The Structure of the Symplectic Groups Over Semilocal Domains

CHAN-NAN CHANG

Department of Mathematics and Statistics, University of Massachusetts, Amherst, Massachusetts 01002

Communicated by J. Dieudonné

Received February 20, 1974

Dickson and Dieudonné have shown in [1] and [2] that all proper normal subgroups of a symplectic group defined over a field with more than three elements are contained in its center \{±1\}.

Attempts at integral analogues of this theorem have been quite successful. In 1963 Klingenberg [3] effected a significant generalization of these results by extending them to an arbitrary local rings whose residue class field has characteristic not equal to two and has more than three elements. However he assumed that the underlying alternating form is unimodular, i.e., has unit discriminant. He was able to show in this case that every normal subgroup of the symplectic group is a congruence subgroup in the usual sense. In 1966 Riehm [5] generalized Klingenberg's results by dropping the requirement that the discriminant be a unit. In doing so he restricted the ring to be a valuation ring with residue class field of characteristic not equal two and having more than three elements. In this situation it turns out that there are normal subgroups which are not congruence subgroups. In order to regain a complete description of the normal subgroups, Riehm generalized the concept of congruence subgroup by allowing the congruence ideal to vary from entry to entry in the matrix. He was able to show that every normal subgroup of the symplectic group is a congruence subgroup in this new sense.

In this paper we shall generalize these results to the case where the underlying ring is a semilocal domain with suitable residue class fields. We also drop the assumption that the discriminant be a unit, but we assume that the module has a canonical splitting (See Section 1). By using Riehm's approach and his notion of congruence subgroup, we are able to prove the following main structure theorem of a symplectic group (listed as Theorem 7.9):

A subgroup of a symplectic group is a normal subgroup if and only if it is a congruence subgroup.
We also include theorems concerning the generators of the congruence subgroups (listed as Theorem 6.5, 6.6, and 7.11) and the relations between two congruence subgroups of the same order (listed as Theorem 4.1).

1. Basic Properties of Lattices

Throughout this paper \( \mathcal{O} \) will be a semilocal domain, i.e., a commutative Noetherian domain which has finitely many maximal ideals. We assume that 2 is a unit in \( \mathcal{O} \) and every residue class field of \( \mathcal{O} \) contains more than three elements. Let \( U \) be the multiplicative group of all units of \( \mathcal{O} \) and \( F \) be the quotient field of \( \mathcal{O} \). If \( S \) is a nonempty subset of \( F \), the fractional ideal generated by \( S \) is denoted by \( [S] \). If \( S = \{a\} \) we use \( [a] \) instead of \( \{[a]\} \). If \( \mathcal{O} \) and \( \mathcal{B} \) are nonzero ideals, we define

\[
\mathcal{O} : \mathcal{B} := \{a \in \mathcal{O} \mid a\mathcal{B} \subseteq \mathcal{O}\}.
\]

If \( \mathcal{O} = [a] \) and \( \mathcal{B} = [b] \) we use \( a : b \) instead of \([a] = [b] \). For any \( a \) in \( F \) we define \( [a]^{-1} = \{b \in F \mid ab \in \mathcal{O}\} \). Clearly \( a : b = [a^{-1}] \cap \mathcal{O} = [b]^{-1}([a] \cap [b]) \) and \( (a : b)b = (b : a)a \).

A lattice \( L \) is a free \( \mathcal{O} \)-module of finite rank, endowed with a nondegenerate alternating bilinear form. Thus there is a bilinear mapping of \( L \times L \) into \( \mathcal{O} \), denoted by \((x, y) \mapsto xy\), such that \( xy = -yx \) for all \( x \) and \( y \) in \( L \), (whence \( x^2 = 0 \) for \( x \) in \( L \)) and such that \( xL = 0 \) only for \( x = 0 \). The rank of \( L \) is denoted by \( \dim L \).

Let \( s \) in \( \mathcal{O} \). A lattice \( L \) is called \( s \)-modular if there exists a basis \( \{x_1, \ldots, x_n\} \) of \( L \) such that \( x_1L = [s] \) for all \( i = 1, \ldots, n \). If \( s \) is a unit then an \( s \)-modular lattice is called unimodular.

We say that \( L \) has a canonical splitting if there exists a nonempty subset \( S = \{s_1, \ldots, s_t\} \) of \( \mathcal{O} \) with the following properties:

(1.1.a) \( s_i = s_j \) if and only if \( i = j \).

(1.1.b) \( \sum_{i=1}^{t} (s : s_i)(s_i : s) = \mathcal{O} \) if and only if

\[
s = s_i \quad \text{for some} \quad i = 1, \ldots, t.
\]

(1.1.c) for any fixed \( j = 1, \ldots, t \)

\[
\sum_{i \neq j} (s_j : s_i)(s_i : s_j) \subseteq \mathcal{O}
\]

such that

\[
L = I_1 \perp I_2 \perp \cdots \perp I_t \quad (1.1)
\]
in which each component \( L_i \) is \( s_i \)-modular for \( i = 1, \ldots, t \). If \( L \) has a canonical splitting (1.1) it may not be unique. However we shall show that the quantities:

\[
t, s_i, \quad \text{and} \quad n_i = \dim L_i, \quad i = 1, \ldots, t, \tag{1.2}
\]

are invariants of \( L \).

The assumptions (1.1.b) and (1.1.c) are made mainly to assure the uniqueness of the quantities in (1.2). If \( R \) is a local domain or \( L \) is \( s \)-modular then (1.1.a) implies (1.1.b) and (1.1.c).

If \( M \) is a nonempty subset of \( L \), \( M^0 \) is its orthogonal complement (in \( L \)), and \( \langle M \rangle \) is the submodule of \( L \) spanned by \( M \). A submodule \( M \) of \( L \) is said to split \( L \) if \( L = M \perp N \) for some submodule \( N \). It is easy to see that if \( L = M \perp N \) then \( N = M^0 \). We call \( M \) an orthogonal component of \( L \) if \( M \) splits \( L \).

Let \( \mathfrak{A} \) be a nonzero ideal and \( M \) a submodule of \( L \). We define

\[
M^{\mathfrak{A}} = \{ x \in M \mid xM \subseteq \mathfrak{A} \}.
\]

Clearly \( M^{\mathfrak{A}} \) is a submodule of \( M \), and \( M^{\mathfrak{A}}M^{\mathfrak{A}} \subseteq \mathfrak{A} \). If \( M = J \perp K \) then \( M^{\mathfrak{A}} = J^{\mathfrak{A}} \perp K^{\mathfrak{A}} \). If \( MM \subseteq \mathfrak{A} \), \( M^{\mathfrak{A}} = M \). If \( M \) is modular with \( \mathfrak{A} \subseteq \mathfrak{B} = MM \), then \( M^{\mathfrak{A}} = (\mathfrak{A} : \mathfrak{B})M \). If \( \mathfrak{A} = [s] \) we write \( M^s \) instead of \( M^{[s]} \).

If \( L \) has a canonical splitting (1.1) with invariants (1.2) then

\[
L^{\mathfrak{A}} = \bigoplus_{1 \leq i \leq t} ((Ls_i^{-1}) \cap \mathfrak{A})L_i, \tag{1.3}
\]

or \( \bigoplus_{1 \leq i \leq t} ([s_i] : \mathfrak{A})L_i \).

A vector \( x \) in \( L \) is called maximal (in \( L \)) if \( x = \alpha y \) with \( y \) in \( L \) implies \( \alpha \) is a unit.

A canonical basis for a lattice \( L \) is a basis \( \{ x_1, \ldots, x_n \} \) with the following properties.

(i) \( n \) is even,

(ii) if \( i < j \) then \( x_i x_j \neq 0 \) if and only if \( i = 2k - 1 \) and \( j = 2k \) for some \( k \), \( 1 \leq k \leq n/2 \).

Let \( \{ p_j \mid j \in J \} \) be the set of all maximal ideals of \( \mathfrak{A} \) and let \( k_j = \mathfrak{A}/p_j \), for \( j \) in \( J \), be the corresponding residue class fields. Let \( -(or \pi_j): \mathfrak{A} \rightarrow k_j \) be the canonical homomorphism. For any lattice \( L \), we define \( L(s, j) = L^s \cap L^{p_j} \) for \( s \) in \( \mathfrak{A} \) and \( j \) in \( J \). We use the same notation \( -(or \pi_j) \) to denote the canonical map: \( L^s \rightarrow L(s, j) \). We define canonically \( \bar{x} + \bar{y} = x + y \), \( \bar{x} \cdot \bar{x} = \bar{xx} \) and \( \bar{xy} = s^{-1}(xy) \) for all \( x, y \) in \( L^s \) and all \( \alpha \) in \( \mathfrak{A} \). It can be shown easily that \( L(s, j) \) is a module with alternating form over \( k_j \). If \( L \) has a canonical splitting (1.1),
then $L(s_i, j) = L_1 \perp L_2 \perp \cdots \perp L_t$ and $n_i \leq \dim(L(s_i, j)) \leq n$. We denote $L(s_i, j)$ by $L(i, j)$.

