ of the sumplectic group $\mathrm{Sp}(2r,F_2,f_0)$ relative to the $f_0=\varphi f$, with $\varphi\in\mathrm{Hom}(K,F_2)$ defined by $\varphi(w_rw_1)=1$ and all $\varphi(w_iw_1)=0$ $(1\leqslant i\leqslant r-1)$ (where we denote by $t_{u,w}$ the 2-transvection in $\mathrm{Sp}(2r,F_2,f_0)$ sendling each x to $x+f_0(x,w)u+f_0(x,u)w$ for any pair of vectors u,w with $f_0(u,w)=0$). For each $\theta\in K^*$ we have $[\begin{pmatrix}I&0\\\theta&I\end{pmatrix},g_2]=[\begin{pmatrix}I&0\\\theta&I\end{pmatrix},\begin{pmatrix}B_{11}&\\&B_{22}\end{pmatrix}]=t_{w_1e_1,w_1(e_2+\theta e_1)}t_{w_1e_1,w_1e_2}\in X$; since $r\geqslant 3$ we can choose $1\neq\theta_1\in K^*$ to make $f_0(w_1(e_2+\theta_1e_1),w_1e_2)=\varphi(\theta_1w_1^2)=0$, thus $t_{w_1e_1,w_1(e_2+\theta_1e_1)}t_{w_1e_1,w_1e_1}=t_{w_1e_1,\theta w_1e_1}\in X$ is a 2-transvection of $\mathrm{Sp}(2r,F_2,f_0)$. We can see that $\langle\mathrm{Sp}(2,K,f),t_{w_1e_1,\theta_1w_1e_1}\rangle=\mathrm{Sp}(2r,F_2,f_0)$, $X\geqslant\mathrm{Sp}(2r,F_2,f_0)$, as desired. Now consider the case in which g_1 has the block $A_{12}=\begin{pmatrix}0^{(r-1)}&\\&1\end{pmatrix}$; we take $g_2=g_1\begin{pmatrix}I&0\\\eta(1)&I\end{pmatrix}g_1^{-1}=\begin{pmatrix}B_{11}&0\\B_{21}&B_{22}\end{pmatrix}\in X$ for

$$\eta(1)=w_r^{-1}w_1=\begin{pmatrix}*&\cdots&*&a_1\\\vdots&\cdots&\vdots&\vdots\\ *&\cdots&*&a_{r-1}\\1&0&\cdots&0\end{pmatrix}\in K^*,$$

with

$$B_{11}=\begin{pmatrix}1&&\\&\ddots&\\1&&1\end{pmatrix}\quad\text{and}\quad B_{22}=\begin{pmatrix}1&&&a_1\\&\ddots&&\vdots\\&&1&a_{r-1}\\&&&1\end{pmatrix}.$$

Define $\varphi\in\mathrm{Hom}(K,F_2)$ by $\varphi(w_r^2)=1$ and all $\varphi(w_iw_r)=0$ $(1\leqslant i\leqslant r-1)$, and define $f_0=\varphi f$. Note that for each $1\leqslant i\leqslant r-1$ we have $w_iw_r^{-1}w_1=w_i\eta(1)=\sum_{j=1}^{r-1}\eta_{ij}w_j+a_iw_r$ (with all $\eta_{ij}\in F_2$), $w_iw_1=\sum_{j=1}^{r-1}\eta_{ij}w_jw_r+a_iw_r^2$, thus $\varphi(w_iw_1)=a_i$, so we can see $\begin{pmatrix}B_{11}&\\&B_{22}\end{pmatrix}=t_{w_1e_1,w_re_2}\in\mathrm{Sp}(2r,F_2,f_0)$, $[\begin{pmatrix}I&0\\I&I\end{pmatrix},g_2]=[\begin{pmatrix}I&0\\I&I\end{pmatrix},\begin{pmatrix}B_{11}&\\&B_{22}\end{pmatrix}]=\begin{pmatrix}I&0\\C&I\end{pmatrix}\in X\cap\mathrm{Sp}(2r,F_2,f_0)$ with

$$C=B_{22}^{-1}B_{11}-I=\begin{pmatrix}a_1&&a_1\\\vdots&0&\vdots\\a_{r-1}&&a_{r-1}\\1&0\ \cdots&0\end{pmatrix};$$

