THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

LINEAR GROUPS OVER LOCAL RINGS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

BY

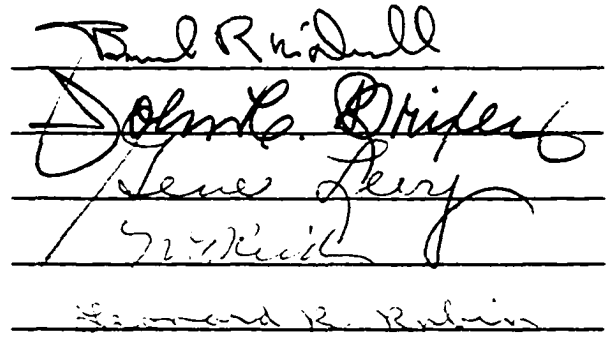
JAMES CHARLES POMFRET

Norman, Oklahoma

1971

LINEAR GROUPS OVER LOCAL RINGS

APPROVED BY



DISSERTATION COMMITTEE

ACKNOWLEDGMENTS

I wish to thank my thesis advisor, Bernard McDonald, who suggested the topic and provided guidance and encouragement throughout the course of the research.

I also thank my wife, Penny, for her help in editing, and my daughter, Deborah, for her understanding.

TABLE OF CONTENTS

INTRODUCTION

The first book devoted to linear groups was Jordan's Traité des Substitutions. In this work is found the famous Jordan Normal Form, results on substitution groups, and a study of quadratic forms. Perhaps the most important results are those dealing with Galois' problem of determining all solvable groups of substitutions. Dickson, in Linear Groups [10], investigates linear groups over Galois fields and derives an extensive collection of finite simple groups. More recently D. Suprenenko, in Soluble and Nilpotent Linear Groups [27], studies the problem of normal series in linear groups.

In 1938 J. Brenner began a study of the general linear group over the particular local ring Z/Zp^n ([6],[7],[8]). For each $t, 1 \leq t \leq n$, he defined normal subgroups of $GL_n(Z/Zp^n)$,

$$N_t^* = \{A : A \text{ is scalar mod } p^t\}$$

$$N_t = \{A : \det A = 1 \text{ and } A \equiv I \bmod p^t\}.$$

It was then shown that the subgroups are characteristic and that G is a normal subgroup of $GL_n(Z/Zp^n)$ if and only if there is an integer t such that,

$$N_t^* \geq G \geq N_t .$$

In Chapter II we sketch a generalization of Brenner's work, due to Klingenberg [18], for the setting of an arbitrary local ring. Klingenberg

defines for each ideal J of the local ring R the congruence subgroups $GC_n(R,J)$ and $SC_n(R,J)$ that are analogous to N_t^* and N_t, respectively; he shows that a subgroup G is normal if and only if

$$GC_n(R,J) \geq G \geq SC_n(R,J)$$

for some ideal J of R.

We raise the question: For an arbitrary local ring are the congruence subgroups characteristic as was the case in $GL_n(Z/Zp^n)$? A partial answer to this question is given in Chapter II showing that the congruence subgroups for R and the maximal ideal are characteristic. Furthermore, it seems unreasonable to expect the congruence subgroups for any ideal in a local ring to be characteristic; rather, it was conjectured that $GC_n(R,M^i)$ and $SC_n(R,M^i)$ are characteristic where M denotes the maximal ideal and i a non-negative integer.

In Chapter III we determine the structure of the kernels of the natural group morphisms

$$h_i : GL_n(R) \to GL_n(R/M^i), \quad i = 1,2,\cdots .$$

In particular, we generalize portions of recent research by Barbee [3] with regard to basic generators, nilpotency, and the central series. It is also shown that when R is a finite local ring the kernel K_1 of h_1 is characteristic. A useful corollary is that when R is finite the automorphisms of $GL_n(R)$ induce the automorphisms of $GL_n(R/M)$. Since in this case R/M is a finite field and here the automorphisms are know, one is able then to "lift" much of the theory to $GL_n(R)$.

The subgroup of the upper triangular matrices over finite fields has been studied by Weir [28] who showed they form a p-Sylow subgroup of the

general linear group. Weir also computed elementary generators and central series. Robertson [22] considered the subgroup of matrices that are upper triangular modulo the Jacobson radical in an arbitrary ring with 1. He showed the subgroup is a Fitting group when the radical is nilpotent. We consider the subgroup S of matrices that are upper triangular modulo the maximal ideal and determine its elementary generators and central series. We show that when the maximal ideal is nilpotent, S is a nilpotent group and when R is finite of characteristic p^t that S is a p-Sylow subgroup.

In 1928 Schreier and van der Waerden [25] determined the automorphisms of the general linear group over an arbitrary commutative field. Later Dieudonné [11] described the automorphisms of the general linear group over a non-commutative field. Hua and Reiner [16] investigated automorphisms of linear groups over the ring of integers; whereas, Landin and Reiner [19] discussed automorphisms of the general linear group over non-commutative principal ideal domains. Automorphisms of the general linear group over integral domains very recently have been determined separately by T. O'Meara [21] and Yen Shih Chien [8].

Finally Cohn [9] was able to compute the automorphisms of $GL_n(R)$ where R is a k-ring (k a field) with a degree function such that R is a discrete normed ring. The general method in each case is first to determine the images of involutions and then the images of transvections. Once the images of transvections are known, O'Meara invoked the Fundamental Theorem of Projective Geometry in the field of quotients of the integral domain to obtain his results. In our study we have no field of quotients to appeal to and since R is not a domain, it was suspected the

presence of torsion elements would produce new automorphisms.

In determining the automorphisms we follow Yen Shih Chien's method which is highly computational. If R/M has characteristic other than 2 and $n \geq 3$, we show that any automorphism of $GL_n(R)$ can be described in terms of a triple (χ,P,α); χ is a group endomorphism of R^*, P a similarity factor in $GL_n(R)$, and α an automorphism of R. The cases $n = 2$ and characteristic of R/M equal to 2 are very difficult and O'Meara [21] was only able to handle involutions for characteristic 2 over domains by employing separate arguments. Cohn [9] discusses in detail the difficulty of determining the automorphisms of $GL_2(R)$.

In the last section the conjecture posed concerning the characteristic subgroups is answered affirmatively by showing that $GC_n(R,M^i)$ and $SC_n(R,M^i)$ are characteristic for all $i \geq 0$, when the characteristic of R/M is not 2 and $n \geq 3$.

Notation and Terminology

Throughout this paper all rings are assumed to have identity element which is denoted by 1 or by I in the case of matrix rings. The symbol R always denotes a local ring with maximal ideal denoted by M.

The following is a list of symbols and a brief description of their use. Precise definitions are given in the text of the paper.

$R^{(n)}$, the n-dimensional free R-module, $\sum_{n\text{-summands}} R$.

GL(V), the group of invertible linear transformations of the free R-module V.

$GL_n(R)$, the group of invertible n by n matrices

$M_n(S)$, the n by n matrix ring over the ring S.

SL(V), the subgroup of GL(V) consisting of elements with determinant 1.

$SL_n(R)$, the subgroup of $GL_n(R)$ consisting of elements of determinant 1.

$GC_n(R,J)$, the general congruence subgroup modulo J, defined in Chapter II.

$SC_n(R,J)$, the special congruence subgroup modulo J, defined in Chapter II.

$H \leq G$, H is a subgroup of G.

$A \subseteq B$, A is a subset of B.

ρ_i, the natural ring morphism of $R \to R/M^i$.

h_i, the natural group morphism induced by ρ_i on $Gl_n(R) \to Gl_n(R/M^i)$.

K_i, the kernel of h_i.

E_{ij}, the matrix having 1 in the (i,j) position and zeros elsewhere.

$B_{ij}(\lambda)$, the elementary transvection, $I + \lambda E_{ij}$.

S_{ij}, the permutation matrix, $B_{ij}(1)B_{ji}(-1)B_{ij}(1)$.

$D_j(r)$, the diagonal matrix, $I + (r-1)E_{jj}$.

J_{ij}, the diagonal involution having - 1 in the (i,i) and (j,j) positions and 1 elsewhere on the main diagonal.

$[a_1,a_2,\cdots,a_n]$, the n by n diagonal matrix having a_i in the (i,i) position.

$\oplus$, direct sum for matrices and for modules. (Use is clear from context.)

CHAPTER I

LOCAL RINGS

PROPERTIES OF LOCAL RINGS

Let S be a commutative ring and J an ideal of S. J is called a maximal ideal of S if $J \neq S$ and J is not a proper subset of any ideal except S.

Definition 1.1. A commutative ring with identity is a local ring if it has exactly one maximal ideal.

An element of a ring is called a unit if it has a multiplicative inverse and the set of units in a ring S is denoted by S^*. It is clear then that a commutative ring S is a local ring if and only if $S - S^*$ is an ideal. Consequently in a local ring a sum of elements is a unit only if one of the summands is a unit.

If R is a local ring with maximal ideal M, then the ring R/M is a field and is called the residue field of R. The natural epimorphism,

$$R \rightarrow R/M$$

will be denoted by ρ and images of elements under ρ will be denoted by bars. That is, if r is in R then the residue class $\rho(r) = r + M$ will be denoted by $\bar{r}$.