We define the dual of $L$ to be

$$L^* := \{ x \in FL \mid xL \subseteq \mathcal{C} \}.$$  

One can easily show that $L^*$ is a lattice and

$$FL^* = FL, \quad L^* = L$$

and

$$(aL)^* = a^{-1}L^*$$

for any $a$ in $F$. If $L$ has a splitting $L = J \perp K$, then

$$L^* = J^* \perp K^*.$$  

Finally, it is easy to see that

$$L \supseteq K \rightarrow L^* \subseteq K^*$$

and

$$(J + K)^* = J^* \cap K^*.$$  

An automorphism $\sigma$ of $L$ is called an isometry if $(ux)(uy) \preccurlyeq sxy$ for all $x$ and $y$ in $L$. The symplectic group of $L$, $S_L(L)$, is the group of all isometries of $L$.

**Theorem 1.1.** Let $L$ be a lattice with a canonical splitting (1.1). Then the quantities $t$, $s_i$, and $n_i = \dim L_i$ for $i = 1, \ldots, t$ of (1.2) are invariants of $L$, i.e., they are independent no matter which canonical splitting of $L$ is used to calculate them.

**Proof.** Let $s$ in $\mathcal{C}$. By (1.3), $L^* = 1_{i \leq i \leq t}(s : s_i)L_i^{s_i}$ and $L^* \cdot L^* = s(\sum_{i=1}^t (s_i : s_i))$. By (1.1.b), $L^* \cdot L^* = s\emptyset$ if and only if $s\emptyset = s_i\mathcal{C}$ for some $i = 1, \ldots, t$. Hence the ideals $[s_1], \ldots, [s_t]$ and $t$ are invariant.

Now we show that $n_i = \dim L_i$ is invariant for $i = 1, \ldots, t$. By (1.1c), for a fixed $i$ there exists a maximal ideal $p_{s_i}$ which contains $\sum_{i=1}^t (s_i : s_i)(s_i : s_j)$. We consider $L(i, \mu) = L_i^n/L_i^{s_ip_{s_i}}$ with the induced alternating form over $k$. By (1.3), $L^{s_i} = 1_{i \leq j \leq t}(s_i : s_j)L_j$. For each $j \neq i$, $(s_i : s_j)p_{s_j} = s_i((s_i : s_j)(s_j : s_j)) \subseteq s_ip_{s_i}$. Hence $L(i, \mu) = L_i(t, k)$ and $\dim L_i = \dim L_i(t, k)$ is an invariant.

**Q.E.D.**

**Proposition 1.2.** Let $M$ be a lattice over $\mathcal{C}$. Suppose that for each orthogonal component $N$ of $M$ there exists $x$ and $y$ in $N$ such that $[x \cdot y] = xN + yN$. Then $M$ has a canonical basis.
Proof. Since $M$ is its own orthogonal component, there exists $x$ and $y$ in $M$ such that $[x \cdot y] = xM + yM$. Let $FM = (Fx + Fy) \perp N$ and let $M' = (Cx + Cy) \perp N \cap M$. We claim that $M = M'$. It suffices to show that $M \subseteq M'$. Let $z$ be an arbitrary element in $M$ and write $z = \alpha x + \beta y + w$, where $\alpha$, $\beta$ are in $F$ and $W$ is in $N$. Since $\alpha(xy) = zy$ is in $[xy]$, $\alpha$ is in $\mathcal{O}$. Similarly $\beta$ is in $\mathcal{O}$. We have $W = z - \alpha x - \beta y$ is in $N \cap M$ and so $M = M'$. The result now follows from the induction on $\dim M$. Q.E.D.

Corollary 1.3. If $\mathcal{O}$ is a principal ideal domain then every lattice $M$ has a canonical basis.

Proof. It is easy to check that every lattice $M$ satisfies the hypothesis of 1.2. Q.E.D.

Lemma 1.4. Suppose $L$ is $s$-modular. Then

$$L = \{x \in FL \mid xL \subseteq s\mathcal{O}\}.$$ 

Proof. Let $M = \{x \in FL \mid xL \subseteq s\mathcal{O}\}$. Since $L$ is $s$-modular we have $L \cdot L \subseteq s\mathcal{O}$ and so $L \subseteq M$. Conversely consider an $x$ in $FL$ with $xL \subseteq [s]$. Then

$$xL_{x} = x(s^{-1}L) \subseteq \mathcal{O},$$

hence $x$ is in $L_{x} = L$.

Now assume that $L$ is $s$-modular. Then $L \cdot s^{-1}L = \mathcal{O}$ so that $s^{-1}L \subseteq L_{x}$. On the other hand for any $y = ax$ in $L_{x}$, where $a \in F$ and $x \in L$, $yL = a(xL) \subseteq \mathcal{O}$, since $xL \subseteq s\mathcal{O}$, $a \in s^{-1}\mathcal{O}$. Hence $s^{-1}L = L_{x}$.

Q.E.D.

Lemma 1.5. Let $M$ be an $s$-modular sublattice of $L$. Then $M$ splits $L$ if and only if $ML = s\mathcal{O}$.

Proof. If $M$ splits $L$ we have $L = M \perp N$ and $ML = MM = s\mathcal{O}$. Conversely suppose $ML = s\mathcal{O}$. Write $FL = FM \perp N$. We claim that

$$L = M \perp (L \cap N).$$

It is enough to show that a typical $x$ in $L$ is also in $M \perp (L \cap N)$. Write $x = y + z$ where $y$ is in $FM$ and $z$ is in $N$. Then

$$yM = xM \subseteq ML \subseteq s\mathcal{O}.$$ 

But $y$ is in $FM$, $y$ is in $M$ by 1.4, so $z$ is in $L$ and we are done. Q.E.D.
**Lemma 1.6.** Let $M$ be a sublattice of $L$. Suppose $M$ has a canonical splitting

$$M = M_1 \perp \cdots \perp M_r.$$ 

Then $M$ splits $L$ if and only if $M_i$ splits $L$.

**Proof.** If $M$ splits $L$ then it is obvious that each $M_i$ does too. Conversely suppose $L$ is split by each $M_i$. Write

$$FL = FM \perp N.$$ 

Similar to the proof of 1.5, one can show that

$$L = M \perp N \cap L.$$ 

**Lemma 1.7.** Let $L$ be a lattice with canonical splitting (1.1) and let $M$ be a sublattice of $L$ which splits $L$. Suppose $M$ has a canonical splitting

$$M = M_1 \perp \cdots \perp M_r.$$ 

Then each component $M_i$ is $s_{(i)}$-modular.

**Proof.** Suppose there exists some $i = 1, \ldots, t$, such that $M_i$ is $s$-modular for some $s \neq s_j$ for $j = 1, \ldots, t$. By (1.1.6), $\sum_{i=1}^t (s : s_j)(s_j : s) \subset C$, we may choose a maximal ideal $p_i$ of $C$ which contains $\sum_{j=1}^t (s : s_j)(s_j : s)$. By (1.3),

$$L^s = \bigcap_{1 \leq j \leq t} (s : s_j) L_j.$$ 

Hence

$$L^s \cdot L^s = \sum_{j=1}^t (s : s_j)^2 s_j = s \left( \sum_{j=1}^t (s_j : s)(s : s_j) \right) \subseteq sp_i.$$

Therefore $L(s, k) - L^s L^s = 0$. On the other hand $L = M \perp M^0$, $L(s, k) = M(s, k) \perp M^0(s, k)$. Since $M_i$ is $s$-modular, $M(s, k) \neq 0$. Thus $L(s, k) = 0$ which is a contradiction. Hence the lemma follows. Q.E.D.

**Lemma 1.8.** Let $L$ be a lattice with canonical splitting (1.1). Let $x$ and $y$ be two maximal vectors in $L$ such that $C x + C y$ splits $L$. Then there exists a canonical basis of $L$ which contains $x$ and $y$ as basic vectors.