such an $\left(\begin{smallmatrix} I & 0 \\ C & I \end{smallmatrix}\right)$ is a product $\rho_{(w_1+w_r)e_1}\rho_{w_re_1}$ of two symplectic transvections $\rho_{(w_1+w_r)e_1}$ and $\rho_{w_re_1}$ in $\mathrm{Sp}(2r, F_2, f_0)$ (where we denote $\rho_u: x \mapsto x + f_0(x, u)u$ for each $u \in V$). Considering the conjugates under $N = \mathrm{Sp}(2, K, f)$ we know that X contains all $\rho_{\alpha u}\rho_u$ with $u \neq 0$ and $\alpha = w_1 w_r^{-1} + 1 \in K^* \setminus \{1\}$. If $\alpha^2 \in F_2 \oplus F_2\alpha$, thus $\alpha^2 = 1 + \alpha$, then we have $F_2[\alpha] = F_4$, thus $2 \mid r$. $W = \{\theta \in K \mid \varphi(\theta) = \varphi(\alpha\theta) = 0\}$ is an $F_2[\alpha]$-space, we can write $K = F_2[\alpha]\beta \oplus W$ for a $\beta \in K^*$ with $\varphi(\beta) = 1$, and we can define $\varphi_1 \in \mathrm{Hom}_{F_2[\alpha]}(K, F_2[\alpha])$ by $\varphi_1(e\beta + w) = e$ for each $e \in F_2[\alpha]$ and $w \in W$ and define $\varphi_2 \in \mathrm{Hom}(F_2[\alpha], F_2)$ by $\varphi_2(e) = \varphi(e\beta)$, thus $\varphi = \varphi_2\varphi_1$. Now all the $\rho_{\alpha u}\rho_u \in X$ act as symplectic transvections of $\mathrm{Sp}(r, F_4, f_1)$ relative to $f_1 = \varphi_1 f$; all such symplectic transvections generate $\mathrm{Sp}(r, F_4, f_1) \leqslant X$, a contradiction. Now suppose $\alpha^2 \notin F_2 \oplus F_2\alpha$, then we can take a $\lambda \in K^*$ with $\varphi(\lambda) = \varphi(\lambda\alpha) = 0 \neq \varphi(\lambda\alpha^2)$ and take $T = \rho_{\alpha\lambda e_2}\rho_{\lambda e_2} \in X$, thus $T^{-1}(\rho_{\alpha e_1}\rho_{e_1})T = \rho_{\alpha e_1 + \alpha\lambda_2}\rho_{e_1} \in X$. Considering the conjugates under N we know $\rho_{\alpha\lambda e_2}\rho_{e_1 - \theta e_2} \in X$ for all $\theta \in K$, thus $(\rho_{\alpha\lambda e_2}\rho_{e_1})^{-1}\rho_{\alpha\lambda e_2}\rho_{e_1 - \theta e_2} = \rho_{e_1}\rho_{e_1 + \theta e_2} \in X$, and $\rho_u\rho_v \in X$ for all u, v with $f(u, v) \neq 0$. For non-zero u, w with $f(u, w) = 0$ we can choose v with both $f(u, v)$ and $f(u, w)$ non-zero, thus $(\rho_u\rho_v)(\rho_v\rho_w) = \rho_u\rho_w \in X$. X contains all the conjugates of $\rho_{e_1}\rho_{e_2}$ in $\mathrm{Sp}(2r, F_2, f_0)$, thus contains the whole $\mathrm{Sp}(2r, F_2, f_0)$, a contradiction again.

2. Overgroups of $\mathrm{Sp}(2v, K)$ in $GL(2vr, F)$

Lemma 4. *Let X be an overgoup of any symplectic group* $\mathrm{Sp}(2m, E)$ *in* $\Gamma L(2m, E)$; *then we have* $X \trianglerighteq \mathrm{Sp}(2m, E)$ *or* $X \trianglerighteq SL(2m, E)$.

Proof. Each transvection in $SL(2m, E)$ has the form $\tau_{u,v}: x \mapsto x + (x, v)u$, associated with a pair of non-zero vectors u, v with $(u, v) = 0$ (where (x, y) denotes the alternating inner product of any pair of vectors x, y in the underlying space of $\mathrm{Sp}(2m, E)$). We see that $\tau_{u,v} \in \mathrm{Sp}(2m, E)$ if and only if u and v are collinear. If $X \ntrianglerighteq \mathrm{Sp}(2m, E)$, then X contains a $\tau_{u_1, v_1} \notin \mathrm{Sp}(2m, E)$, with u_1, v_1 non-collinear, and X contains all $g^{-1}\tau_{u_1, v_1} g = \tau_{u_1 g, v_1 g}$ with $g \in \mathrm{Sp}(2m, E)$. Since $\{u_1 g, v_1 g\}$ $(g \in \mathrm{Sp}(2m, E))$ ranges over all the non-collinear and orthogonal pairs of vectors, we know $\tau_{u_1 g, v_1 g}$ ranges over all transvections in $SL(2m, E)$ not lying in $\mathrm{Sp}(2m, E)$; X contains all the transvections of $SL(2m, E)$, thus contains the whole of $SL(2m, E)$.

In Lemma 5 we shall write $A_{(k)}$ to suggest that $A \in \mathrm{Mat}_k R$.

Lemma 5. *Let R be a ring with* 1, *D a field lying in R as a subring. Let $n = 2v \geqslant 4$, Γ be the normalizer of $GL(n, D)$ in $GL(n, R)$, and $g_1 = (a_{ij})_{n \times n} \in GL(n, R) \setminus \Gamma$ with all $a_{1j} = 0$ $(j \geqslant 2)$. Let* $\mathrm{Sp}(n, D) =$

$\{A \in GL(n, D) \mid AHA' = H\}$ *for* $H = \begin{pmatrix} 0 & I_{(v)} \\ I_{(v)} & 0 \end{pmatrix} \in GL(n, D)$, $Y = \langle \mathrm{Sp}(n, D), g_1 \mathrm{Sp}(n, D) g_1^{-1} \rangle$. *Then*

(1) *Y contains a*

$$T = \left(\begin{array}{c|c} I_{(v)} & 0 \\ \hline \begin{matrix} S & \\ & 0 \end{matrix} & I_{(v)} \end{array}\right)$$

with $S = \begin{pmatrix} a & b \\ c & 0 \end{pmatrix} \in (\mathrm{Mat}_2 R) \setminus \mathrm{Mat}_2 D)$;