The following are examples of local rings:

Example 1.1 Any field is a local ring with maximal ideal (0).

Example 1.2 The rational integers modulo a power of a prime, Z/Zp^n, is a local ring with maximal ideal Zp/Zp^n.

Example 1.3 If S is a commutative ring and P is a prime ideal in S then S_p, the ring of fractions with denominators from S - P, is a local ring. The maximal ideal consists of all elements $\frac{p}{t}$ where p is in P. S_p is called the localization of S at P.

Local rings play the role of elementary building blocks for commutative rings that satisfy the descending chain condition on ideals, that is, Artinian rings.

Theorem 1.1 (The Structure Theorem for Artinian Rings) An Artinian ring is uniquely (up to isomorphism) a finite direct product of Artinian local rings.

Proof: See Atiyah and MacDonald [2].

Finite commutative rings obviously are Artinian and therefore Theorem 1.1 reduces the study of these finite rings to the study of their finite local components. We also note that Artinian local rings have nilpotent maximal ideals and thus Theorem 1.1 reduces the study of Artinian rings to the study of local components having nilpotent maximal ideals. If R is a local ring with nilpotent maximal ideal M then the smallest positive integer β satisfying $M^\beta = 0$ is called the degree of nilpotency of M.

We note also that structures "over" a ring S are often productive and thus Theorem 1.1 simplifies their study. In particular, for a ring S let $M_n(S)$ denote the n by n matrix over S and S[X] the polynomial ring.

If S is isomorphic to a direct product,

$$S = \prod_{i=1}^{t} S_i,$$

then

$$M_n(S) \cong \prod_{i=1}^{t} M_n(S_i) \qquad \text{(ring isomorphism)}$$

and

$$S[X] \cong \prod_{i=1}^{t} S_i[X] \qquad \text{(ring isomorphism)}$$

Suppose that R is a finite local ring with maximal ideal M. The residue field, R/M, is a finite field and therefore has characteristic p for some prime number p. It can be shown that if R/M has characteristic p, then R has characteristic p^t for some t and cardinality p^s for some s. The following theorem, proven by Ganske in [13], gives the structure of finite local rings.

Theorem 1.2 Let R be a finite local ring with characteristic p^t and maximal ideal M having minimal generating set of cardinality n. Then there is a subring T of R such that T is a separable extension of Z/Zp^t and R is a homomorphic image of $T[X_1,\cdots,X_n]$.

Modules Over Local Rings

The following theorem is fundamental for finitely generated modules over local rings. A detailed treatment of the theorem and proof can be found in Nagata [20].

Theorem 1.3. Let R be a local ring with maximal ideal M and let E be a finitely generated R-module.

(i) A subset $\{u_i\}_{i=1}^{n}$ of E is a generating set for E if and only if their residue classes $\{\bar{u}_i\}_{i=1}^{n}$ generate the vector space E/ME over the field R/M.

(ii) A subset $\{u_i\}_{i=1}^{n}$ of E is a minimal generating set for E if and only if $\{\bar{u}_i\}_{i=1}^{n}$ is a linearly independent basis for E/ME over R/M.

(iii) Any generating set for E contains a minimal generating set; if $u_1,\cdots,u_n$ and $v_1,\cdots,v_m$ are both minimal generating sets for E then $m = n$ and there is an R-isomorphism of $E \to E$ that maps $u_i \to v_i$, $1 \leq i \leq n$.

Let V denote the n-dimensional free module over the local ring R. If $e_1,\cdots,e_n$ is an R-basis for V then $Re_i \cong R$ for all $1 \leq i \leq n$ and $V = Re_1 \oplus \cdots \oplus Re_n \cong R^{(n)}$. When dealing with V, we will often assume that a free basis has been chosen and will treat elements of V as n-component vectors.

Observe that there is no ambiguity in the dimension of a finitely generated free R-module because by Theorem 1.3 (iii) each basis will have the same number of elements.

Definition 1.2 K is a subspace of free R-module V if K is a direct summand of V.

The following theorem, due to Kaplansky [17], has as a consequence that subspaces are free R-modules.

Theorem 1.4 Projective modules over local rings are free.

A subspace K of the free R-module V is a summand of the free module V and is therefore projective, hence free. Furthermore if $K \neq V$, then the dimension of K is strictly smaller than that of V. The situation appears to be very much the same as with a finite dimensional vector space over a field. However, it is not the case that submodules of V are necessarily subspaces as happens over fields; for example, if e_1 is a free basis element and m is in the maximal ideal, then $R(e_1m)$ is a non-trivial submodule but not a subspace.

Definition 1.3 Let V be an n-dimensional free module over the local ring R. An (n-1)-dimensional subspace of V is called a hyperplane

of V; a 1-dimensional subspace is called a line of V ; and a cyclic submodule of V is called a ray of V.

Theorem 1.5 Let V be an n-dimensional free module over the local ring R with maximal ideal M. If $m \leq n$ and $u_1,\cdots,u_m$ are linearly independent modulo M, then $u_1,\cdots,u_m$ can be extended to a free R-basis of V.

Proof: Applying Theorem 1.3, $u_1,\cdots,u_m$ are linearly independent in V/MV over R/M and thus can be extended to a basis of the vector space V/MV. If $\{\bar{u}_1,\cdots,\bar{u}_n\}$ is such a basis, then again by Theorem 1.3, $\{u_1,\cdots,u_n\}$ is a free generating set for V over R.

Linear Groups

Throughout this section R will denote a local ring with maximal ideal M and V will denote an n-dimensional free R-module, $n \geq 2$.

Definition 1.4 The general linear group of V over R is the group of all invertible R-linear mappings of V to V. The general linear group is denoted GL(V).

We use $GL_n(R)$ to denote the group of invertible $n \times n$ matrices over R and observe that by fixing a basis for V there is a natural matrix representation such that $GL(V) \cong GL_n(R)$. It is often convenient to consider that a basis has been fixed and to use $GL_n(R)$ to obtain results for GL(V).

Definition 1.5 The special linear group of V over R is the subgroup of GL(V) consisting of elements having determinant 1. It is denoted by SL(V).

The set of $n \times n$ matrices having determinant 1 is denoted by $SL_n(R)$ and $SL(V) \cong SL_n(R)$ under the above isomorphism $GL(V) \cong GL_n(R)$.

We will use $GL_n(R)$ to describe basic generators for the general and special linear groups. The letter I will denote the identity matrix; and E_{ij} will denote the $n \times n$ matrix having zeros everywhere except in the (i,j) position where a 1 appears. In the $n \times n$ matrix ring $M_n(R)$ we have,

$$E_{ij}\, E_{km} = \delta_{jk}\, E_{im} \quad \text{where} \quad \delta_{jk} = \begin{cases} 0, j \neq k \\ 1, j = k \end{cases}.$$

Consider two "elementary" matrices:

$$B_{ij}(\lambda) = I + \lambda E_{ij}, \quad i \neq j \quad \text{and } \lambda \in R;$$

and

$$D_i(u) = I + (u - 1)E_{ii}, \quad u \in R^*.$$

The matrix $B_{ij}(\lambda)$ is called an <u>elementary transvection</u>. A casual observation yields,

$$B_{ij}(\lambda)^{-1} = B_{ij}(-\lambda),$$

$$B_{ij}(\lambda_1) \cdot B_{ij}(\lambda_2) = B_{ij}(\lambda_1 + \lambda_2),$$

and

$$\det (B_{ij}(\lambda)) = 1.$$

Furthermore, if A is an element of $GL_n(R)$ then $B_{ij}(\lambda) \cdot A$ is the matrix obtained from A by adding λ times the j-th row to the i-th row. Similarly, right multiplication by $B_{ij}(\lambda)$ adds λ times the i-th column to the j-th column.

The matrix $D_i(u)$ is called an <u>elementary diagonal element</u> and we have

$$D_i(u)^{-1} = D_i(u^{-1}),$$

$$D_i(u_1) \cdot D_i(u_2) = D_i(u_1 u_2),$$

and

$$\det D_i(u) = u.$$

If A is in $GL_n(R)$, then $D_i(u)\cdot A$ is the matrix obtained from A by multiplying the i-th row by u. Right multiplication by $D_i(u)$ multiplies the i-th column by u.

If P is a subset of M then $B_{ij}(P)$ denotes an arbitrary elementary transvection with an element from P in the (i,j) position. The notation $D_i(1 - P)$ is analogous.

If A is an element of $GL_n(R)$, then det A is in R^*. Since R is a local ring, it follows that A must have at least one unit in each row and in each column. This observation allows us to establish the following theorem on elementary generators for $GL_n(R)$.

Theorem 1.6 If R is a local ring, then $GL_n(R)$ is generated by the set

$$\{B_{ij}(\lambda), D_k(u) \mid \lambda \in R,\ u \in R^*,\ i \neq j,\ 1 \leq i,j,k \leq n\}.$$

Proof: The proof proceeds by showing that a matrix can be reduced to the identity matrix under left multiplication by elements from the asserted set of generators.