**Proof.** Let $M = Cx + Cy$ and write $L = M \perp M^0$. By 1.7, we may assume $xy = s_i$ for some $i = 1, \ldots, t$. It is easy to see that $L(i, j) = M(i, j) \perp M^0(i, j)$ for all $j$ in $J$. If $n_i = 2$ then the result follows from 1.1 and induction on $t$. If $n_i > 2$ then $M^0(i, j) \neq 0$ for all $j$ in $J$. For each $j$ in $J$, there exists $z_j$ and $w_j$ in $M^0(i, j)$ with $z_j \cdot w_j \neq 0$ in $k_j$. We choose $z_j$ and $w_j$ in $(M^0)^i$ with canonical images $\bar{z}_j$ and $\bar{w}_j$ in $M^0(i, j)$, respectively. By Chinese Remainder theorem, there exists, for each $j$ in $J$, an element $d_j$ in $C$ such that $d_j \equiv 1$ (mod $p_j$) and $d_j \equiv 0$ (mod $p_\mu$) for all $\mu \neq j$. We consider

$$z = \sum_{j=1}^t d_j z_j \quad \text{and} \quad w = \sum_{j=1}^t d_j w_j.$$
Then $z$ and $w$ are in
\[ \pi_j(s_j^{-1}(z \cdot w)) = \pi_j \left( s_j^{-1} \left( \sum d_j d_j z_j w_j \right) \right) = \pi_j(s_j^{-1} d_j z_j w_j) \neq 0 \]
for all $j$ in $J$. Hence $s_j^{-1}(z \cdot w)$ is a unit and $[z \cdot w] = s_j \mathcal{O} = zM^0 + wM^0$.

By 1.2, $\langle z, w \rangle$ splits $M^0$. Now the result follows from the induction on $n_i$ and $t$.

**Q.E.D.**

**PROPOSITION 1.9.** Let $L$ be a lattice with canonical splitting (1.1). Let $M$ be a sublattice which splits $L$. If $M$ has a canonical basis (or canonical splitting) then so does $M^0$.

**Proof.** Apply 1.6, 1.7, 1.8 and induction on dim $M$.

**Q.E.D.**

**COROLLARY 1.10.** Let $L$ be a lattice with canonical splitting (1.1). Then $L$ has a canonical basis.

**Proof.** Apply 1.8 and induction on dim $L$.

**Q.E.D.**

**LEMMA 1.11.** Let $a$ and $b$ be in $\mathcal{O}$ with $a\mathcal{O} + b\mathcal{O} = \mathcal{O}$. Then there exists a unit $d$ in $\mathcal{O}$ such that $a + bd$ is a unit.

**Proof.** If $\mathcal{O}$ is a field then $\mathcal{O}$ contains more than three elements. Hence the lemma follows easily. Assume $\mathcal{O}$ is not a field. For each $j$ in $J$, $k_j$ contains more than three elements, there exists $d_j$ in $\mathcal{O}$ but not in $p_j$ such that $a + d_j b$ is not in $p_j$. By Chinese Remainder theorem, there exists $d$ in $\mathcal{O}$ such that $d - d_j$ is in $p_j$ for $j$ in $J$. Hence $d$ and $a + bd$ both are not in $p_j$ for all $j$ in $J$. Therefore $d$ and $a + bd$ both are units.

**Q.E.D.**

Throughout the rest of the paper we shall be working with a lattice $L$ which has a canonical splitting (1.1). Whenever the symbols $t$, $s_i$, $n_i$ appear, they will be the invariants of $L$ given in (1.2) unless stated otherwise explicitly. We also assume that $n_i \geq 4$ for all $i = 1,\ldots, t$.

2. **Invariant Modules**

An $\mathcal{O}$-submodule $M$ of $FL$ is called invariant if $\sigma M = M$ for all in $S_p(L)$. We shall determine all such invariant modules.

A pure vector $x$ of $L$ is one which is in some canonical basis for $L$; $x$ is pure of type $i$ (or $i$-pure) if $xL = [s_i]$.

A transvection is an element $\sigma \in S_p(L)$ of the form: $\sigma(x) = x + w(wx)\lambda$ with
Let $x$ be an $i$-pure vector contained in the invariant module $M$. Choose a canonical basis for $L$ containing $x$ and let $y$ be the vector in this basis with $xy = \mu \neq 0$. Then $L^{s_i} = (Cy + Cy) \cap N$ by 1.2 and 1.8. Let $z$ be any vector in $N$ and let $w = x - y - z$. Then $\tau_{w, -u}$ in $S_p(L)$. Carries $x$ to $y + z$. Thus $M$ contains $y + z$ and $z$, so $M$ contains $L^{s_i}$. Q.E.D.

**Lemma 2.2.** Let $M$ be an $s$-modular submodule of $L$ and $2^G$ be an ideal in $G$. Then $G!M = \{x \in M \mid s^{-1}(xM) \subseteq 2^G\}$.

**Proof.** Let $N = \{x \in M \mid s^{-1}(xM) \subseteq 2^G\}$. $G!M \subseteq N$ is clear. Let $y \in N$. Choose a canonical basis $B$ for $M$ and write $y$ as linear combination of vectors in $B$. Then it follows easily that $y \in G!M$ and so $N \subseteq G!M$. Q.E.D.

**Theorem 2.3.** If $M$ is an invariant module, there are ideals $g_1, \ldots, g_t$ such that

$$M = g_1 L^{s_1} + g_2 L^{s_2} + \cdots + g_t L^{s_t},$$

(2.1)

where the ideals $g_1, \ldots, g_t$ are uniquely determined and are given by

$$g_j = s^{-1}_j(ML^{s_j}) \quad j = 1, \ldots, t.$$  

(2.2)

If $L = L_1 \perp \cdots \perp L_t$ is an arbitrary canonical splitting, then

$$M = g_1 L_1 \perp g_2 L_2 \perp \cdots \perp g_t L_t,$$

(2.3)

and

$$g_i(s_i : s_j) \subseteq g_j \quad \text{for all } i \text{ and } j.$$  

(2.4)

**Proof.** Let $g_j = s_j^{-1}(ML^{s_j})$. We divide the proof into three steps.

**Step 1.** We show that $\sum_{i=1}^t g_i L^{s_i} \subseteq M$. Let $a \in g_1$. Since $M$ is invariant, there exists an $i$-pure vector $x$ in $L$ and a vector $y$ in $M$ with $xy = a s_i$. $\tau_{x, s_i^{-1}}$ is in $S_p(L)$ and carries $y$ to $y + ax$. Thus $ax$ in $M$. The assertion follows from 2.1.

**Step 2.** We show that $M \supseteq \sum_{i=1}^t g_i L_i$. Let $x = \sum_{i=1}^t a_i x_i$ in $M$ with $x_i$ is a maximal vector in $L_i$ for $i = 1, \ldots, t$. Choose $y_i$ in $L$ with $y_i x_i = s_i$. Then $a_i y_i = \tau_{y_i, s_i^{-1}}(x) - x$ in $M \cap g_i L_i$. By 2.2, $a_i \in g_i$, so $x \in \sum_{i=1}^t g_i L_i$. This proves (2.1), (2.2), and (2.3).
Step 3. Let \( a \in g_i(s_i : s_j), a = b(s_i^{-1}(xy)) \) where \( b \in (s_i : s_j), \ x \in M \) and \( y \in L^{s_i}. \) Rewrite \( a = s_j^{-1}(x \cdot (bs_i^{-1}s_j)y). \)

It is easy to see that \( bs_i^{-1}s_j \in \mathcal{O} \) and \( bs_i^{-1}s_jy \in L^{s_i}, \) (2.4) is now immediate.

Q.E.D.

Remark. Since the ideals \( g_1, \ldots, g_t \) characterize \( M \) completely, we shall often write \( M = (g_1, \ldots, g_t). \) Notice that it follows from the theorem that a nonzero invariant module contains a nonzero vector from every line of \( FL. \)

If \( X \) is a nonempty subset of \( L, \) we denote by \( l(X) \) the smallest invariant module containing \( X. \) It is the intersection of all invariant modules containing \( X \) and is also given as the set of all finite sums.

\[
Z(X) = \sum_{a_i \sigma_i(x_i)} a_i \text{ in } \mathcal{O}, \ x_i \text{ in } X^t.
\] (2.5)

We note that \( Z(X) = Z(X). \)

**Lemma 2.4.** Let \( X \) be a subset of \( L \) and let \( g_j = a_j^{-1}(XL^{s_j}) \) for \( j = 1, \ldots, t. \) Then

\[
l(X) = (g_1, \ldots, g_t).
\] (2.6)

**Proof.** It is easy to see that \( g_j = s_j^{-1}(l(X)L^{s_j}) \) for \( j = 1, \ldots, t. \) Hence the lemma follows from 2.3.

Q.E.D.