(2) *for each* $S = \begin{pmatrix} a & b \\ c & 0 \end{pmatrix} \in \mathrm{Mat}_2 R$ *with*

$$\left(\begin{array}{c|c} I_{(v)} & 0 \\ \hline \begin{matrix} S & \\ & 0 \end{matrix} & I_{(v)} \end{array}\right) \in Y,$$

Y contains all $T_{v+1,1}(\theta c + b\theta)$ $(\theta \in D)$ *and Y contains* $T_{v+1,1}(a)$ *and all* $T_{v+1,1}(b\theta c)$ $(\theta \in D)$ *when* $D \neq F_2$ *or* $v \geqslant 3$;

(3) *when* $D \neq F_2$ *or* $v \geqslant 3$ *we have* $T_{v+1,1}(aa_1a) \in Y$ *for any* $T_{v+1,1}(a), T_{v+1,1}(a_1) \in Y$.

Proof. (1) We point out that we can replace g_1 by any gg_1z with $g \in Y$ and $z\,\mathrm{Sp}(n, D)\,z^{-1} = \mathrm{Sp}(n, D)$, without changing Y. We can also replace g_1 by a $g \in Y$, thus replace Y by a $Y_1 \leqslant Y$, provided that we can find a needed T in Y_1.

For each $\theta \in D^*$, consider

$$g_2 = g_1 T_{v+1,1}(\theta) g_1^{-1} = \begin{pmatrix} 1 & & & \\ b_2 & 1 & & \\ \vdots & & \ddots & \\ b_n & & & 1 \end{pmatrix} \in Y$$

with $b_i = a_{i,v+1}\theta a_{11}^{-1}$ $(i \geqslant 2)$. Consider $[T_{i\pm v,i}(1), g_2] = T_{i\pm v,1}(b_i) \in Y$ for each $i \notin \{1, v+1\}$ (where

$$i \pm v = \begin{cases} i + v & \text{when} \quad i \leqslant v \\ i - v & \text{when} \quad i > v \end{cases},$$

thus $1 \leqslant i \pm v \leqslant 2v$). If we can choose $\theta \in D^*$ to make $b_i \notin D$ for an $i \notin \{1, v+1\}$, then $[T_{v+1,i+v}(1)\, T_{i1}(\mp 1),\ T_{i\pm v,1}(b_i)] = T_{v+1,1}(b_i) \in Y$ is just the needed T. Suppose $b_i = a_{i,v+1}\theta a_{11}^{-1} \in D$ for all $\theta \in D$ and all

$i\notin\{1,\nu+1\}$. If for an $i\notin\{1,\nu+1\}$ we have $a_{i,\nu+1}\neq 0$, thus $b_i\in D^*$, $T_{i\pm\nu,1}(b_i)\in SL(n,D)\setminus \mathrm{Sp}(n,D)$, $Y\geqslant\langle \mathrm{Sp}(n,D),\ T_{i\pm\nu,1}(b_i)\rangle = SL(n,D)$ by Lemma 4, we can replace g_1 by zg_1 to annihilate all $a_{i,\nu+1}$ $(i\neq\nu+1)$ for a suitable $z=(\alpha_{ij})_{n\times n}\in SL(n,D)<Y$ with all $\alpha_{1j}=0$ $(j\geqslant 2)$. So we may begin by assuming all $a_{i,\nu+1}=0$ $(i\neq \nu+1)$. Since g_1 is invertible we have $a_{\nu+1,\nu+1}\neq0$ and $g_2=T_{\nu+1,1}(b_{\nu+1})\in Y$ with $b_{\nu+1}=a_{\nu+1,\nu+1}\theta a_{11}^{-1}\neq 0$ for all $\theta\in D^*$. If we can choose θ to make $b_{\nu+1}\notin D$, $T=T_{\nu+1,1}(b_{\nu+1})$ is just what we need. Suppose $b_{\nu+1}\in D^*$ for all chosen $\theta\in D^*$. Specifically, we have $a_{\nu+1,\nu+1}a_{11}^{-1}\in D^*$, thus $a_{11}\theta a_{11}^{-1}=(a_{\nu+1,\nu+1}a_{11}^{-1})^{-1}a_{\nu+1,\nu+1}\theta a_{11}^{-1}\in D^*$ for all $\theta\in D^*$. Now, for each $A\in \mathrm{Sp}(n,D)$ we have $a_{11}Aa_{11}^{-1}\in GL(n,D)$ and $(a_{11}Aa_{11}^{-1})\,H(a_{11}Aa_{11}^{-1})'=a_{11}A(a_{11}^{-1}Ha_{11})\,A'a_{11}^{-1}=H$, thus $A\in \mathrm{Sp}(n,D)$. This shows that $a_{11}\mathrm{Sp}(n,D)a_{11}^{-1}=\mathrm{Sp}(n,D)$, $g_1\,\mathrm{Sp}(n,D)\,g_1^{-1}=(g_1a_{11}^{-1})\,\mathrm{Sp}(n,D)(g_1a_{11}^{-1})^{-1}$, we can replace g_1 by $g_1a_{11}^{-1}$ to reduce to the case $a_{11}=1$, and $a_{\nu+1,\nu+1}\in D^*$. Since $g_1\notin\Gamma$ we have some $a_{ij}\notin D$. If $a_{ij}\in D$ for all $i\neq\nu+1$, we have some $a_{\nu+1,j}\notin D$; since $(a_{11},\dots,a_{1n})=(1,0,\dots,0)$ and $a_{2,\nu+1}=0$ we can take a $z\in \mathrm{Sp}(n,D)$ having the same first two rows as g_1, and can replace g_1 by g_1z^{-1} to reduce to the case $(a_{21},\dots,a_{2n})=(0,1,0,\dots,0)$. Still we have some $a_{\nu+1,j}\notin D$ and can replace g_1 by $P_{12}P_{\nu+1,\nu+2}\,g_1(P_{12}P_{\nu+1,\nu+2})^{-1}$ to reduce to the case in which some $a_{\nu+2,j}\notin D$. So we may start by assuming that there exists an $a_{kl}\notin D$ with $k\neq\nu+1$ (and, of course, $k\neq1$). When $k\geqslant\nu+2$ we replace g_1 by $P_{k-\nu,k}\,g_1$ to reduce to the case $2\leqslant k\leqslant\nu$. Suppose $k\leqslant\nu$; then we replace g_1 by $(P_{2k}P_{\nu+2,\nu+k})\,g_1$ to reduce to the case $k=2$. So we suppose some $a_{2l}\notin D$. If $a_{2l}\notin D$ for an $l\neq1$ (of course $l\neq\nu+1$, since $a_{2,\nu+1}=0$), we replace g_1 by $g_1T_{l1}(1)\,T_{\nu+1,l\pm\nu}(\mp1)\,g_1^{-1}=(b_{ij})_{n\times n}\in Y$ with $b_{11}=b_{22}=1$, $b_{21}=a_{2l}\notin D$, and $b_{1j_1}=b_{2j_2}=0$ for all $j_1\geqslant2$ and $j_2\geqslant3$. Suppose all $a_{2j}\in D$ $(j\geqslant2)$ but $a_{21}\notin D$; then we can choose a $z=(\theta_{ij})_{n\times n}\in \mathrm{Sp}(n,D)$ with $(\theta_{11},\dots,\theta_{1n})=(1,0,\dots,0)$ and $(\theta_{21},\dots,\theta_{2n})=(0,a_{22},\dots,a_{2n})$, and can replace g_1 by g_1z^{-1} to reduce to the case $(a_{21},\dots,a_{2n})=(a_{21},1,0,\dots,0)$ with $a_{21}\notin D$. Anyway, we may reduce to the case in which $(a_{21},\dots,a_{2n})=(a_{21},1,0,\dots,0)$ with $a_{21}\notin D$. Now we take