If $A = [a_{ij}]$ is in $GL_n(R)$, then some element in the first column of A is a unit. The product $P_{ij} = B_{ij}(1)B_{ji}(-1)B_{ij}(1)$ permutes the i-th and j-th rows while changing the sign on the i-th. We can arrange to have a unit in the (1,1) position then by multiplying A by P_{1j} for suitable choice of j. Now assume a_{11} is a unit and place 1 in the (1,1) position by left multiplying by $B_{21}(a_{11}^{-1}(1 - a_{21}))$ and then by $B_{12}(1 - a_{11})$. With 1 in the (1,1) position, the remainder of first column is reduced

to zeros by left multiplying by $B_{i1}(-a_{i1})$.

This same process is applied to all of the columns, putting 1 on the diagonal and zero elsewhere, until only the n-th column remains. Note that we have only used elementary transvections in the reduction. Now the (n,n) position must contain a unit so that left multiplication by $B_{ij}(a_{nn}^{-1}(-a_{in}))$, for $1 \leq i \leq n-1$, reduces the last column to zeros except for a_{nn}. Finally $D_n(a_{nn}^{-1})$ completes the reduction of A to I. We have then that,

$$D_n(a_{nn}^{-1}) \cdot B \cdot A = I,$$

where B denotes a product of elementary transvections. Therefore,

$$A = B^{-1} \cdot D_n(a_{nn}).$$

The proof is completed by recalling that $B_{ij}(\lambda)^{-1} = B_{ij}(-\lambda)$.

Corollary 1. Any element A of $GL_n(R)$ can be written as $BD_n(u)$ where B is a product of elementary transvections and $u = \det A$.

Proof: In the proof of the theorem, A is expressed precisely in this form; furthermore, $\det(B) = 1$ and thus $\det A = u$.

Corollary 2. $SL_n(R)$ is the group generated by all elementary transvections.

Proof: Clearly each $B_{ij}(\lambda)$ is an element of $SL_n(R)$. Further, by Corollary 1, any element of $GL_n(R)$ having determinant 1 is a product of elementary transvections.

Commutators

In this section we list some basic relations for commutators of the elementary generators.

Definition 1.6 If G is a group and x and y elements of G, then the

product $x^{-1}y^{-1}xy$ is called the commutator of x and y and is denoted by [x,y]. If H and K are subgroups of G, the commutator subgroup generated by H and K is the subgroup of G generated by elements [h,k] where h is in H and k is in K; it is denoted by [H,K].

In the following lemma we give several identities that will be important throughout the remainder of the paper. The proofs will not be given since all are routine calculations. The equations are true in any commutative ring.

Lemma 1.1 (The Basic Identities).

(i) Let $A = [a_{ij}]$ be in $GL_n(R)$ and $A^{-1} = [\alpha^{ij}]$; then

$$[A,B_{ij}(\lambda)] = B_{ij}(\lambda) - \lambda^2 a_{ji} \sum_t \alpha^{ti} E_{tj} - \lambda \sum_{t,s} \alpha^{ti} a_{js} E_{ts}.$$

(ii) If $1 - r$ is a unit and λ is in R, then

$$[D_k(1-r),B_{ij}(\lambda)] = \begin{cases} I, & \text{if } k \neq i \text{ and } k \neq j \\ B_{ij}(-r\lambda(1-r)^{-1}), & \text{if } k = i \\ B_{ij}(r\lambda), & \text{if } k = j \end{cases}$$

(iii) If λ_1, λ_2 are in R and i, j, k are distinct indices, then

$$[B_{ij}(\lambda_1),B_{jk}(\lambda_2)] = B_{ik}(\lambda_1\lambda_2).$$

(iv) If $i \neq k$ and $j \neq m$,

$$[B_{ij}(\lambda_1),B_{mk}(\lambda_2)] = I.$$

(v) $$[B_{ij}(\lambda_1),B_{ji}(\lambda_2)] = I + (\lambda_1\lambda_2 - \lambda_1^2\lambda_2^2)E_{ii} + \lambda_1^2\lambda_2 E_{ij} - \lambda_1\lambda_2^2 E_{ji} - \lambda_1\lambda_2 E_{jj}.$$

(vi) If 1 - ru is a unit, then

$$D_i\left(\frac{1}{1-ru}\right)D_j(1-ru) = B_{ij}\left(\frac{-r^2u}{1-ru}\right)[B_{ij}(r),B_{ji}(u)]B_{ji}\left(\frac{ru^2}{1-ru}\right).$$

(vii) For A, B, C matrices in $GL_n(R)$,

$$[AB,C] = [A,C][[A,C],B][B,C].$$

Individual parts of this lemma will be referred to as "Basic Identity (j)".

The Finite Case

As mentioned previously if R is a finite local ring, then there is a prime p such that the cardinality of R is a power of p and R/M is a finite field of characteristic p. The natural map of $R \to R/M$ induces a group morphism denoted by h_M: $GL_n(R) \to GL_n(R/M)$. The kernel of h_M is the group of matrices that are congruent modulo M to the identity. When R is finite, we have $GL_n(R)$ finite. The following theorem gives the cardinality of $GL_n(R)$.

Theorem 1.7. Let R be a finite local ring with R/M having characteristic p and cardinality $q = p^t$; then

$$|GL_n(R)| = |R|^{n^2} \prod_{i=0}^{n-1} (1 - q^{i-n}).$$

Proof: In [10] Dickson shows that,

$$|GL_n(R/M)| = q^{n^2} \prod_{i=0}^{n-1} (1 - q^{i-n}).$$

This is easy to see by considering the number of choices for linearly independent rows. The map

$$h_M: GL_n(R) \to GL_n(R/M)$$

has a kernel, as described above, of cardinality $|M|^{n^2}$. Since h_M is a surjection, we have,

$$|GL_n(R)| = |M|^{n^2} |GL_n(R/M)|$$

$$= |M|^{n^2} q^{n^2} \prod_{i=0}^{n-1} (1 - q^{i-n})$$

$$= |R|^{n^2} \prod_{i=0} (1 - q^{i-n}).$$

In order to count $SL_n(R)$, note that for the units R^* of R we have,

$$|R^*| = |R| - |M| = |R|(1 - |R/M|).$$

Theorem 1.8 If R is a finite local ring with $|R/M| = q = p^t$, then

$$|SL_n(R)| = |R|^{n^2-1} \prod_{i=0}^{n-1} (1 - q^{i-n})/(1 - \frac{1}{q}).$$

Proof: The determinant map,

$$\det: GL_n(R) \to R^*$$

is an epimorphism with kernel $SL_n(R)$. Thus

$$[GL_n(R): SL_n(R)] = [R^*: 1].$$

Therefore

$$|SL_n(R)| = \frac{|GL_n(R)|}{|R^*|}$$

and the conclusion follows from the previous theorem.

CHAPTER II

THE NORMAL SUBGROUP STRUCTURE OF GL(V)

In a series of three papers Brenner [5], [6], [7] described the lattice of normal subgroups of $GL_n(Z/Zp^n)$ in terms of certain maximal and minimal congruence subgroups. Brenner also determined characteristic subgroups of $GL_n(Z/Zp^n)$. Klingenberg [18] has generalized many of the results on normal subgroups to the case of an arbitrary local ring.

Congruence Subgroups of GL(V)

Let R be a local ring with maximal ideal M, and V be an n-dimensional free R-module. We shall sketch Klingenberg's results on normal subgroups of GL(V). For our purposes it suffices to consider $n \geq 3$, although most of the following theorems remain true when $n = 2$ if one assumes R/M does not have characteristic 2.

If J is an ideal of R, we let V/J denote the n-dimensional free module over R/J. The map ρ_J: $R \to R/J$ induces natural morphisms

$$g_J: V \to V/J \qquad \text{(module morphism)}$$

and

$$h_J: GL(V) \to GL(V/J) \qquad \text{(group morphism)}.$$

When a basis has been fixed, g_J merely reduces components of a vector modulo J and h_J reduces entries of a matrix modulo J. More precisely h_J is defined by,

$$(h_J\sigma)(g_JX) = g_J(\sigma X) \quad \text{for all } \sigma \text{ in } GL(V),\ X \text{ in } V.$$

If $J = (0)$, then g_J and h_J are identity morphisms; if $J = R$, $GL(V/R)$ is defined to be the identity group.

Definition 2.1. If X is an element of V, the order of X, $O(X)$, is the smallest ideal J such that $g_J(X) = 0$. If σ is in $GL(V)$, the order of σ, $O(\sigma)$, is the smallest ideal such that $h_J\sigma$ is in the center of $GL(V/J)$. If G is a subgroup of $GL(V)$, the order of G, $O(G)$, is the ideal generated by $\{O(\sigma)\colon \sigma \text{ is in } G\}$.

If X is in V and has representation $(a_1,\cdots,a_n)$ with respect to some basis, then $O(X)$ is precisely the ideal generated by the set $\{a_i\colon 1 \le i \le n\}$. Note also that an element X can be extended to a basis of V if and only if $O(X) = R$. If σ is in $GL(V)$ and has matrix $[a_{ij}]$ with respect to some basis, then $O(\sigma)$ is the smallest ideal J such that $[a_{ij}]$ is scalar modulo J. Equivalently $O(\sigma)$ is generated by $\{a_{ij},\ a_{ii} - a_{jj}\colon i \ne j,\ 1 \le i, j \le n\}$.