3. **Orders and Tableaus**

We use \( \delta \) to denote \( \pm 1 \) and \( \delta_L \) (or \( \delta \)) to denote the mapping \( x \rightarrow \delta x \) of \( L. \) Let \( \sigma \) in \( S_p(L) \): we define

\[
O_i(\sigma) = \bigcap_\delta l((\sigma - \delta L)L^{s_i}), \quad i = 1, \ldots, t.
\] (3.1)

\( O_i(\sigma) \) is an invariant module and

\[
O_i(\sigma) = \min_\delta l((\sigma - \delta L)L^{s_i}), \quad L^{s_i} = \max_\delta l((\sigma - \delta L)L^{s_i}).
\] (3.2)

We define the **order** of \( \sigma, \ O(\sigma), \) to be the function which assigns to each ordered pair of integers \( (i, j), \ 1 \leq i, j \leq t, \) the ideal \( O_{i,j}(\sigma) \) defined by \( O_{i,j}(\sigma) = (O_{i,j}(\sigma), \ldots, O_{i,j}(\sigma)). \)

It follows from (3.2) and Lemma 2.4 that

\[
s_jO_{i,j}(\sigma) = \min_\delta l((\sigma - \delta L)L^{s_i})L^{s_j}.
\]
Let us temporarily put \( g_i = O_i(\sigma) \). Then by (3.1),
\[
g_i \subseteq L^i; \quad i = 1, \ldots, t, \tag{3.3}
\]
and since \((s_j : s_i)L^i \subseteq L\)
\[
(s_j : s_i)g_i \subseteq g_j \quad i, j = 1, \ldots, t. \tag{3.4}
\]

If we put \( g_{ij} = O_{ij}(\sigma) \), these conditions can be restated as
\[
g_{ij} \subseteq (s_i : s_j) \tag{3.5}
\]
\[
(s_k : s_i)g_{ij} \subseteq g_{kj} \quad \text{for all } i, j, k. \tag{3.6}
\]

The condition (2.4) leads to
\[
g_{ij}(s_j : s_k) \subseteq g_{ik} \quad \text{for all } i, j, k. \tag{3.7}
\]

**Lemma 3.1.** Suppose that \( \sigma \) in \( S_p(L) \) and that for some choice of \( \delta = \pm 1 \), we have \( O_i(\sigma) = l((\sigma - \delta_i)L^i) \) and \( O_j(\sigma) = l((\sigma + \delta_j)L^j) \). Then \( O_{ij}(\sigma) = s_i : s_j \).

**Proof.** By (3.2), \( L^1 = l((\sigma - \delta_i)L^i) \), whence by (3.4), \( (s_i : s_j)L^i \subseteq O_i(\sigma) \) and \( (s_i : s_j) \subseteq O_{ij}(\sigma) \). The lemma now follows from (3.5). Q.E.D.

We put \( O_{ij}(\sigma) = g_{ij} \) as above. It follows from Lemma 3.1 and from the identity \((\alpha x - \delta x)y = +\delta(\alpha y - \delta y)\alpha x\) that
\[
s_i g_{ij} = s_j g_{ji}. \tag{3.8}
\]

Let \( g \) be a mapping from the pairs of integers \((i, j)\) with \( i, j = 1, \ldots, t \), into the set of ideals of \( L \). Write \( g_{ij} \) for the image of \((i, j)\) under \( g \). Then \( g \) is called a tableau (for \( L \)) if it satisfies the conditions (3.5)–(3.8). If \( g \) is a tableau, \( g_i \) will denote the invariant lattice \( g_i = (g_{i1}, \ldots, g_{it}) \).

By what we have proved above, the order \( O(\sigma) \) of any fixed \( \sigma \) in \( S_p(L) \) is a tableau.

The set of tableaux (for \( L \)) is partially ordered by defining \( g \leq h \) if \( g_{ij} \subseteq h_{ij} \) for all \( i \) and \( j \).

Let \( G \) be any non-empty subset of \( S_p(L) \). The mapping
\[
(i, j) \rightarrow O_{ij}(G) = \sum_{\sigma \in G} O_{ij}(\sigma)
\]
is the order \( O(G) \) of \( G \). Clearly \( O(G) \) is a tableau.

The following lemma follows from the identity
\[
\sigma \rho - (\delta \rho')_L = (\sigma - \delta_L)\rho + \delta_L(\rho - \delta_L'). \tag{3.9}
\]
**Lemma 3.3.** Suppose $\sigma, \rho$ are in $S_p(L)$ and $\delta, \delta' = \pm 1$. Then

$$l((\sigma \rho - (\delta \delta')) L) \subseteq l((\sigma - \delta_L) L^{\delta_i}) + l((\rho - \delta_L) L^{\delta_i})$$

and $O_i(\sigma \rho) \subseteq O_i(\sigma) + O_i(\rho)$ for $i = 1, \ldots, t$.

### 4. Congruence Subgroups of $S_p(L)$

Let $g = (g_i) = (g_{ij})$ be a tableau. The set of all $\sigma$ in $S_p(L)$ with $O(\sigma) \leq g$ is called the (general) congruence group of $S_p(L)$ corresponding to $g$. It is denoted by $G_S(L, g)$ or $G_{S_p}(g)$, or $G_{S_p}$, when $L$ and $g$ are clearly determined by the context.

Since $S_p(g)$ is also characterized as the set of those $\sigma$ such that $O_i(\sigma) \subseteq g_i$ for $i = 1, \ldots, t$, it follows from Lemma 3.3 that it is closed under multiplication. Moreover $O(\sigma^{-1}) = O(\sigma)$ since $(\sigma^{-1} - \delta_L) L^{\delta_i} = (\sigma - \delta_L) L^{\delta_i}$, whence $S_p(g)$ is a subgroup of $S_p(L)$. That it is actually a normal subgroup follows from

$$\rho \sigma \rho^{-1} - \delta_L = \rho(\sigma - \delta_L) \rho^{-1}. \quad (4.1)$$

The special congruence group of $S_p(L)$ corresponding to $g$ is the set $SS_p(L, g)$ of those $\sigma$ in $S_p(L)$ with the property $(\sigma - 1) L^{\delta_i} \subseteq g_i$ for $i = 1, \ldots, t$. Again we often write $SS_p(g)$ or $SS_p$ instead of $SS_p(L, g)$. One can show that $SS_p \subseteq GS_p$ and $SS_p$ is a normal subgroup of $S_p(L)$.

We define a third congruence subgroup, $\Omega(L, g)$, to be the mixed commutator subgroup $[S_p(L), G_S_p(g)]$. It is a normal subgroup of $S_p(L)$ and the formula

$$\sigma \rho \sigma^{-1} \rho^{-1} - 1_L = (\sigma \rho \sigma^{-1} - \delta_L) \rho^{-1} + \delta_L(\rho^{-1} - \delta_L)$$

shows that $\Omega(g) \subseteq SS_p(g)$.

If $t = 1$, our notations of order and of congruence subgroups coincide with those of Klingenberg in [3]. If $\mathcal{O}$ is a complete discrete valuation ring, our notations coincide with those of Riehm in [4].

We use $\subset$ to denote strict set inclusion.

Fix a tableau $g$. Let $I = \{i \mid 1 \leq i \leq t \text{ and } g_{ii} \subseteq \mathcal{O}\}$. Define a relation $\sim$ on $I$ by putting $i \sim j$ if either $i = j$ or $g_{ij} \subseteq (s_i : s_j)$ when $i \neq j$. By (3.8), $\sim$ is symmetric. Let $i, j,$ and $k$ be distinct elements of $I$. If $i \sim j$ and $j \sim k$, then by (3.6), $g_{ik} \subseteq (s_i : s_j) g_{jk} \subseteq (s_i : s_j) (s_j : s_k) \subseteq (s_i : s_k)$, so $\sim$ is transitive. Hence $\sim$ is an equivalent relation on $I$. Let $r$ be the number of the equivalent classes of $\sim$ on $I$.

Let $\sigma$ in $G_{S_p}(g)$. It follows from (3.2) that for each $i$ in $I(\sigma - \delta_L) L^{\delta_i} \subseteq g_i$.
for one and only one value of $\delta = \pm 1$; define $\Delta_i(\sigma)$ to be this value of $\delta$. It is easy to see that $\Delta: GS_p \to Z_2$ is a surjective homomorphism. It is also easy to see that $\Delta_i(SS_p) = 1$ for all $i$ in $I$. Thus $SS_p$ is the kernel of the mapping $\Delta: GS_p \to Z_2$ defined by $\Delta(\sigma) = (\Delta_i(\sigma))_{i \in I}$.

**Theorem 4.1.** (i) $GS_p/SS_p = \bigoplus Z_2$

(ii) Every subgroup of $GS_p$ which contains $SS_p$ is a normal subgroup of $S_p(L)$.

**Proof.** We may assume that $g$ is not the maximum tableau. Then the set $I$ is nonempty. We need only find the image of $\Delta$ to prove (i).

Suppose that $A_i(\sigma) = -\Delta_i(\sigma)$ with $i \neq j$ in $I$. By 3.1, $i \neq j$, the image of $A_i$ is contained in $\bigoplus Z_2$. It suffices to find, for each $i$ in $I$, a mapping $\sigma$ in $GS_p$ such that $\Delta_i(\sigma) = -1$ for all $j \sim i$ and $\Delta_i(\sigma) = 1$ otherwise.