$$g_2=g_1T_{\nu+2,2}(1)\,g_1^{-1}=(b_{ij})_{n\times n}=\begin{pmatrix}1&&&&\\0&1&&&\\b_{31}&b_{32}&1&&\\\vdots&\vdots&&\ddots&\\b_{n1}&b_{n2}&&&1\end{pmatrix}\in Y,$$

with $b_{i2}=a_{i,\nu+2}$ and $b_{i1}=-a_{i,\nu+2}a_{21}$ for $3\leqslant i\leqslant n$. Since g_1 is invertible we have some $a_{k,\nu+2}\neq0$ $(k\geqslant3)$. For each such k, if $b_{k2}=a_{k,\nu+2}\in D^*$ we have $b_{k1}=-a_{k,\nu+2}a_{21}\notin D$ (since $a_{21}\notin D$). Namely, we have either $b_{k2}\notin D$ or $b_{k1}\notin D$, $g_2\notin\Gamma$ in any case. If $b_{k2}\neq0$ for a $k\notin\{\nu+1,\nu+2\}$, we have

$$\tilde{g}_2=[g_2, T_{21}(1)\,T_{\nu+1,\nu+2}(-1)]$$

$$=\begin{bmatrix} 1 & & & & & & \\ 0 & 1 & & & & & \\ b_{32} & & 1 & & & & \\ \vdots & & & \ddots & & & \\ b_{\nu+1,2}+b_{\nu+2,1} & b_{\nu+2,2} & & & 1 & & \\ \vdots & & & & & \ddots & \\ b_{n2} & & & & & & 1 \end{bmatrix}\in Y,$$

thus $[T_{k\pm\nu,k}(1), \tilde{g}_2]=T_{k\pm\nu,1}(b_{k2})\in Y$. When $b_{k2}\in D^*$ we have $Y\geqslant\langle \mathrm{Sp}(n, D), T_{k\pm\nu,1}(b_{k2})\rangle=SL(n, D)$ by Lemma 4. Applying Lemma 3 we know $Y\geqslant\langle SL(n, D), g_2 SL(n, D)\, g_2^{-1}\rangle\ni T_{\nu+1,1}(c)$ for a $c\in R\backslash D$, as desired. When $b_{k2}\notin D$, take $T=[T_{\nu+1,k\mp\nu}(1)\,T_{k1}(\mp 1), T_{k\pm\nu,1}(b_{k2})]=T_{\nu+1,1}(b_{k2})\in Y$, still we are done. Now we suppose $b_{i2}=0=b_{i1}$ for all $i\notin(\nu+1, \nu+2)$,

$$g_2=\left(\begin{array}{c|c} I_{(\nu)} & 0 \\ \hline \begin{matrix} C \\ 0 \end{matrix} & I_{(\nu)} \end{array}\right)\in Y$$