Definition 2.2. For each ideal J of R the general congruence group modulo J, $GC_n(R,J)$, is the group h_J^{-1} (center of $GL(V/J)$).

$GC_n(R,J)$ is obviously a normal subgroup of $GL(V)$. For the extreme cases we have that $GC_n(R,0)$ = center of $GL(V)$ and $GC_n(R,R) = GL(V)$. In terms of matrix representations, $GC_n(R,J)$ is the set of matrices that are scalar modulo J.

Definition 2.3. An element τ of $GL(V)$ is called a transvection if there is a hyperplane H of V such that $\tau X - X$ is in H for all X in V and τ is the identity map when restricted to H. If a basis is chosen, say $\{e_1,\cdots,e_n\}$, then for $i \ne j$, $B_{ij}(\lambda)$ is a transvection with hyperplane generated by $\{e_1,\cdots,e_{j-1},e_{j+1},\cdots,e_n\}$ and $B_{ij}(\lambda)(X) - X$ is contained in

the line generated by e_i.

Lemma 2.1. Let τ be a transvection in GL(V) with fixed hyperplane H. Then there is a functional ϕ in the dual space of V and an element A of H such that

$$\tau X = X + \phi(X)A, \quad \text{for all } X \text{ in } V.$$

Proof: There is obviously a functional, ϕ, such that $H = \phi^{-1}(0)$ and $\phi(B) = 1$ for some element B in V-H. For example, choose B such that $V = H \oplus R \cdot B$ and simply define $\phi(B) = 1$ and $\phi(X) = 0$ for X in H. Then extend ϕ to V linearly. Now let $A = \tau B - B$. Then A is in H and $X - \phi(X)B$ is in H. Thus for all X in V,

$$\tau(X - \phi(X)B) = X - \phi(X)B$$

and therefore

$$\tau X = X + \phi(X)(\tau B - B) = X + \phi(X)A.$$

This completes the proof of the lemma.

If τ is a transvection with

$$\tau X = X + \phi(X)A,$$

then A is said to belong to τ and is determined by H up to a unit multiple. That is, if ϕ is replaced by $u\phi$, u in R^*, then A is replaced by $u^{-1}A$. It is also clear that the ideals $0(A)$ and $0(\tau)$ are identical for $h_J\tau$ is scalar if and only if $g_J A = 0$.

Conversely, if ϕ is a surjective element of the dual space, then $\phi^{-1}(0)$ is a summand of V of dimension n - 1. If $H = \phi^{-1}(0)$ and A is any element of H, then the map τ described by,

$$\tau X = X + \phi(X) \cdot A,$$

is a transvection with A the vector belonging to τ and hyperplane H.

If τ is a transvection with hyperplane H, then by choosing a basis for V, say $\{e_1,\cdots,e_n\}$, with e_i in H for $1 \le i \le n-1$ and $\phi(e_n) = 1$ we have the matrix representation,

$$\text{mat } (\tau) = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & a_1 \\ 0 & 1 & 0 & \cdots & 0 & a_2 \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & \cdot & \cdots & 1 & a_{n-1} \\ 0 & 0 & \cdot & \cdots & 0 & 1 \end{bmatrix}$$

where $A = \sum_{i=1}^{n-1} a_i e_i$. It follows that det $(\tau) = 1$ for any transvection τ.

Definition 2.4. For each ideal J of R the special congruence group modulo J, $SC_n(R,J)$, is the normal subgroup of GL(V) generated by all transvections τ such that $0(\tau) \subseteq J$.

For the extreme cases we have that $SC_n(R,(0)) = I$ and $SC_n(R,R) = SL(V)$.

Klingenberg's Results

In this section we will sketch Klingenberg's results on normal subgroups of GL(V). All of the propositions appear in [18] and we shall only indicate the method of proof in each case. R is a local ring with maximal ideal M and V is an n-dimensional free R-module.

Lemma 2.2 (Klingenberg). Two elements X, Y of V have the same order if and only if there is a σ in GL(V) such that $\sigma X = Y$.

Proof: (Klingenberg) Let $\sigma X = Y$. Note that $(h_J\sigma)(g_J X) = 0$ if and only if $g_J(X) = 0$. But $(h_J\sigma)(g_J X) = g_J(\sigma X) = g_J(Y)$. Thus $0(X) = 0(Y)$.

Conversely, supposing that $0(X) = 0(Y) = J$, choose a minimal

generating set $\{u_i\}_{i=1}^m$ for J and show there are bases $\{e_i\}_{i=1}^n$ and $\{f_i\}_{i=1}^n$ for V such that $X = \sum_{i=1}^m u_i e_i$ and $Y = \sum_{i=1}^m u_i f_i$. Then take σ to be the change of basis map, $\sigma(e_i) = f_i$, $1 \leq i \leq n$, and have $\sigma X = Y$.

Theorem 2.1 (Klingenberg). Two transvections τ_1 and τ_2 of GL(V) have the same order if and only if they are conjugate in GL(V). Furthermore, if $n \geq 3$ and $O(\tau_1)$ is a principal ideal, then τ_1 and τ_2 are conjugate in SL(V).

Proof: (Klingenberg) As in the lemma if τ_1 and τ_2 are conjugate, then $O(\tau_1) = O(\tau_2)$.

For the converse suppose,

$$\tau_1 X = X + \phi_1(X)A_1$$

and

$$\tau_2 X = X + \phi_2(X)A_2.$$

Then by lemma 2.1 there is a σ in GL(V) such that $\sigma A_1 = A_2$ and $\sigma H_1 = H_2$. Then a calculation yields, $\tau_2 = \sigma\tau_1\sigma^{-1}$. When the order of τ is principal, then σ can be chosen to have determinant 1.

The next theorem characterizes the special congruence subgroups and part (ii) gives rise to an easy matrix interpretation of $SC_n(R,J)$.

Theorem 2.2 (Klingenberg). Let J be an ideal of the local ring R and G be a subgroup of GL(V). The following are equivalent:

(i) $G = SC_n(R,J)$

(ii) G consists of elements σ such that $h_J\sigma = I$ and $\det \sigma = 1$.

(iii) G is the mixed commutator group $[GL(V), GC_n(R,J)]$.

Proof: (Klingenberg) Let G_2 denote the group defined in (ii) and G_3 the group defined in (iii). It is clear from our representation of

transvections that each transvection of order J is contained in G_2. Thus $SC_n(R,J)$ is a subgroup of G_2. But any matrix representation of an element of G_2 can be reduced to the identity by left multiplication by elementary transvections of order contained in J and by $GL_n(R)$ conjugates of elementary transvections of order contained in J. Therefore G_2 is a subgroup of $SC_n(R,J)$ and we have $G_2 = SC_n(R,J)$.

For part (iii), it is clear that G_3 is a subgroup of $G_2 = SC_n(R,J)$. For the reverse we suppose that τ is a transvection of order contained in J and

$$\tau X = X + \phi(X)A, \quad O(A) = O(\tau).$$

There exist A_1 and A_2 in V such that $A = A_2 - A_1$ and $O(A) = O(A_1) = O(A_2)$. Define τ_1, τ_2 by

$$\tau_i X = X + \phi(X)A_i$$

and note there is a σ in GL(V) such that $\tau_2 = \sigma\tau_1\sigma^{-1}$. Then we have $\tau = \tau_2\tau_1^{-1} = \sigma\tau_1\sigma^{-1}\tau_1^{-1}$ and the latter is in G_3. Thus $SC_n(R,J)$ is a subgroup of G_3 and it follows that (i), (ii), (iii) are equivalent.

Theorem 2.3 (Klingenberg). Let J be an ideal of the local ring R.

(i) $GC_n(R,J)/SC_n(R,J)$ is isomorphic to the subgroup of $R^* \times (R/J^*)$ that is formed by the elements (a,b) such that $\rho_J(a) = b^n$.

(ii) The group $GC_n(R,J) \cap SL_n(R)$ is isomorphic to the n-th roots of unity in $(R/J)^*$.

Proof (Klingenberg): The proof of both parts follows immediately from the observation that

$$\text{center } (GL(V/J)) \cong (R/J)^*$$

and by considering the image of the morphism

$$f: GC_n(R,J) \to R^* \times \text{center } (GL(V/J))$$

where $\sigma \longrightarrow (\det \sigma, h_J\sigma)$. The morphism f has kernel $SC_n(R,J)$ and image $\{(a,b)\}$ such that $a = \det \sigma$ and $[b,\cdots,b] = h_J\sigma$.

An immediate consequence of Theorem 2.3 is that $GC_n(R,J)/SC_n(R,J)$ is an Abelian group and thus any group G satisfying,

$$(*) \qquad GC_n(R,J) \geq G \geq SC_n(R,J)$$

is a normal subgroup of GL(V). It is obvious also that if G satisfies (*), then $0(G) = J$. The consequence of the next theorem is that conversely for any SL(V) invariant subgroup G of GL(V), there is an ideal J for which G satisfies (*).