Take a canonical splitting (1.1) and define for each $i$ in $I$.

$$L(i) = \prod_{j \sim i} L_j, \quad \text{and} \quad \sigma = -1_{L(i)} \perp 1_{L(i)}.$$

Trivially $\sigma$ is in $S_p(L)$. It is easy to check that $\sigma$ has the required property. This shows (i).

Let $G$ be a group between $GS_p$ and $SS_p$. If $\sigma$ is in $S_p(L)$ and $T$ is in $G$, it follows from (4.1) that $\Delta_i(\sigma T \sigma^{-1}) = \Delta_i(T)$ for all $i$ in $I$, whence $\Delta(GS_p) = \Delta(G)$. Since $\Delta(GS_p)$ is an abelian group and $G$ contains the kernel of $\Delta$, so $G$ is a normal subgroup of $S_p(L)$. Q.E.D.

5. **Transvections**

Let $y$ and $z$ be vectors of $L$ with $yz = 0$ and let $\lambda$ in $F$. If

$$T_{y,z,\lambda}(x) = x + \lambda(yx)z + \lambda(xz)y$$

for $x$ in $L$ is in $L$, $T_{y,z,\lambda}$ is an element of $S_p(L)$ and is called a quasi transvection. A quasi-transvection is called a $(k, l)$-transvection if it is of the form $T_{y,z,\lambda}$ with the following properties: (i) there is a canonical basis containing both $y$ and $z$; (ii) $y$ is $k$-pure and $z$ is $l$-pure; (iii) $y = z$ if $k = l$. Whenever we refer to the fact that $T_{y,z,\lambda}$ is $(k, l)$-pure, it will be implicit in this statement that $y$ and $z$ have these properties unless mentioned otherwise.

The following identities are easy to verify:

$$T_{y,z,\lambda} = T_{z,y,\lambda}; \quad T_{y,z,\lambda} = T_{y,z,\mu}; \quad T_{y,z,\lambda} = T_{y,z,\mu}; \quad T_{y,z,\lambda} = T_{y,z,\nu}; \quad T_{y,z,\lambda} = T_{y,z,\nu}, \quad n \in Z;$$

$$T_{u,z,\lambda} T_{x,w,\mu} = T_{x,w,\mu} T_{u,z,\lambda}.$$
if $x, y, z$ and are mutually orthogonal;

$$\sigma T_{y,z,x} \sigma^{-1} = T_{cy,az,x} \quad \text{if} \quad \sigma \in S_p(L). \quad (5.4)$$

$T_{y,v,z}$ will usually be written as $\tau_{y,2z}$. Thus $\tau_{y,2z}(x) = x + \lambda(yx)y$. One can easily check the following:

$$T_{z,z,x} T_{z,v,z} = \tau_{z,2(yz)} T_{x+y,z} \quad (5.5)$$

$$T_{x,z,x} T_{x,v,z} = T_{x+y,z} \quad \text{if} \quad xy = 0, \quad (5.6)$$

$$T_{y,z,x} = \tau_{y,2(z+y)}^{-1} T_{x+y,z} \quad \text{if} \quad xy = 0. \quad (5.7)$$

$$\tau_{y,z,x} = \tau_{x,2(yx+y)}^{-1} T_{x+y,z} \quad \text{if} \quad xy = 0. \quad (5.8)$$

We simply call an $(i,j)$-pure quasitransvection ($i$-pure transvection $(i, j)$-transvection ($i$-transvection, respectively).

**Lemma 5.1.** Let $T = T_{y,z,x}$ be an $(i,j)$-transvection. Then for $k = 1,..., t$, $O_k(T) = ((T - 1_L)L^{s_k}, \lambda L^{s_k})$, i.e., $T$ is in $SS_p(O(T))$. Also $O_k(T) = [\lambda s_i]$, and if $g$ is a tableau, then $O(T) \leq g$ if and only if $O_{ij}(T) = [\lambda s] \subseteq g_{ij}$.

**Proof.** The first part of the lemma follows from 2.1, (3.3) and definitions.

It is easy to see that $(T - 1_L)L^{s_k} = \lambda(yL^{s_k})z + \lambda(zL^{s_k})y$. By 2.3, we get $O_i(T) = [\lambda s_i]$. Also $O_k(T) = \lambda((s_k : s_i)s_i)L^{s_i} + \lambda((s_k : s_i)s_i)L^{s_i}$. Suppose $[\lambda s_i] \subseteq g_{ij}$ and put $g_k = (g_{k1}, ..., g_{kt})$ for $k = 1,..., t$. By (3.8), $[\lambda s_i] \subseteq g_{ij}$ if and only if $[\lambda s_i] \subseteq g_{ij}$. Hence

$$O_k(T) \subseteq (s_k : s_i)g_{ij}L^{s_i} + (s_k : s_i)g_{ij}L^{s_i}$$

$$\subseteq (s_k : s_i)g_i + (s_k : s_i)g_j$$

This finishes the proof. Q.E.D.

6. GENERATION OF $SS_p$ BY TRANSVECTIONS

Fix a tableau $g = (g_i) = g_{ij}$.

Let $(x,y)$ be a pair of $i$-pure vectors with $[xy] = [s_i]$ and let $\alpha$ be a unit in $\mathfrak{c}$. We use $P_{(x,y)}(\alpha)$ to denote the mapping which carries $x$ to $\alpha x$, $y$ to $\alpha^{-1}y$ and $z$ to $z$ for $z$ in $(x,y)^0$. We write $P_{\alpha}$ instead of $P_{(x,y)}(\alpha)$. 
LEMMA 6.1. Let \( \{x, y\} \) be two i-pure vectors with \( xy = \mu s_i \), \( \mu \in U \). Let \( \alpha = 1 + \beta \) be a unit. Then

\[
P_{(x,y)}(\alpha) = \tau_{x,\mu^{-1}} \tau_{y,\mu^{-1}} \tau_{x+y,\mu^{-1}}.
\]

Proof. Straight forward computation.

LEMMA 6.2. Let \( \sigma \in GS_{\mu} \) and let \( x \) be an i-pure vector in \( L \). Then there exists a canonical basis \( \mathcal{B} = \{x_i\}_{i=1}^n \) of \( L \) such that \( x_1 = x \) and \( \sigma(x_1) \cdot x_2 \) is a unit congruent to 1 modulo \( g_{ii} \).

Proof. Let \( \mathcal{B}' = \{y_i\}_{i=1}^n \) be a canonical basis of \( L \) with \( y_1 = x, y_j \) is pure of type \( \nu(j) \) for \( j = 1, \ldots, n \) and \( y_{2j-1} y_{2j} = s_{\nu(2j)} \) for \( j = 1, \ldots, n/2 \). Write \( \sigma(x) = \sum_{i=1}^n \beta_i y_i \) and \( \sigma(y_1) = \sum_{i=1}^n \gamma_i y_i \), where \( \beta_1 = \delta, \gamma_2 = \delta \in g_{ii} \) and \( \beta_j, \gamma_j \in g_{ii}(\nu) \) for all \( j \neq 1, 2 \). Let \( \gamma_1 = y_1 y_2 = (\sigma y_1)(\sigma y_2) = \sum (\beta_{2j-1} y_{2j} - \beta_{2j} y_{2j-1}) s_{\nu(2j)} \).

Then \( \gamma_1 \) is in \( \mathcal{B} \) and \( \gamma_1 + \gamma_2 = 0 \). There exists by 1.11 some unit \( a = \beta_1 + dr \) with \( d \) in \( \mathcal{O} \) and \( dr \) in \( g_{ii} \) by (3.7). For each \( j > 1 \) we define \( x_{2j-1} = y_{2j-1} - d \beta_{2j-1} s_{\nu(2j)} s_{\nu(2j-1)} y_1 \) and \( x_{2j} = y_{2j} + d \beta_{2j-1} s_{\nu(2j)} s_{\nu(2j-1)} y_1 \). It follows from 1.4 and (3.7) that \( x_j \) is \( \nu(j) \)-pure for \( j > 1 \). Define \( x_1 = y_1 \) and \( x_2 = 2d s_{\nu(1)} s_{\nu(2)} + \sum_{j=2}^{n/2} (\beta_{2j-1} y_{2j-1} - \beta_{2j} y_{2j-1}) s_{\nu(2j)} \).

It can be shown that \( x_1, x_2 \) are i-pure. Furthermore one also can show that \( \mathcal{B} = \{x_i\}_{i=1}^n \) is a canonical basis of \( L \) and \( (\sigma x) x_2 = a \) which is a unit with \( a = \delta \) (mod \( g_{ii} \)). Q.E.D.

LEMMA 6.3. Let \( x, y \) be two i-pure vectors with \( x - y \in g_{ii} \). Then there exists a product of pure transvections, each in \( SS_{\mu} \), which carries \( x \) to \( y \). Moreover the types of the pure transvections involved are all of the form \( (i, j), \) \( i < j \leq t \).