with $C=\begin{pmatrix} b_{\nu+1,1} & b_{\nu+1,2} \\ b_{\nu+2,1} & b_{\nu+2,2}\end{pmatrix}\notin \mathrm{Mat}_2 D$. If $b_{\nu+2,2}\in D$ we have $T_{\nu+2,2}(-b_{\nu+2,2})\in \mathrm{Sp}(n, D)$, thus can take $T=g_2 T_{\nu+1,1}(-b_{\nu+2,2})\in Y$. When $b_{\nu+2,2}\notin D$ we have

$$T=[T_{21}(1)\,T_{\nu+1,\nu+2}(-1), g_2^{-1}]=\left(\begin{array}{c|c} I_{(\nu)} & 0 \\ \hline \begin{matrix} S \\ 0 \end{matrix} & I_{(\nu)} \end{array}\right)\in Y$$

with $S=\begin{pmatrix}1&1\\0&1\end{pmatrix} C \begin{pmatrix}1&0\\1&1\end{pmatrix}-C=\begin{pmatrix} * & b_{\nu+2,2} \\ b_{\nu+2,2} & 0\end{pmatrix}\notin \mathrm{Mat}_2 D$, as required.

(2) Denote

$$L=\left\{ S\in \mathrm{Mat}_\nu R \;\middle|\; \left(\begin{array}{c|c} I_{(\nu)} & 0 \\ \hline S & I_{(\nu)} \end{array}\right)\in Y\right\}$$

and

$$L_2=\left\{ S\in \mathrm{Mat}_2 R \;\middle|\; \begin{pmatrix} S & \\ & O_{(\nu-2)}\end{pmatrix}\in L\right\}.$$

One can see that L and L_2 are additive groups. For each

$$g=\left(\begin{array}{c|c} A_{(v)} & 0 \\ \hline * & B_{(v)} \end{array}\right)\in Y \qquad \text{and} \qquad S\in L$$

we have $g\left(\begin{smallmatrix} I & 0 \\ S & I \end{smallmatrix}\right) g^{-1}=\left(\begin{smallmatrix} I & 0 \\ S_1 & I \end{smallmatrix}\right)\in Y$ with $S_1=BSA^{-1}$, thus $BSA^{-1}\in L$. In particular, for each $A\in GL(v, D)$ we have $\operatorname{diag}(A'^{-1}, A)\in \operatorname{Sp}(n, D)$; thus $ASA'\in L$ for all $S\in L$. And we have $ASA'\in L_2$ for all $A\in GL(2, D)$ and $S\in L_2$. Now, for each $S=\left(\begin{smallmatrix} a & b \\ c & 0 \end{smallmatrix}\right)\in L_2$ and $\theta\in D$, we have $\left(\begin{smallmatrix} 1 & \theta \\ 0 & 1 \end{smallmatrix}\right) S\left(\begin{smallmatrix} 1 & 0 \\ \theta & 1 \end{smallmatrix}\right)-S=\left(\begin{smallmatrix} \theta c+b\theta & \\ & 0 \end{smallmatrix}\right)\in L_2$, i.e., $T_{v+1,1}(\theta c+b\theta)\in Y$. When $D\neq F_2$ we can take $1\neq\theta\in D^*$, thus $\left(\begin{smallmatrix} a & b \\ c & 0 \end{smallmatrix}\right)\in L_2$ implies $\left(\begin{smallmatrix} 1 & \\ & (\theta-1)^{-1} \end{smallmatrix}\right)\left(\left(\begin{smallmatrix} 1 & 0 \\ 0 & \theta \end{smallmatrix}\right)\left(\begin{smallmatrix} a & b \\ c & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 0 \\ 0 & \theta \end{smallmatrix}\right) - \left(\begin{smallmatrix} a & b \\ c & 0 \end{smallmatrix}\right)\right)\left(\begin{smallmatrix} 1 & \\ & (\theta-1)^{-1} \end{smallmatrix}\right)=\left(\begin{smallmatrix} 0 & b \\ c & 0 \end{smallmatrix}\right)\in L_2$ and $\left(\begin{smallmatrix} a & \\ & 0 \end{smallmatrix}\right)=\left(\begin{smallmatrix} a & b \\ c & 0 \end{smallmatrix}\right)-\left(\begin{smallmatrix} 0 & b \\ c & 0 \end{smallmatrix}\right)\in L_2$, i.e., $T_{v+1,1}(a)\in Y$. When $v\geqslant 3$, $\left(\begin{smallmatrix} a & b \\ c & 0 \end{smallmatrix}\right)\in L_2$ implies

$$\begin{pmatrix} 1 & & & \\ & 0 & 1 & \\ & 1 & 0 & \\ & & & I \end{pmatrix}\left(\begin{pmatrix} 1 & & & \\ & 1 & 0 & \\ & 1 & 1 & \\ & & & I \end{pmatrix}\begin{pmatrix} a & b & \\ c & 0 & \\ & & 0 \end{pmatrix}\right.$$

$$\left.\times\begin{pmatrix} 1 & & & \\ & 1 & 1 & \\ & 0 & 1 & \\ & & & I \end{pmatrix}-\begin{pmatrix} a & b & \\ c & 0 & \\ & & 0 \end{pmatrix}\right)\begin{pmatrix} 1 & & & \\ & 0 & 1 & \\ & 1 & 0 & \\ & & & I \end{pmatrix}=\begin{pmatrix} 0 & b & \\ c & 0 & \\ & & 0 \end{pmatrix}\in L;$$