Theorem 2.4 (Klingenberg). (i) Let $\{\tau_\alpha\}_{\alpha \in q}$ be any set of transvections. The SL(V) invariant subgroup G of GL(V) generated by the is precisely $SC_n(R,J)$ where J is the ideal generated by $\{0(\tau_\alpha)\}_{\alpha \in q}$.

(ii) Let G' be an SL(V) invariant subgroup of GL(V) of order J. Then G' contains the group $SC_n(R,J)$.

Proof (Klingenberg): (i) First assume that there is only one transvection, τ, in the collection. If $0(\tau)$ is principal then Theorem 2.1 implies that G, being SL(V) invariant, contains all transvections of $0(\tau)$. If $\tau X = X + \phi(X)A$ and $A = \sum_{i=1}^{n} a_i e_i$ with respect to the basis $\{e_1,\cdots,e_n\}$, then there is an SL(V) conjugate of τ having order (a_1) for each $1 \leq i \leq n$. Thus by Theorem 2.1, G contains all transvections of order (a_i), $1 \leq i \leq n$, and hence G contains all transvections of order 0(A) the ideal generated by $\{a_1,\cdots,a_n\}$. Finally if $\{\tau_\alpha\}_{\alpha \in q}$ is an arbitrary collection and $0(\tau)$ is contained in the ideal generated by $\{0(\tau_\alpha)\}_{\alpha \in q}$, then

$$\tau X = X + \phi(X)\cdot A$$

where $A = \sum_{i=1}^{n} A_{\alpha_i}$ is a finite sum with A_{α_i} having order contained in $O(\tau_{\alpha_i})$. By the first part G contains each of the transvections τ'_{α_i} defined by,

$$\tau'_{\alpha_i} = X + \phi(X)\cdot A_{\alpha_i}.$$

But then G contains τ since

$$\prod_{i=1}^{n} \tau'_{\alpha_i}(X) = X + \phi(X)\left(\sum_{i=1}^{n} A_{\alpha_i}\right) = \tau(X).$$

Since G contains all transvections of order J, $G = SC_n(R,J)$.

The proof of (ii) relies heavily on (i). Let σ be an element of the SL(V) invariant group G. The ideal $O(\sigma)$ is generated by elements a of G satisfying the property that there is a free element E of V and a surjective element of the dual space of V, ϕ, such that $\phi(E) = 0$ and $\phi(\sigma E) = a$. On the basis of this we show that G contains a transvection of order containing (a). Thus by (i) G' contains $SC_n(R,(a))$ for each generator, (a), of $O(\sigma)$. Thus G' contains $SC_n(R,O(\sigma))$ for all σ in G and, since $O(G') = J = (\{O(\sigma): \sigma \text{ in } G\})$, we have $SC_n(R,J)$ is contained in G'.

Corollary A subgroup G of GL(V) is normal if and only if G is SL(V) invariant.

The final theorem of this section is stated to summarize the foregoing results. The conclusions have been proven or are direct consequences of the previous results.

Theorem 2.5 (Klingenberg). Let R be a local ring.

(i) A subgroup G of GL(V) which is invariant under SL(V) determines

an ideal J in R such that

$$GC_n(R,J) \geq G \geq SC_n(R,J).$$

Conversely every subgroup G of GL(V) that satisfies this relation is normal and has order J.

(ii) A normal subgroup, G, of SL(V) determines an ideal J such that

$$GC_n(R,J) \cap SL(V) \geq G \geq SC_n(R,J).$$

Conversely every subgroup of SL(V) satisfying R relation is normal and has order J.

Characteristic Subgroups

In [6] Brenner gives the theorems of the previous section for the case where R is Z/Zp^r. In addition Brenner shows that the subgroups $GC_n(R,Zp^t)$ and $SC_n(R,Zp^t)$ are invariant under group isomorphisms of $GL_n(R)$. His proof depends heavily on two properties of Z/Zp^r. The first property is that the groups are finite, so that counting arguments can be used; the second property is that the ideals are totally ordered. It is obvious from the results of the last section and the fact that $GC_n(R,Zp^t)$ and $SC_n(R,Zp^t)$ map to themselves under group automorphisms that automorphisms of $GL_n(Z/Zp^r)$ inner automorphisms. What subgroups are preserved in the general case?

Definition 2.5. A subgroup H of a group G is called characteristic if for all automorphisms, Λ: G $\to$ G, ΛH is a subset of H. H char G will denote the phrase "H is a characteristic subgroup of G".

Note that any characteristic subgroup of G is normal since it is invariant under conjugation, that is, under inner automorphisms. Also observe that if H char G and K char G, then also [H,K] char G.

We now ask the following question: What conditions on the local ring

R are necessary to assure that $GC_n(R,M^i)$ and $SC_n(R,M^i)$ are characteristic? Obviously from Brenner's work R being finite and having ideals totally ordered is enough. Note also that by Theorem 2.2 we need only show that $GC_n(R,M^i)$ char GL(V) because $SC_n(R,M^i) = [GL(V), GC_n(R,M^i)]$.

Observation 2.1 Let R be a local ring. The subgroups $GC_n(R,M)$, $SC_n(R,M)$ and SL(V) are characteristic subgroups of GL(V).

Proof: $SL(V) = [GL(V),GL(V)]$ by Theorem 2.2 and thus is characteristic. For $GC_n(R,M)$, let Λ be an automorphism and note that $\Lambda(GC_n(R,M))$ is a normal subgroup. By Klingenberg's theory $\Lambda GC_n(R,M)$ is either between GL(V) and SL(V) or contained in $GC_n(R,M)$. The former is impossible, for then it would follow that $SL(V) \subseteq GC_n(R,M)$. Thus $\Lambda GC_n(R,M)$ is contained in $GC_n(R,M)$ and hence $GC_n(R,M)$ char GL(V). Finally $SC_n(R,M)$ is characteristic, since $SC_n(R,M) = [GL(V),GC_n(R,M)]$.

For higher powers of M the general problem is greatly complicated by the abundance of ideals not lying in the chain $M \supseteq M^2 \supseteq M^3 \supseteq \cdots$. The question is slightly easier in the finite case as will be shown in the next chapter. The final affirmative answer to the above question will not follow, however, until the group of automorphisms of GL(V) has been determined. This will be done in the last chapter.

The Finite Case

Throughout this section R will denote a finite local ring with maximal ideal M. We shall count the congruence subgroups and determine the structure of $GC_n(R,M^i)/GC_n(R,M^{i+1})$ and $SC_n(R,M^i)/SC_n(R,M^{i+1})$.

Theorem 2.6. Let R be a finite local ring and J be an ideal of R.

(i) $|GC_n(R,J)| = |R^*| \cdot |J|^{n^2-1}$

(ii) $|SC_n(R,J)| = |J|^{n^2-1}$

Proof: (i) Elements of $GC_n(R,J)$ are scalar modulo J and thus each off-diagonal position can be filled in any of $|J|$ ways. Along the diagonal each position must contain a unit and the units must be pairwise congruent modulo J. After the unit in the (1,1) position has been chosen there are $|J|$ ways to fill each remaining diagonal position. Therefore there are $|R^*|\cdot|J|^{n-1}$ possible diagonals and, for each choice of diagonal, $|J|^{n^2-n}$ possibilities for the remaining positions. Thus $|GC_n(R,J)| = |R^*|\cdot|J|^{n-1}|J|^{n^2-n} = |R^*||J|^{n^2-1}$.

(ii) Let K_J denote the set of matrices in $GL_n(R)$ that are congruent to 1 modulo J. Obviously $|K_J| = |J|^{n^2}$. Now consider

$$\det : K_J \to H = \{r \in R^* \mid r \equiv 1 \bmod J\}.$$

Clearly elements of K_J have determinants congruent to 1 modulo J. Moreover the kernel of this group morphism is precisely $SC_n(R,J)$. Observing that $|H| = |J|$ we have

$$|SC_n(R,J)| = \frac{|K_J|}{|H|} = \frac{|J|^{n^2}}{|J|} = |J|^{n^2-1}.$$

Note that when R is finite, the quotients M^i/M^{i+1} are naturally R/M-vector spaces of finite dimension. Suppose $|R/M| = q = p^t$ and that $\psi_i = \dim_{R/M}(M^i/M^{i+1})$, $i \geq 1$.

Theorem 2.7. Let R be a finite local ring with $|R/M| = q = p^t$.

(i) $SC_n(R,M^i)/SC_n(R,M^{i+1})$ is an elementary Abelian p-group and is a direct product of $t\cdot\psi_i\cdot(n^2-1)$ cyclic groups of order p.

(ii) $GC_n(R,M^i)/GC_n(R,M^{i+1}) \cong SC_n(R,M^i)/SC_n(R,M^{i+1})$.