Proof. If \( g_{ii} = \mathcal{O} \) then the lemma follows from the proof of 2.1. We may assume that \( g_{ii} \subseteq \mathcal{O} \). Choose a canonical splitting \( L = \bigcup_{1 \leq j \leq t} L_j \) with \( x \in L_i \), and write \( y - x = \sum_{j=1}^t r_j z_j \), where \( z_j \) is a maximal vector in \( L_i \) and \( r_j \) in \( g_{ij} \), for \( j = 1, \ldots, t \). We shall first find a product \( \rho \) of the required kind which carries \( x \) to \( x + r_j z_j \) and is identity on \( L_j \), \( j \neq i \). If \( [z_j] = [s_i] \), \( \rho \) can be chosen as the transvection \( \tau_{z_j, r_j, [s_i]^{-1}} \). Assume that \( [z_j] \subseteq [s_i] \). By 6.2, we may assume that there exists i-pure vector \( \omega \) in \( L_i \) such that \( x \omega = s_i \) and...
\[ \alpha = (ax) \omega \xi^{-1} \] is a unit. Write \( z_i = ax + bw + z_i \), where \( z \) in \( \langle x, w \rangle^0 \), \( a, b \) in \( 0 \) and \( \alpha = 1 + r, a \). Let \( \rho = T_{w, z, -s, -t, \sigma -1, \tau w, \xi^{-1} \tau, \alpha - t} P_{(x, w)}(\alpha) \). Then it follows from (5.7), 6.1 and straightforward computation that \( \rho \) is the required map.

Next suppose that \( j \neq i \). We shall find a transvection \( T_j \) of the required type which carries \( x \) to \( x + r, z_j \) and leaves \( L, h \neq i, j \) fixed. Choose \( \omega \) as above, define \( T_j = T_{z_j, \omega, (xw) -1, t} \), we have \( T_j \) in \( SS_p \) by 5.1. The product \( \rho \prod_{j \neq i} T_j \) satisfies the condition of the lemma.

**Lemma 6.4.** In addition to the hypothesis of 6.3, suppose that \( \omega \) is an \( i \)-pure vector with \( [xw] = [s] \) and \( xw = yw \). Then the product described in 6.3 can be chosen to have the additional property that it leaves \( \omega \) fixed.

**Proof.** Suppose \( g_{ii} \subseteq \mathcal{C} \), \( \mathcal{C} x + \mathcal{C} w \) splits \( L \) by 1.2, and we may assume that both \( x \) and \( w \) are in \( L_i, w z_j = 0 \) for \( j = 1, \ldots, t \) since \( xw = yw \). The proof of 6.3 can be applied here almost word for word, once the \( \omega \) in 6.3 has been replaced by the \( \omega \) here.

Suppose \( g_{ii} = \emptyset \). If \( [xy] = [s] \), then we choose \( \rho = \tau y - x, (xw) -1 \). If \( [xy] \subseteq [s] \), there exists, by 1.11, some \( a \) in \( \mathcal{C} \) with \( (xy) + a(yw) \) being a unit. We choose \( \rho = \tau y - x, (y (x + aw)) -1, w, a (xw)) -1 \). Then \( \rho \) is the required map.

**Theorem 6.5.** \( SS_p \) is generated by the pure transvections contained in it.

**Proof.** Let \( \sigma \) in \( SS_p \) and suppose that \( 0 \neq L = (ax + \sigma y) \perp M \) with \( xy = s_1 \). By the definition of \( SS_p \), \( ax - x \) and \( \sigma y - y \) are in \( g_{11} = (g_{11}, \ldots, g_{1s}) \). By 6.3, there is a product \( \rho_1 \) of transvections, each in \( SS_p \), which carries \( x \) to \( ax \) and let say \( y \) to \( y' \). Therefore \( y' - y \) is in \( g_1 \) and so \( y' - \sigma y = (y' - y) - (\sigma y - y) \) is also in \( g_1 \). Since \( (ax) y' = xy = (ax)(\sigma y) \), by 6.4, there exists another product \( \rho_2 \) which carries \( y' \) to \( \sigma y \) and leaves \( ax \) fixed. Then \( \rho = \rho_2 \rho_1 \) carries \( x \) to \( ax \) and \( y \) to \( \sigma y \). Thus \( \tau = \rho^{-1} \sigma \) in \( SS_p \) and is identity on \( \mathcal{C} x + \mathcal{C} y \). Now the theorem follows from the induction on \( \dim L \).

By 5.1, (5.7), and 6.5, we get:

**Corollary 6.6.** \( S_p(L) \) is generated by pure single transvections.

### 7. The Determination of Normal Subgroups

From now on the assumption \( \dim L_i \geq 4 \) for \( i = 1, \ldots, t \), is in force, let \( g = (g_i) = g_{ij} \) be a fixed tableau where \( g_i = (g_{i1}, \ldots, g_{is}) \) is an invariant lattice for \( i = 1, \ldots, t \).

Let \( S \) be a nonempty subset of \( S_p(L) \). We define \( G(S) \) to be the smallest normal subgroup of \( S_p(L) \) containing \( S \). If \( S = \{ \sigma \} \) we write \( G(\sigma) \) instead of \( G(\{\sigma\}) \).
Lemma 7.1. Let \( x \) be an \( i \)-pure vector and \( \sigma := \tau_{x,u} \) in \( S_\nu(L) \). Then \( G(\sigma) \) contains all pure single transvections \( \tau_{y,v} \) with \( O(\tau_{y,v}) \leq O(\sigma) \).

Proof. (1) Suppose \( y \) is \( i \)-pure and \( v = \mu \). Choose \( \tau \) in \( S_\nu(L) \) with \( \tau(x) = y \), then \( \tau_{y,v} = \tau \sigma \tau^{-1} \in G(\sigma) \).

(2) Suppose \( y \) is \( i \)-pure and \( v = \mu x^2 \) for some \( \alpha \) in \( C \). Choose an \( i \)-pure vector \( z \) in \( L \) with \( yz = 0 \). Then \( \tau_{y,v} = \tau_{x,y,v} = \tau_{2,y,z} \tau_{y,z} \tau_{x,y} \tau_{x,z} \tau_{y,x} \) by (5.8). By (1), \( \tau_{y,v} \) is in \( G(\sigma) \).

(3) Suppose \( y \) is \( s \)-pure and \( v = p \) for some \( p \) in \( L^* \). Choose an \( i \)-pure vector \( x \) in \( L \) with \( xy \neq 0 \). Then \( \tau_{y,v} = \tau_{x,y,v} = \tau_{y,z} \tau_{y,z} \tau_{x,y} \) by (5.8). \( \tau_{y,v} \) is in \( G(\sigma) \) by (1).

Q.E.D.

Lemma 7.2. Let \( \{x, y\} \) be two \( i \)-pure vectors with \( [xy] = [s_i] \). Let \( \sigma := \tau_{x,u} \tau_{y,v}, \) where \( \tau_{x,u} \) and \( \tau_{y,v} \) are in \( S_\nu(L) \). Then \( G(\sigma) \) contains all \( j \)-pure transvections \( \tau_{z,\lambda} \) with \( O(\tau_{z,\lambda}) \leq j \).

Proof. Let \( \{x', y'\} \) be two \( i \)-pure vectors in \( \langle x, y \rangle \) with \( [x'y'] = [s_i] \). By 6.4, there exists \( \tau_1 \) in \( S_\nu(L) \) with \( \tau_1(x) = y \). Consider \( \sigma_1 = \sigma^{-1} \tau_1 \sigma \tau_{1}^{-1} = \tau_{y,u} \tau_{y,v} = \tau_{z,\lambda} \). By 2.1 again, there exists \( \tau_2 \) in \( S_\nu(L) \) with \( \tau_2(y) = y' \). Consider \( \sigma_2 = \sigma(y, \tau_2^{-1}) = \tau_{x,u} \tau_{y,v} \tau_{y,v}^{-1} \) in \( G(\sigma) \), \( \tau_{2,z} \tau_{y,v} \tau_{y,v}^{-1} \) in \( G(\sigma) \) for all units \( \alpha, \beta \). Choose a unit \( \beta \) such that \( \beta^2 - 1 \) is still a unit. Then \( \tau_{x,u} \tau_{y,v} \tau_{y,v}^{-1} = \tau_{y,1} \) by 7.1, \( \tau_{y,v} \) in \( G(\sigma) \), and so \( \tau_{z,\lambda} \) in \( G(\sigma) \). Now result follows from 7.1.

Q.E.D.

Proposition 7.3. Let \( P = \langle x, y \rangle \) be a plane spanned by two \( i \)-pure vectors \( x, y \) with \( [xy] = [s_i] \). Assume that \( a \) is in \( S_\nu(L) \) with \( a \mid P^0 = \) identity. Then \( G(\sigma) \) contains all \( i \)-single transvections \( \tau_{z,\lambda} \) with \( O(\tau_{z,\lambda}) \leq O(\sigma) \).