thus $\left(\begin{smallmatrix} 0 & b \\ c & 0 \end{smallmatrix}\right)\in L_2$ and $\left(\begin{smallmatrix} a & \\ & 0 \end{smallmatrix}\right)\in L_2$ again. Now, $\left(\begin{smallmatrix} 0 & c \\ b & 0 \end{smallmatrix}\right)\in L_2$ implies

$$P_{2,v+2}^{-1}\left(\begin{array}{ccc|c} & I_{(v)} & & \\ \hline 0 & b & & \\ c & 0 & & I_{(v)} \\ & & 0 & \end{array}\right)P_{2,v+2}$$

$$=\operatorname{diag}\left(\begin{pmatrix} 1 & & \\ -c & 1 & \\ & & I_{(v-2)} \end{pmatrix}, \begin{pmatrix} 1 & b & \\ & 1 & \\ & & I_{(v-2)} \end{pmatrix}\right);$$

thus $\left(\begin{smallmatrix} 1 & b \\ 0 & 1 \end{smallmatrix}\right) S\left(\begin{smallmatrix} 1 & 0 \\ c & 1 \end{smallmatrix}\right)\in L_2$ for all $S\in L_2$. In particular, for each $\theta\in D^*$ we have $\left(\begin{smallmatrix} 1 & b \\ 0 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & \\ & \theta \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 0 \\ c & 1 \end{smallmatrix}\right) - \left(\begin{smallmatrix} 0 & \\ & \theta \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & b \\ c & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & \\ & \theta \end{smallmatrix}\right) = \left(\begin{smallmatrix} b\theta c & \\ & 0 \end{smallmatrix}\right)\in L_2$, $T_{v+1,1}(b\theta c)\in Y$.

(3) $T_{v+1,1}(a)\in Y$ implies $\left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} a & \\ & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)-\left(\begin{smallmatrix} a & \\ & 0 \end{smallmatrix}\right)-\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} a & \\ & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)=\left(\begin{smallmatrix} 0 & a \\ a & 0 \end{smallmatrix}\right)\in L_2$; thus $\left(\begin{smallmatrix} 1 & a \\ 0 & 1 \end{smallmatrix}\right) S\left(\begin{smallmatrix} 1 & 0 \\ a & 1 \end{smallmatrix}\right)\in L_2$ for all $S\in L_2$ (by the proof of (2)).

$\begin{pmatrix} a_1 & \\ & 0 \end{pmatrix} \in L_2$ leads to $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a_1 & \\ & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \\ & a_1 \end{pmatrix} \in L_2$, thus $\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & \\ & a_1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ a & 1 \end{pmatrix} = \begin{pmatrix} aa_1a & aa_1 \\ a_1a & 0 \end{pmatrix} \in L_2$, leading to $\begin{pmatrix} aa_1a & \\ & 0 \end{pmatrix} \in L_2$ (i.e., $T_{v+1,1}(aa_1a) \in Y$) when $D \neq F_2$ or $v \geqslant 3$.

Proof of Theorem 2. Let $N_1 = \mathrm{Sp}(n, K, f) \leqslant X \leqslant G = GL(nr, F)$. Choose a minimal intermediate field E between F and K such that $X \geqslant \mathrm{Sp}(nd, E, f_E)$ (where $d = \dim_E K$) for an f_E. If $E = F$, $X \geqslant \mathrm{Sp}(nr, F)$, apply Lemma 4. So we suppose $E \supsetneq F$ and replace K by E to reduce to the case in which X contains no $\mathrm{Sp}(nd, E)$ with $E \subsetneq K$. If $X \leqslant \Gamma = GL(n, K) \rtimes \mathrm{Aut}\, K/F$, by Lemma 4 we know $X \trianglerighteq \mathrm{Sp}(n, K, f)$ or $X \trianglerighteq SL(n, K)$, so Theorem 2 holds. So we suppose $X \nleqslant \Gamma$, and try to find an $\mathrm{Sp}(nd, E) \leqslant X$ with $E \subsetneq K$, thus obtaining a contradiction and completing the proof of Theorem 2.

When $n = 2$, by the proof of Theorem 1 in Section 1 we know X contains an $\mathrm{Sp}(nd, E)$ with $E \subsetneq K$, a contradiction already. So we suppose $n = 2v \geqslant 4$ in the following. We shall choose the basis $\{e_1, ..., e_{2v}\}$ of $V(2v, K)$ such that the inner products $f(e_i, e_{i+v}) = 1 = -f(e_{i+v}, e_i)$ for $1 \leqslant i \leqslant v$, and $f(e_i, e_j) = 0$ for $1 \leqslant i$, $j \leqslant 2v$ and $i \neq j \pm v$. Relative to this basis we have $N_1 = \{A \in GL(n, K) \mid AHA' = H\}$ for $H = \begin{pmatrix} 0 & I_{(v)} \\ -I_{(v)} & 0 \end{pmatrix}$ (where A' is the transpose of A in $\mathrm{Mat}_n K$ and the identity $I_{(v)} \in GL(v, K)$).