Proof: From the previous theorem we have

$$\left|GC_n(R,M^i)/GC_n(R,M^{i+1})\right| = \left|M^i/M^{i+1}\right|^{n^2-1}$$
$$= \left|SC_n(R,M^i)/SC_n(R,M^{i+1})\right|.$$

Thus if $|R/M| = p^t$, then

$$\left|M^i/M^{i+1}\right|^{n^2-1} = (p^t)^{\psi_i \cdot (n^2-1)}.$$

It only remains to show the quotient groups are elementary Abelian. Let $I + N_1$ and $I + N_2$ be elements of $SC_n(R,M^i)$ where N_1 and N_2 are n by n matrices with elements from M^i. Since M is nilpotent, the representations

$$(I + N_j)^{-1} = I - N_j + N_j^2 - \cdots, \quad j = 1,2$$

are finite sums. We now compute the commutator,

$$\begin{aligned}
[I+N_1, I+N_2] &= (I+N_1)^{-1}(I+N_2)^{-1}(I+N_1)(I+N_2) \\
&= (I+N_1)^{-1}(I+N_2)^{-1}(I+N_2+N_1+N_1N_2) \\
&= (I+N_1)^{-1}(I+N_1+N_1N_2-N_2N_1-N_2N_1N_2+\cdots) \\
&= (I+N_1N_2-N_2N_1-N_2N_1N_2+\cdots-N_1^2N_2+N_1N_2N_1+N_1N_2N_1N_2-\cdots) \\
&\equiv I \pmod{M^{i+1}}.
\end{aligned}$$

That is, elements of $SC_n(R,M^i)$ commute modulo $SC_n(R,M^{i+1})$ and thus $SC_n(R,M^i)/SC_n(R,M^{i+1})$ is Abelian. Furthermore since R/M has characteristic p, $p \cdot 1$ is an element of M and hence

$$(I + N_1)^p = I + pN_1 + \cdots$$
$$\equiv I \pmod{M^{i+1}}.$$

Thus elements of $SC_n(R,M^i)/SC_n(R,M^{i+1})$ have order 1 or p.

Completely analogous computations yield that $GC_n(R,M^i)/GC_n(R,M^{i+1})$

is elementary Abelian; since the cardinalities are the same, the groups are isomorphic. Furthermore the decomposition theorem for finite Abelian groups implies $SC_n(R,M^i)/SC_n(R,M^{i+1})$ is a direct product of cyclic groups which must be of order p.

CHAPTER III

SUBGROUPS OF $GL_n(R)$

In this chapter we study in detail two subgroups of $GL_n(R)$. First we determine central series in the kernel K_1 of the morphism $h_M: GL_n(R) \to GL_n(R/M)$ and show that K_1 is a nilpotent group when M is a nilpotent ideal. Secondly we investigate the group S of matrices that are congruent modulo M to upper triangular matrices and show that when R is a finite local ring, S is a p-Sylow subgroup of $GL_n(R)$. We also continue the discussion of characteristic subgroups begun in Chapter II. R will always be used to denote a local ring with maximal ideal denoted by M. Again we assume that $n \geq 3$ and note many of the theorems are true for $GL_2(R)$ but the proofs are different.

The Kernels

Let h_i $(i \geq 0)$ be the group morphism induced by $\rho_i: R \to R/M^i$ on the linear groups; that is

$$h_i : GL_n(R) \to GL_n(R/M^i)$$

where $h_i([a_{ts}]) = [\rho_i(a_{ts})]$. The purpose of the present section is to determine the structure of the kernels: K_i = kernel of h_i, $i = 1,2,\cdots$. It is immediate that K_i has order ideal M^i and consists of all matrices that are congruent to the identity matrix modulo M^i. For generators of

K_i we have the following lemma.

Lemma 3.1. For $i > 0$, K_i is generated by the set $\{B_{kj}(\lambda), D_k(1+\lambda) : k \neq j,\ 1 \leq k, j \leq n,\ \lambda \text{ in } M^i\}$.

Proof: If A is in K_i then A has the form

$$A = \begin{bmatrix} 1+\lambda_{11} & \lambda_{12} & \cdots & \lambda_{1n} \\ \lambda_{21} & 1+\lambda_{22} & \cdots & \lambda_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \lambda_{n1} & \lambda_{n2} & \cdots & 1+\lambda_{nn} \end{bmatrix}$$

where λ_{kj} is in M^i for $1 \leq k, j \leq n$. A series of left multiplications by elementary transvections of the form $B_{kj}(-\lambda_{kj}(1+\lambda_{jj})^{-1})$ will diagonalize A. Then left multiplying by elementary diagonal elements of the form $D_k((1+\lambda_{kk})^{-1})$ will reduce A to I. Thus $A = B(M^i)D(1+M^i)$ where $D(1+M^i)$ denotes a product of $D_k(1+\lambda_{kk})$, λ_{kk} in M^i and $B(M^i)$ denotes a product of elementary transvections each with off diagonal entry from M^i.

When R is a finite local ring with characteristic p^t, p a prime, then K_i is a finite group. In fact $|K_i| = |M^i|^{n^2}$ and thus K_i is a p-group.

Theorem 3.1. (i) For $i \geq 1$ the groups K_i/K_{i+1} are Abelian.

(ii) If M is nilpotent of degree β and R/M has characteristic p then K_i/K_{i+1} is elementary Abelian and for $i \geq \beta/2$, K_i is Abelian.

(iii) If R is finite and R/M has cardinality p^m, then K_i/K_{i+1} is a direct product of $m \cdot n^2 \cdot \dim_{R/M}(M^i/M^{i+1})$ cyclic groups of order p.

Proof: The basis of proof of all three parts is that K_i/K_{i+1} is Abelian if and only if the commutator $[K_i, K_i]$ is contained in K_{i+1}.

The Basic Identities give,

$$[B_{ts}(M^i), B_{kj}(M^i)] \equiv I \pmod{M^i \cdot M^i},$$

$$[D_k(1+M^i), B_{ts}(M^i)] \equiv I \pmod{M^i \cdot M^i},$$

and

$$[D_k(1+M^i), D_j(1+M^i)] = I$$

where $B_{ts}(M^i)$ denotes an arbitrary elementary transvection with an element of M^i in the (t,s) position and $D_k(M^i)$ is analogous.

Thus if $i \geq 1$, we have shown that $[K_i, K_i] \subseteq K_{2i} \subseteq K_{i+1}$ and part (i) is established. Furthermore if M is nilpotent of degree β, then we have for $i \geq \beta/2$ that $K_{2i} = \{I\}$ and hence that $[K_i, K_i] \subseteq \{I\}$. That is if M is nilpotent, then K_i is Abelian for $i \geq \beta/2$.

If R/M has characteristic p, then $p \cdot 1$ is in M. Let $A = I + N$ be an element of K_i, where N is a matrix having all entries from M^i. Then,

$$A^p = (I + N)^p = I + pN + (\text{higher powers of } N).$$

The matrix pN has all entries in M^{i+1} and hence,

$$A^p \equiv I \pmod{M^{i+1}}.$$

We have shown that each element of K_i/K_{i+1} has order p and therefore that K_i/K_{i+1} is elementary Abelian. Part (ii) is now complete.

For part (iii) we have already shown that K_i/K_{i+1} is elementary Abelian, so that when R is finite we conclude immediately that K_i/K_{i+1} is a direct product of finitely many cyclic groups of order p. To determine the number of factors, we count the cardinality of K_i/K_{i+1} obtaining,

$$|K_i/K_{i+1}| = \frac{|M^i|^{n^2}}{|M^{i+1}|^{n^2}} = |M^i/M^{i+1}|^{n^2}.$$

As previously noted, M^i/M^{i+1} is an R/M vector space and therefore has order $(p^m)^{\psi_i}$ where $|R/M| = p^m$ and $\dim_{R/M}(M^i/M^{i+1}) = \psi_i$. Thus,

$$|K_i/K_{i+1}| = (p)^{m\psi_i n^2},$$

and it follows K_i/K_{i+1} is a product of $m \cdot n^2 \cdot \psi_i$ cyclic groups of order p.

Central Series in K_i

Definition 3.1. In a group G a series of normal subgroups, $G = G_0 > G_1 > G_2 > \cdots > G_m$, is called a normal series. Such a series is called a central series if $G_i/G_{i+1} \leq$ center (G/G_{i+1}) for all i. The group G is said to be nilpotent of class m if there is a central series of distinct subgroups such that

$$G = G_0 > G_1 > G_2 > \cdots > G_m = \{1\}.$$

Definition 3.2. In a group G the lower central series is a chain of subgroups, $\{\gamma_i(G)\}$ $i \geq 0$, defined inductively by $\gamma_0(G) = G$ and $\gamma_{i+1}(G) = [\gamma_i(G), G]$.

The lower central series of a group G is indeed a central series and is a minimal central series. Note G is nilpotent of class n if and only if $\gamma_n(G) = \{1\}$. (See Hall, [15]). We notice also that the lower central series is a characteristic series since it is obviously invariant under automorphisms of G.

Theorem 3.2. Let R be a local ring and K_1 the kernel of $h_1 : GL_n(R) \to GL_n(R/M)$.

(i) The lower central series of K_1 is

$$K_1 > SC_n(R,M^2) > SC_n(R,M^3) > \cdots > SC_n(R,M^i) > \cdots.$$

(ii) If M has degree of nilpotency β, then K_1 is a nilpotent group of class $\beta - 1$.