Proof. Let \( \alpha x = ax + by \). By a change of basis and 1.11, we may assume without loss of generality that \( a \) is a unit. Let \( \tau_{x,-a^{-1}b} \sigma(\alpha) = ax \). Let \( \tau_{y,-a^{-1}b} \sigma \) and \( \sigma(\alpha) = cx + a^{-1}y \). Consider \( \sigma_2 = \tau_{x,u} \sigma \). Let \( f_\sigma \) be the mapping that carries \( x \) to \( \nu x \) and \( y \) to \( \nu^{-1}y \) for unit \( \nu \). It is easy to see that \( \sigma_2 = f_\sigma \). Hence \( \sigma = \tau_{x,-c} \sigma_2 \). Let \( \tau = \tau_{x,\nu} \sigma_2 \) in \( G(\sigma) \). \( \tau = \tau_{x,\nu} \sigma_2 \) in \( G(\sigma) \).

If \( O_{\nu}(\sigma) = [b] + [d] \) then we are done by 7.2. If \( O_{\nu}(\sigma) \neq [b] + [d] \) then
Let $\alpha \in S_0(L)$ with $O(\alpha) \subseteq \mathfrak{g}$. Then there exists finitely many $i$-pure vectors $x_k$ in $L$ and $\sigma_k$ in $G(\alpha)$ such that $\sum_k [\sigma_k(x_k) \cdot x_k] = \sigma g_{\mathfrak{sl}}$ for $i = 1, \ldots, t$.

Proof. Fix an $i$ between 1 and $t$. If $g_{ii} = 0$ then by a change of basis the result follows easily. We may assume that $g_{ii} \subseteq \mathfrak{g}$. It follows from the definitions of order and tableau that $g_{ii}$ is finitely generated. It will be sufficient to show that for an element $\mu$ in a minimal set of generators of $g_{ii}$, there exists an $i$-pure vector $x$ and some $\rho$ in $G(\alpha)$ such that $[\rho(x) \cdot x] \supseteq [\mu \mathfrak{s}_i]$. By 2.3 there exists $y$ in $L^{g_i}$ such that $((\sigma - \delta_L)y) \supseteq [\mu \mathfrak{s}_i]$ where $\delta = \pm 1$. Splitting $y$ into components with respect to a canonical splitting, we may assume that $y$ is $i$-pure. Let $[yz] = [s_i]$ with $z$ in $L^{g_i}$. Since $((\sigma y)y) = ((\sigma - \delta_L)y)y$, we see that $(\sigma y)y$ is in $s_i g_{\mathfrak{sl}}$. If $[(\sigma y)y] \supseteq [\mu \mathfrak{s}_i]$, we are finished, so assume

$[\sigma y]y \subseteq [\mu \mathfrak{s}_i].$ (7.1)

Let $w \in L^{g_i}$ satisfy $[((\sigma - \delta_L)y)w] \supseteq [s_i \mu]$ and define $x$ as follows: put $x = z$ if $[((\sigma - \delta_L)y)z] \supseteq [s_i \mu]$; otherwise put $x = w$ if $[yw] = [s_i]$ or $x = z \cdot w$ if $[yw] \subseteq [s_i]$. Therefore $[((\sigma - \delta_L)y)x] \supseteq [\mu \mathfrak{s}_i]$ and $[xy] = [s_i]$, whence (7.2) $(\sigma y)x = (yx)(\delta + \gamma)$, $[\gamma] \supseteq [\mu]$. Now put $\lambda = (yx)^{-1}$ and define $\rho = \sigma \tau_{y, \lambda} \sigma^{-1} \tau_{y, \lambda} = \tau_{y, \lambda} \tau_{y, \lambda}^{-1}$ in $G(\alpha)$. A direct computation using (7.1) and (7.2) shows that $[(\rho x)x] \supseteq [\mu \mathfrak{s}_i]$ as required.

Q.E.D.

Lemma 7.5. Suppose that $\mathfrak{g}$ is in $S_\mathfrak{g}(L)$. Then $G(\mathfrak{g})$ contains all $i$-transvections $\tau$ with $O_{\mathfrak{sl}}(\tau) \subseteq O_{\mathfrak{sl}}(\mathfrak{g})$ for $i = 1, \ldots, t$.

Proof. (1) Assume $O_{\mathfrak{sl}}(\mathfrak{g}) = \emptyset$. There exists $i$-pure vector $x \in L$ with $x L^{g_i} = ((\sigma - 1_L)x)L^{g_i} = [s_i]$, and $\{x, (\sigma - 1_L)x\}$ is linearly independent. If $[sx \cdot x] = [s_i]$, apply 7.2. If $[sx, x] \subset [s_i]$ we can find an $i$-pure vector $y$ with $yxy = 0$ and $yx = s_i$, as can be seen by constructing a canonical basis containing $x$ as basic vector. Let $\tau = \tau_{y, \sigma, s_i}^{-1}$ and $\rho = \sigma \tau_{y, \lambda} \sigma^{-1} = \tau_{g_i, \lambda} \tau_{g_i, \lambda}^{-1}$ in $G(\mathfrak{g})$. A direct computation using (7.1) and (7.2) shows that $[\rho x] = [\mu \mathfrak{s}_i]$ and we are finished.

(2) Assume $O_{\mathfrak{sl}}(\mathfrak{g}) \subset \emptyset$. Let $\mu$ be an element of a minimal set of generators of $O_{\mathfrak{sl}}(\mathfrak{g})$. By 7.4 we can find an $i$-pure vector $x$ and $\rho$ in $G(\mathfrak{g})$ with $[\rho x] = [\mu \mathfrak{s}_i]$. It follows from 6.2 that there exists an $i$-pure vector $y$ in $L$ such that $xy = s_i$ and $s_i^{-1} y = a$ is a unit. Since $\sigma \rho = ((\rho x)x) \rho = \emptyset$. By 1.11, there exists some unit $\lambda$ such that $a + \lambda((\rho x)x) s_i = b$ is a unit.
Let $\tau = \tau_x,\lambda x^{-1}$ and let $\sigma' = \rho^{-1}\tau\rho^{-1}\tau \in G(\sigma)$. It follows that $\sigma'x = x + \lambda(x(\rho x))(\rho^{-1}x)$ and $\sigma'y = (y - \lambda x) + \lambda s^{-1}(x(\rho y - \lambda(\rho x))(\rho^{-1}y)) = (y - x) + b(\rho^{-1}x)$.

Put $v = s^{-1}(x(\rho x))(\rho b)^{-1}, [v] = [(\rho x)x]$. Let $z = x + vy$. Then $z$ is an $i$-pure vector and $\sigma'z = (1 - \lambda x)z + \lambda s^{-1}$.

It follows from 7.3 that $G(\sigma)$ contains all pure transvections $\tau_{x,\lambda}$ with $O_{\mu}(\tau_{v,\lambda}) \subseteq [v] = [\mu]$. Q.E.D.

**Lemma 7.6.** Let $T = T_{x,z,\lambda}$ be an $i,j$-pure transvection in $S_p(L)$ with $O(T_{x,z,\lambda}) = g$ and $i \neq j$. Then $G(T)$ contains all $(k,l)$-pure transvections $T'$ with $O(T') \leq g$.

Proof. (1) Let $T' = T_{x,w,\alpha}$ be an $i,j$-pure transvection in $S_p(L)$ with $\alpha$ being a unit. Choose $\tau$ in $S_p(L)$ with $T(y) = x$ and $T(z) = x\alpha$, then $T' = T\tau^{-1}$ is in $G(T)$.

(2) Let $T' = T_{x,w,\alpha}$ be an $i,j$-transvection with $\alpha$ in $\mathcal{C}$. Choose $j$-pure vector $w'$ in $L$ with $xw' = 0 = ww'$. Then $T' = T_{x,w',\alpha}$ is in $G(\sigma)$ by (1).

(3) Let $T'' = T_{x,w',\lambda}$ be a $(k,l)$-transvection. By 5.1, $g_{ik} \cdot s_k = ((s_i : s_j) s_j(s_k : s_j) s_j + (s_i : s_j) s_j(s_k : s_i) s_j).$ Suppose $\lambda s_k \subseteq \mu(s_i : s_j) s_j(s_k : s_i) s_j$, $s_k = \mu(\alpha(s_i : s_j) s_j(s_k : s_i) s_j)$ where $x \in (s_i : s_j)$ and $y \in (s_k : s_i)$. Let $a = o_{\alpha(s_i : s_j)}$ and $b = \beta(s_k : s_i)$ and follows that $a, b$ is in $\mathcal{C}$, $\alpha x$ in $L^\mathcal{C}$ and $\alpha y$ in $L^{\mathcal{C}}$. Choose $i$-pure vector $x', j$-pure vector $w'$ such that $x, w, x', w'$ are mutually orthogonal. Then $x' + ax = x''$ is $i$-pure and $w' + bw = w''$ is $j$-pure. By (5.6) $T'' = T_{x'',w'',\alpha}T_{x'',w'',\alpha}T_{x'',w'',\alpha}T_{x'',w'',\alpha}$ is in $G(\sigma)$ by (1) and (2). The case $x, w, x', w'$ is done similarly. The result now follows from (5.2).