Take a $g_1 = (A_{ij}^{(r)})_{n \times n} \in X_{\langle e_{11}, ..., e_{1k} \rangle} \setminus \Gamma$ with maximal $k \leqslant r$. We want to prove that $k = r$. Suppose $k \leqslant r - 1$; we try to obtain a contradiction. Denote by u_{ij} the $((i-1)r + j)$th row of g_1. We can take $z \in N_1$, sending e_1 and $u_{1,k+1}$ into the K-space $Ke_1 \oplus Ke_2 \oplus Ke_{v+2}$. In $g_1 z = (B_{ij}^{(r)})_{n \times n}$ the block $B_{1,v+1}$ has the first $k+1$ rows zero, thus $g_2 = (g_1 z) T_{v+1,1}(I^{(r)})(g_1 z)^{-1} \in X_{\langle e_{11}, ..., e_{1,k+1} \rangle}$. By the maximality of k we must have $g_2 \in \Gamma$, and we see $g_2 \in \Gamma_{Ke_1}$. Applying Lemma 2 we know that g_2 is a transvection of $SL(n, K)$. If $g_2 \notin N_1$ we have $X \geqslant \langle N_1, g_2 \rangle = SL(n, K)$ by Lemma 4; X is known by Theorem 1. Suppose $g_2 \in N_1$; we have some $z_1 \in N_1$ such that $z_1 g_2 z_1^{-1} = T_{v+1,1}(\alpha)$ for an $\alpha^{(r)} \in K^*$, $T_{v+1,1}(\alpha) = \tilde{g}_1 T_{v+1,1}(I)\, \tilde{g}_1^{-1}$ for $\tilde{g}_1 = z_1 g_1 z \in X \setminus \Gamma$. From $\tilde{g}_1(T_{v+1,1}(I) - I) = (T_{v+1,1}(\alpha) - I)\, \tilde{g}_1$ we see $\tilde{g}_1 \in X_{Ke_1}$, $k = r$ as desired.

So we can always find a $g_1 = (A_{ij}^{(r)})_{n \times n} \in X_{Ke_1} \setminus \Gamma$. Applying Lemma 5 to the case $D = K \neq F_2$ and $R = \mathrm{Mat}_r F$ we know X contains a

$$T = \left(\begin{array}{c|c} I^{(vr)} & \\ \hline S & \\ 0 & I^{(vr)} \end{array}\right)$$

with

$$S = \begin{pmatrix} A_0^{(r)} & B_0^{(r)} \\ C_0^{(r)} & 0 \end{pmatrix} \notin \mathrm{Mat}_2\, K,$$

and X contains $T_{\nu+1,1}(A_0^{(r)})$, all $T_{\nu+1,1}(\theta C_0+B_0\theta)$ and all $T_{\nu+1,1}(B_0\theta C_0)$ $(\theta\in K)$. We claim that X contains a $T_1=T_{\nu+1,1}(A_1^{(r)})$ with $A_1\notin K$. When $A_0\notin K$ resp. $C_0+B_0\notin K$ we can take $T_1=T_{\nu+1,1}(A_0)$ resp. $T_{\nu+1,1}(C_0+B_0)$. Suppose both A_0 and C_0+B_0 lie in K; since $S\notin \mathrm{Mat}_2 K$ we must have $B_0\notin K$ and $C_0\notin K$. We can choose an $\alpha^{(r)}\in K$ having the same first row as B_0, and can replace T by $TT_{\nu+1,2}(-\alpha)\,T_{\nu+2,1}(-\alpha)\in X$, thus replacing $S=\left(\begin{smallmatrix}A_0 & B_0\\ C_0 & 0\end{smallmatrix}\right)$ by $S-\left(\begin{smallmatrix}0 & \alpha\\ \alpha & 0\end{smallmatrix}\right)=\left(\begin{smallmatrix}A_0 & B_0-\alpha\\ C_0-\alpha & 0\end{smallmatrix}\right)$, to annihilate the first row of B_0, thus reducing to the case in which B_0 is singular. So we suppose B_0 is singular, and B_0 and C_0 are non-zero (since they are not in K). We can choose a $\theta\in K^*$ sending a non-zero row u of B_0 to $u\theta\notin \mathrm{Ker}\,C_0$, thus $B_0\theta C_0\neq 0$. But $B_0\theta C_0$ is singular (since B_0 is singular), thus $B_0\theta C_0\notin K$, $T_1=T_{\nu+1,1}(B_0\theta C_0)\in X$ is just what we need.

So we can always find a $T_1=T_{\nu+1,1}(A_1^{(r)})\in X$ with $A_1\notin K$. We can take a $\theta_1\in K$ having the same first row as A_1 and replace T_1 by $T_1T_{\nu+1,1}(-\theta_1)\in X$, thus replacing A_1 by $A_1-\theta_1$, to annihilate the first row of A_1. We have $0<\mathrm{rank}\,A_1<r$ for such A_1. Choose a $T_{\nu+1,1}(A)$ with smallest $h=\mathrm{rank}\,A>0$, then $h<r$. By Lemma 5 (3) we know that $T_{\nu+1,1}(A\theta A)\in X$ for all $\theta\in K$. Since $\mathrm{rank}\,(A\theta A)\leqslant \mathrm{rank}\,A=h$, we must have either $\mathrm{rank}\,(A\theta A)=h$ or $A\theta A=0$. By Lemma 1 we know that $\mathrm{Im}\,A$ and $\mathrm{Ker}\,A$ are subspaces of K over the field E generated by all $\beta\beta_1^{-1}$ $(\vec\beta\in \mathrm{Im}\,A)$ for a given non-zero $\vec\beta_1\in \mathrm{Im}\,A$, with $\dim_E(\mathrm{Im}\,A)=1$ and $\dim_E(\mathrm{Ker}\,A)=d-1$ (where $d=\dim_E K=r/h$). We can choose a basis $\{w_1, ..., w_d\}$ of K/E and a basis $\{\varepsilon_1, ..., \varepsilon_h\}$ of E/F to construct a basis $\{\varepsilon_j w_i \mid 1\leqslant i\leqslant d,\ 1\leqslant j\leqslant h\}$ of K/F to replace $\{k_1, ..., k_r\}$, to reduce A to the form $\left(\begin{smallmatrix}\delta & 0\\ & \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix}0 & \\ & \delta\end{smallmatrix}\right)$ with $\delta\in GL(h, F)$ and $K\subset \mathrm{Mat}_d E$. We can choose an $\alpha^{(h)}\in E$ having the same first row as $\delta^{(h)}$, thus $\mathrm{rank}\,(\delta-\alpha)<h$. When $A=\left(\begin{smallmatrix}\delta & 0\\ & \end{smallmatrix}\right)$, take