Proof: In the proof of Theorem 3.1 we have already observed that $[K_1,K_1] \leq K_2$. Since elements of $[K_1,K_1]$ have determinant 1, we have $[K_1,K_1] \leq SC_n(R,M^2)$. Conversely if $B_{ij}(\lambda)$ is an elementary transvection of order M^2, then we may assume $\lambda = \lambda_1\lambda_2$, λ_i in M and from Basic Identity (ii),

$$B_{ij}(\lambda_1\lambda_2) = [D_j(1-\lambda_1),B_{ij}(\lambda_2)].$$

Thus $B_{ij}(\lambda_1\lambda_2)$ is in $[K_1,K_1]$ for any (i,j) and hence $SC_n(R,M^2) \leq [K_1,K_1]$. Therefore $SC_n(R,M^2) = [K_1,K_1]$.

Similarly the above equation with λ_1 in M and λ_2 in M^i shows that

$$SC_n(R,M^{i+1}) \leq [K_1,SC_n(R,M^i)].$$

Also, just as in Theorem 3.1, $[K_1,K_i] \leq K_{i+1}$ and therefore,

$$[K_1,SC_n(R_1M^i)] \leq [K_1,K_i] \leq K_{i+1}.$$

Since elements of $[K_1,SC_n(R,M^i)]$ have determinant 1, we have that

$$[K_1,SC_n(R,M^i)] \leq SC_n(R,M^{i+1})$$

and thus $[K_1,SC_n(R,M^i)] = SC_n(R,M^{i+1})$ for $i \geq 2$. The proof of part (i) is now complete.

For part (ii) we merely note that $M^\beta = 0$ and $M^{\beta-1} \neq 0$. Thus $SC_n(R,M^{\ }) = \{I\}$ and $SC_n(R,M^{\beta-1}) \neq \{I\}$; that is, the lower central series stops with $SC_n(R,M^\beta)$ and hence K_1 is a nilpotent group of class $\beta - 1$.

Corollary 1. The lower central series of $K_t (t \geq 1)$ is $\{K_t,\{SC_n(R,M^{st}) : s = 2,3,\cdots\}\}$.

Proof: As above,

$$SC_n(R,M^{2t}) \leq [K_t,K_t] \leq K_{2t}$$

and since elements of $[K_t,K_t]$ have determinant 1,

$$SC_n(R,M^{2t}) = [K_t,K_t].$$

Likewise $[SC_n(R,M^{(s-1)t}),K_t] = SC_n(R,M^{st})$.

Letting [t] denote the greatest integer function, we have the following corollary:

Corollary 2. If M is nilpotent of degree β then K_t is a nilpotent group of class $\left[\frac{\beta-1}{t}\right]$.

Proof: The smallest non-trivial subgroup in the series of Corollary 1 is $SC_n(R,M^{st})$ where s is the greatest integer such that $st \leq \beta - 1$.

The derived group of a group G is the commutator subgroup [G,G] and is denoted by G'. The derived series in a group G is defined inductively by $G_o = G$ and $G_i = G'_{i-1}$, $i \geq 1$.

Corollary 3. The derived series for K_1 is $\{K_1,\{SC_n(R,M^{2^k}): k \geq 1\}\}$.

Proof: We already have $K_1' = SC_n(R,M^2)$. Now

$$SC_n(R,M^{2^k}) \leq [SC_n(R,M^{2^{k-1}}), SC_n(R,M^{2^{k-1}})]$$

$$\leq K_{2^{k-1}}.$$

As noted before, elements of the commutator group have determinant 1 and thus,

$$SC_n(R,M^{2^k}) = [SC_n(R,M^{2^{k-1}}), SC_n(R,M^{2^{k-1}})].$$

Definition 3.3. Using Z(G) to denote the center of a group G, define the upper central series in G inductively by $Z_o = \{1\}$, $Z_1 = Z(G)$, and $Z_{i+1}(G)$ is the preimage of $Z(G/Z_i(G))$ under the natural epimorphism

$G \to G/Z_i(G)$.

The series $\{1\} < Z_1(G) < Z_2(G) < \cdots$ is a central series and, in fact, a maximal central series. Further, G is a nilpotent group if and only if the upper central series reaches G; G is nilpotent of class m if $Z_m(G) = G$. Observe that $Z_{i+1}(G)$ can be characterized in terms of commutators in the following way:

$$Z_{i+1}(G) = \{X \in G \mid [X,A] \text{ is in } Z_i(G) \text{ for all } A \text{ in } G\}.$$

If R is a local ring and J an ideal of R, the annihilator of J, $\{r \in R : rJ = 0\}$, is denoted by A(J) and is an ideal of R. We shall be interested in the groups $H_i \leq K_1$ defined by

$$H_i = \{X \in K_1 : X \text{ is scalar modulo } A(M^i),\ i \geq 1\}$$

and $H_o = \{I\}$. Note that when M is nilpotent of degree β, $A(M^i)$ contains the ideal $M^{\beta-i}$ but contains no larger power of M. Furthermore when M is nilpotent and $R = Z/Zp^n$, we have that $A(M^i) = M^{\beta-i}$.

In terms of congruence subgroups,

$$H_i = K_1 \cap GC_n(R,A(M^i)),\ i \geq 1.$$

Lemma 3.2. For $i \geq 1$, H_i is the subgroup of K_1 generated by all $B_{jk}(\lambda)$, λ in $A(M^i)$ and all products $\prod_{j=1}^{n} D_j(1 - \lambda_{jj})$ for λ_{jj} in M and $\lambda_{jj} - \lambda_{kk}$ in $A(M^i)$.

Proof: Clearly for $i \geq 1$, H_i contains these generators. If $[a_{ts}]$ is in H_i, then for $t \neq s$ $a_{ts} \equiv 0 \pmod{A(M^i)}$ and $a_{tt} \equiv 1 \pmod{M}$. Thus $[a_{ts}]$ reduces to a diagonal matrix, $[1-\lambda_{11},\cdots,1-\lambda_{nn}]$, under left multiplication by elementary transvections $B_{ij}(A(M^i))$. Then since $[a_{ts}]$ is scalar modulo $A(M^i)$, it follows that $\lambda_{jj} - \lambda_{kk} \equiv 0 \pmod{A(M^i)}$. That is,

$$[a_{ts}] = \prod_{i \neq j} B_{ij}(\lambda_{ij}) \prod_{j=1}^{n} D_j(1 - \lambda_{kk})$$

where the λ_{ij}'s are in $A(M^i)$, the λ_{kk}'s are in M, and $\lambda_{kk} - \lambda_{jj}$ is in $A(M^i)$ for all j, k.

Theorem 3.3. The series,

$$\{I\} < H_1 < H_2 < \cdots < H_i < \cdots,$$

is the upper central series for K_1. Furthermore if M is nilpotent of degree β, then $H_{\beta-1} = K_1$.

Proof: We start by showing that for $i \geq 1$ each element of H_i maps to the center of K_1/H_{i-1}. For this it suffices to show that for each generator X of H_i and each generator A of K_1 the commutator [X,A] is in H_{i-1}. For the generators $B_{ts}(\lambda)$ of K_1 and $\prod_{j=1}^{n} D_j(1 - \lambda_j)$ of H_i using the Basic Identity (i) we have,

$$\left[\prod_{j=1}^{n} D_j(1-\lambda_j),\ B_{ts}(\lambda)\right] = B_{ts}(\lambda) - \lambda\left(\frac{1-\lambda_s}{1-\lambda_t}\right) E_{ts}$$

$$= I + \lambda\left(\frac{\lambda_s - \lambda_t}{1 - \lambda_t}\right) E_{ts}$$

$$\equiv I \pmod{A(M^{i-1})}.$$

The latter congruence follows from the facts that $\lambda_s - \lambda_t \equiv 0 \pmod{A(M^i)}$ and λ is in M. Thus the commutator is scalar modulo $A(M^{i-1})$ and is therefore in H_{i-1}.

For the generators $B_{st}(\mu)$ of H_i and $B_{jk}(\lambda)$ of K_1, by Basic Identities (iii) - (v), for each choice of (s,t) and (j,k) there is a matrix N such that,

$$[B_{st}(\mu), B_{jk}(\lambda)] = I + \mu\lambda N;$$

however, $\mu \equiv 0 \pmod{A(M^i)}$ and λ is in M. Therefore the commutator is congruent to I mod $A(M^{i-1})$ and thus is in H_{i-1}.

For $D_k(1-\lambda)$ generating $K_1(\lambda$ in M) and $B_{st}(\mu)$ generating H_i (μ in $A(M^i)$) by Basic Identity (ii),

$$[D_k(1-\lambda), B_{st}(\mu)] \equiv I \ (A(M^{i-1}));$$

hence the commutator is in H_{i-1}.

Since commutators of two diagonal matrices are always I, we have shown that all commutators of combinations of generators of H_i with generators of K_1 are in H_{i-1}. Thus $H_i \leq$ preimage $(Z(K_1/H_{i-1}))$ for $i \geq 1$. In particular, $H_1 \leq Z(K_1)$.