Q.E.D.

**Lemma 7.7.** Let $\{x,y\}$ be $i$-pure vectors with $[xy] = [s_j]$ and let $T = T_{x,z,\lambda}$ in $S_p(L)$. Then $G(\sigma)$ contains all $(i,j)$-pure transvections which have order $\leq O(\sigma)$, $j = 1, \ldots, t$. In particular if $\{x_i\}$ is a canonical basis for $\langle x,y \rangle^0$ and $w = \sum_1^t x_i, z = \sum_1^t \beta x_i$, then all the transvections $T_{x,v,\alpha}$ and $T_{z,\alpha}z,\beta$ are in $G(\sigma)$.

Proof. We may assume that $z$ and $w$ are in $\langle x, y \rangle^0$.

(1) First we show that $T_{x,z,\alpha} \in G(T)$ if $z$ is a pure vector. Let $\tau_1$ be the mapping which sends $x$ to $x$, $y$ to $x + y$ and $v$ to $y$ for all $v$ in $\langle x, y \rangle^0$. It is easy to see that $\tau_1$ is in $S_p(L)$. Let $\sigma_1 = T^{-1}_1T^{-1}_1 = T_{x+y,z,-\mu}T_{y,z,\mu}$ in $G(\sigma)$. One can easily check that $O(\tau_{z,\mu}^2(x,y)) \leq O(\sigma), \tau_{z,\mu}^2(x,y)$ is in $G(\sigma)$ by 7.5. It follows from (5.5) that $T_{x,z,\mu}^2(x,y) \sigma_1 = T_{x,y,\mu}T_{y,z,\mu}T_{y,z,\mu}$ is in $G(\sigma)$.

(2) Let $\{x_i\}$ be a canonical basis for $\langle x, y \rangle^0$. If $w = \sum \alpha x_i$ then
\( T_{x,w,u} \) is the product of the transvections \( \sum u = T_{x,\alpha_x^2u-1^2x_{2u-1}^2+2x_{2u}} \). Similarly if \( x = \sum \beta u x_u \) then \( T_{x,\sum \alpha x_u} \) is the product of \( T_u = T_{y,\beta u^2y_{2u-1}^2x_{u-1}^2+3y u_{2u}^2} \). Let \( \tau_u \) be the mapping: \( \tau_u(x_{2u}) = x_{2u} + x_{2u-1} \) and \( \tau_u(x_v) = x_v \) for all \( u \neq 2u \). We consider \( \rho_u = \sum v (\tau_u T_{r}^{-1} T^{-1}) \sum u \in G(\sigma) \). It follows from (5.5) and direct computation that \( \rho_u = \tau_1 T_{x,\alpha x_{2u}, x_{2u-1}^2+2x_{2u}^2} \tau_1 T_{y,\beta y_{2u}, y_{2u-1}^2+2y_{2u}^2} \). Let \( \tau_1 \) and \( \tau_2 \) be the mapping: \( \tau_1(x_{2u}) = x_{2u} + x_{2u-1} \) and \( \tau_2(x_{2u}) = x_{2u}^2 + x_{2u} \).

By Lemmas 5.1, 5.2, and 7.5, \( \tau_1 \) and \( \tau_2 \) are in \( G(\sigma) \), so is

\[ T_{x,\alpha x_{2u}, x_{2u-1}^2+2x_{2u}^2} T_{y,\beta y_{2u}, y_{2u-1}^2+2y_{2u}^2} \].

Now we are done by (1).

PROPOSITION 7.8. If \( \sigma \in S_p(L) \) with \( O(\sigma) = g \), then \( G(\sigma) \) contains \( SS_p(g) \).

**Proof.** (1) By 6.5, it suffices to show that \( G(\sigma) \) contains all pure transvections with order \( \lesssim g \). By 7.5, we only have to show that \( G(\sigma) \) contains all \( (i, j) \)-pure transvections with order \( \lesssim g \) for all \( i \neq j \). Fix \( i \) and \( j \). If \( g_{ij} = 0 \) then \( G(\sigma) \) contains all \( (i, j) \) transvections with order \( \lesssim g \) by (5.7), (2.1), and 2.1. We assume that \( g_{ij} \subseteq G \). Let \( \mu \) be an element in a minimal set of generators of \( g_{ij} \). By 2.3 there exists \( x \in L^s \) such that \( [\mu s] \subseteq ((\sigma - \delta x) x \in L^s \). Splitting \( x \) into components with respect to a canonical basis, we see that we may suppose that \( x \) is pure of type \( i \). By 6.2 there exists an \( i \)-pure vector \( y \) in \( L \) with \( xy = s_i \) and \( \alpha = (\sigma x) \cdot y \) a unit.

(2) Let us write \( \sigma \) as a product of certain standard mappings. Let \( \sigma_x = \alpha P_{x,y}(x^{-1}) \), then \( (\sigma_x x) y = x y \), \( \sigma_x x = x + \beta y + z \) for some \( \beta \in g_{ij} \) and \( z \in \langle x, y \rangle \). Let \( T_{x} = T_{y,x,-x^{-1}} \) and \( \sigma_x = T_{x}^{-1} \sigma_1 \). Then \( \sigma_2(y) = \gamma x + y + w \) with \( \gamma \) in \( g_{ij} \) and \( w \in \langle x, y \rangle \). Let \( T_{x} = T_{x,w,-w^{-1}} \) and \( \rho = T_{x}^{-1} \sigma_2 \). Then \( \rho(x) = x \), \( \rho(y) = y \) and \( \rho(\langle x, y \rangle) = \langle x, y \rangle \).

(3) Let \( \tau = T_{x}(P_{-1}-\alpha P_{-1}^{-1}) T_{x}^{-1} = T_{x,z,-z^{-1}} T_{x,w,-w^{-1}} \in G(\sigma) \). Thus if \( O_{ij}(\tau) \supseteq [\mu] \) the proposition follows by 7.8. So assume \( O_{ij}(\tau) \not\subseteq [\mu] \). By 7.8 \( T_{y} T_{x} \in G(\sigma) \) whence \( \sigma = \rho \rho_{a} \) is in \( G(\sigma) \), and \( O_{ij}(T_{y} T_{x}) \subseteq [\mu] \). Hence \( ((\alpha - \delta x) x \in L^s \supseteq [\mu s] \) by virtue of the identity

\[ \alpha \pm \delta L = T_{y} T_{x}(\sigma \pm \delta L) + \delta L(T_{y} T_{x} - 1 L), \]

the choice of \( x \) and the condition \( O_{ij}(T_{y} T_{x}) \subseteq [\mu] \).

(4) Let \( \mu = \alpha - \delta ; G(\sigma) \) contains all pure transvections \( \tau \) with \( O_{ij}(\tau) \subseteq [\mu] \) by 7.6. Let \( \beta \) be an element in a minimal set of generators of \( s_i, s_j \), and let \( u \) and \( v \) be pure vectors of type \( i \) and \( j \) in \( \langle x \rangle \), respectively. It is easy
to see that $T_{x,c,u \partial s_i^{-1}} = T_{x,y \cdot \partial r, u \partial s_i^{-1}} T_{x,y, -u \partial s_i^{-1}}$ is in $G(\sigma)$ by (5.7). Now the proposition follows from 7.7.

Q.E.D.

**Theorem 7.9.** Let $G$ be a subgroup of $S_p(L)$ with $O(G) = g$. Then $G$ is a normal subgroup of $S_p(L)$ if and only if $SS_p(g) \subseteq G \subseteq GS_p(g)$.

**Proof.** Necessity follows from 7.5 and 7.8. Sufficiency follows from 4.1, 5.1 and 6.3. Q.E.D.

**Corollary 7.10.** Let $G$ be a normal subgroup of order $g$. Then the mixed commutator subgroup $[S_p(L), G]$ is equal to $SS_p(g)$. In particular $S_p(L)$ equals its own derived group $\Omega(S_p(L))$.

**Proof.** $[S_p(L), G] \subseteq SS_p(g)$ is obvious.

Conversely given a pure transvection $T$ in $G$, it is easy to find some $\rho$ in $S_p(L)$ such that $O(\rho T_{p^{-1} T^{-1}}) = O(T)$ — for example by using (5.4), (5.6), and 5.1. Then the result follows immediately from 7.9. Q.E.D.

**References**