$$\eta=\begin{pmatrix} * & \cdots & * & I^{(h)}\\ & & & 0\\ \vdots & \cdots & \vdots & \vdots\\ * & \cdots & * & 0\end{pmatrix}\in K^*\subset \mathrm{Mat}_d E.$$

By Lemma 5 (3) we have $T_{\nu+1,1}(\eta A\eta)\in X$ with

$$\eta A\eta=\begin{pmatrix} * & \cdots & * & \delta\\ & & & 0\\ & & \cdot^{\cdot^{\cdot}} & \\ 0 & & & \end{pmatrix}.$$

Take a $\theta_1=(\alpha_{ij})_{d\times d}\in K^*$ with $\alpha_{1d}=\alpha$, then by Lemma 5 (3) we have $T_{\nu+1,1}(A_2)\in X$ for $A_2=A(\eta A\eta-\theta_1)A=\left(\begin{smallmatrix}\delta_1 & 0\\ & \end{smallmatrix}\right)$ with $\delta_1=\delta(\delta-\alpha)\delta$, rank $A_1 = \text{rank}\,(\delta - \alpha) < h$; by the minimality of h we must have $\delta - \alpha = 0$, $\delta = \alpha \in E^*$. When $A = (^{0}\ {}_{\delta})$ we take $\alpha^{(r)} = \text{diag}(\alpha^{(h)}, ..., \alpha^{(h)}) \in K^*$. Since $T_{v+1,1}(A)$ and $T_{v+1,1}(A - \alpha)$ lie in X, we know $T_{v+1,1}(A(A-\alpha)A) \in X$ by Lemma 5 (3), with $A(A-\alpha)A = (^{0}\ {}_{\delta_1})$, $\delta_1 = \delta(\delta - \alpha)\delta$, still leading to $\delta = \alpha \in E^*$. So we have $\delta \in E^*$ anyway, $T_{v+1,1}(A) \in X$ is a transvection of $SL(nd, E)$. Furthermore, $T_{v+1,1}(A)$ is just the symplectic transvection $\rho_{w_t e_1, 1}$: $x \mapsto x + f_E(x, w_t e_1)\, w_t e_1$ of the symplectic group $\text{Sp}(2vd, E, f_E)$ relative to the alternating E-form $f_E = \varphi_E f$, with $\varphi_E \in \text{Hom}_E(K, E)$ defined by

$$\varphi_E(w_i w_t) = \begin{cases} 0 & \text{when} \quad i \neq d \\ \delta & \text{when} \quad i = d, \end{cases}$$

where $t = 1$ when $A = ({}_{\delta}\ ^{0})$, $t = d$ when $A = (^{0}\ {}_{\delta})$. X contains the conjugates $g^{-1} \rho_{w_t e_1, 1}\, g = \rho_{w_t e_1 g, 1}$ of $\rho_{w_t e_1, 1} = T_{v+1,1}(A) \in X$ by all $g \in N_1 < \text{Sp}(2vd, E, f_E)$. Since $w_t e_1 g$ $(g \in N_1)$ ranges over all the non-zero vectors we know $\rho_{w_t e_1 g, 1}$ ranges over all the conjugates of $\rho_{w_t e_1, 1}$ in $\text{Sp}(2vd, E, f_E)$; all such $\rho_{w_t e_1 g, 1}$ generate a normal subgroup of $\text{Sp}(2vd, E, f_E)$ which must coincide with $\text{Sp}(2vd, E, f_E)$, $X \geqslant \text{Sp}(2vd, E, f_E)$, contradicting our assumption (remember that we assume X contains no $\text{Sp}(2vd, E)$ with $E \subsetneq K$), thus completing the proof of Theorem 2.

REFERENCES

1. M. ASCHBACHER, On the maximal subgroups of finite classical groups, *Invent. Math.* **76** (1984), 469–514.
2. W. M. KANTOR, Linear groups containing a singer cycle, *J. Algebra* **62** (1980), 232–234.
3. R. H. DYE, A maximal subgroup of $PSp_6(2^m)$ related to a spread, *J. Algebra* **84** (1983), 128–135.
4. R. H. DYE, Maximal subgroups of symplectic groups stabilizing spreads, *J. Algebra* **87** (1984), 493–509.
5. R. H. DYE, Maximal subgroups of $PSp_{6n}(q)$ stabilizing spreads of totally isotropic planes, *J. Algebra* **99** (1986), 191–209.