Conversely we suppose X maps to the center of K_1/H_{i-1} and show X must be in H_i. The proof is a calculation based on the idea that if we know what X commutes with, then we know X. If X maps to the center of K_1/H_{i-1}, then X commutes modulo H_{i-1} with all elementary transvections of K_1. That is, if λ is in M, then $[X, B_{jk}(\lambda)]$ is scalar mod $A(M^{i-1})$ for all possible choices of (j,k). Let $X = [a_{jk}]$, $X^{-1} = [\alpha^{jk}]$, and recall the Basic Identity (i),

$$[X, B_{jk}(\lambda)] = B_{jk}(\lambda) - \lambda^2 a_{kj} \sum_t \alpha^{tj} E_{tk} - \lambda \sum_{t,s} \alpha^{tj} a_{ks} E_{ts}.$$

If $[X, B_{jk}(\lambda)]$ is in H_{i-1}, it is scalar modulo $A(M^{i-1})$ and congruent to I modulo M. The following congruences must hold:

(i) For the (j,k) position,

$$\lambda(1 - a_{kk}\alpha^{jj} - \lambda a_{kj}\alpha^{jj}) \equiv 0 \pmod{A(M^{i-1})};$$

(ii) For the (t,k) position, $t \neq j$ and $t \neq k$,

$$\lambda\alpha^{tj}(a_{kk} + \lambda a_{kj}) \equiv 0 \pmod{A(M^{i-1})};$$

(iii) For the (t,s) position, $s \neq k$, $t \neq s$,

$$\lambda\alpha^{tj}a_{ks} \equiv 0 \pmod{A(M^{i-1})};$$

(iv) For diagonal positions, $t \neq k$,

$$\lambda\alpha^{kj}(a_{kk} + \lambda a_{kj}) \equiv \alpha^{tj}a_{kt} \pmod{A(M^{i-1})}.$$

Since λ is arbitrary in M these reduce to the following congruences modulo $A(M^i)$:

(i) $(1 - a_{kk}\alpha^{jj} - \lambda a_{kj}\alpha^{jj}) \equiv 0 \pmod{A(M^i)}$,

(ii) $\alpha^{tj}(a_{kk} + \lambda a_{kj}) \equiv 0 \pmod{A(M^i)}$,

(i i) $\alpha^{tj}a_{ks} \equiv 0 \pmod{A(M^i)}$.

(iv) $\alpha^{kj}(a_{kk} + \lambda a_{kj}) \equiv \alpha^{tj}a_{kt} \pmod{A(M^i)}$,

with the same restrictions on the indices as stated previously.

The congruences (i) - (iv) are sufficient to show X must be in H_i. Since X is in K_1, a_{kk} is a unit; hence, $a_{kk} + \lambda a_{kj}$ is a unit and (ii) then implies that α^{tj} is in $A(M^i)$ for all t except $t = k$ and $t = j$. But from (iv) we conclude α^{kj} is in $A(M^i)$. Also (iii) implies $a_{ks} \equiv 0 \pmod{A(M^i)}$ except for a_{kj} and a_{kk}; but again, (iv) implies $a_{kj} \equiv 0 \pmod{A(M^i)}$. Since the indices j, k are arbitrary, we now have all off-diagonal elements of X in $A(M^i)$. For diagonal elements we use the first congruence,

$$a_{kk}\alpha^{jj} \equiv 1 - \lambda a_{kj}\alpha^{jj} \pmod{A(M^i)}.$$

But a_{kj} is in $A(M^i)$ and thus,

$$a_{kk}\alpha^{jj} \equiv 1 \pmod{A(M^i)}.$$

Letting k vary, we conclude that

$$a_{kk} - a_{k'k'} \equiv 0 \pmod{A(M^i)}$$

for $1 \leq k, k' \leq n$. Thus we have shown that if $X = [a_{jk}]$ maps to $Z(K_1/H_{i-1})$, then a_{jk} is in $A(M^i)$ and $a_{jj} - a_{kk}$ is in $A(M^i)$ for any $j \neq k$, which is precisely the condition for X to be in H_i. We have shown H_i is the preimage of $Z(K_1/H_{i-1})$ under the natural projection and hence, that the upper central series is

$$\{I\} < H_1 < H_2 < \cdots < H_i < \cdots .$$

If M has degree of nilpotency β, then $A(M^{\beta-1})$ contains M and is not all of R. Thus $A(M^{\beta-1}) = M$, and $H_{\beta-1} = K_1$.

Corollary 1. The center of K_1 consists of all elements of K_1 that are scalar modulo A(M).

Corollary 2. When M is nilpotent,

$$SC_n(R,M^i) \leq K_i \leq H_{\beta-i} .$$

Corollary 3. The upper central series in K_t is $\{I, \{K_t \cap H_{ts} : s = 1,2,3,\cdots\}\}$.

The Upper Triangular Subgroup

Weir [28] investigated the subgroup of upper triangular matrices in $GL_n(F)$ where F is a finite field with p^t elements. He showed that the subgroup is a p-Sylow subgroup, determined basic generators, and computed central series. In $GL_n(Z/Zp^n)$ the subgroup of matrices congruent modulo (p) to upper triangular has been shown by Barbee [3] to be a p-Sylow subgroup. E. Robertson [22] has shown that over any ring with 1 and with nilpotent Jacobson radical the group of matrices congruent modulo

the Jacobson radical to upper triangular matrices is generated by its normal subgroups, that is, is a Fitting group.

Throughout this section R is a local ring with nilpotent maximal ideal M of nilpotency β. We also assume that the characteristic of R is not a power of 2, equivalently, that 2 is a unit. Define S to be the set of elements of $GL_n(R)$ that are congruent modulo M to the set of upper triangular matrices having 1 in each main diagonal position. Schematically, an element X of S has the form,

$$X = \begin{bmatrix} 1+m_{11} & r_{12} & r_{13} & \cdot & \cdot & r_{1n} \\ m_{21} & 1+m_{22} & r_{23} & \cdot & \cdot & r_{2n} \\ m_{31} & m_{32} & 1+m_{33} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & r_{n-1,n} \\ m_{n1} & m_{n2} & \cdot & \cdot & m_{n,n-1} & 1+m_{nn} \end{bmatrix},$$

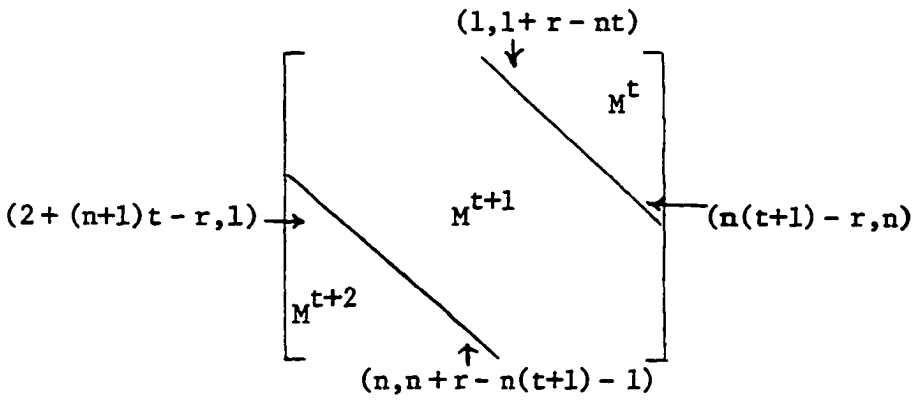
where r_{ij} denotes an element of R and m_{ij} denotes an element of M. S is obviously a subgroup of $GL_n(R)$ and X is a product of $B_{ij}(M)$ $(j - i < 0)$, $B_{ij}(R)$ $(j - i > 0)$, and $D_K(1 + M)$ $(k \geq 1)$.

To generate central series in S we let N_r denote the collection of n by n matrices $[a_{ij}]$ for which,

$$a_{ij} \equiv \begin{cases} 0 \pmod{M^t} \\ 0 \pmod{M^{t+1}} \text{ when } j - i < r - nt \\ 0 \pmod{M^{t+2}} \text{ when } j - i < r - n(t+1) \end{cases}$$

where $t = \left[\frac{r}{n}\right]$ (the greatest integer).

An element of N_r has the form,



where M^q denotes "entries from M^q". The positions labeled indicate the diagonals by which the triangles are bounded.

It is obvious that $N_{r+1} \subset N_r$; since $M^\beta = 0$, we have $N_{n\beta} = 0$. Furthermore $N_{n\beta-1} \neq 0$. For $1 \leq r \leq n_\beta - 1$, note that N_r and N_{r+1} differ only in the diagonals where $j - i = r - nt$ and $j - i = r - n(t+1)$ with N_{r+1} having entries form one higher power of M in these positions.

Lemma 3.3. If X_r is in N_r and Y_s is in N_s, then $X_r Y_s$ is in N_{r+s}.

Proof: Let $X_r = [x_{ij}]$, $Y_s = [y_{ij}]$ and $Z = X_r \cdot Y_s = [z_{ij}]$. Also let $t = \left[\frac{r}{n}\right]$, $u = \left[\frac{s}{n}\right]$ and $w = \left[\frac{r+s}{n}\right]$. Note that $w = t+u$ or $w = t+u+1$. We have $z_{ij} = \sum_{k=1}^{n} x_{ik}y_{kj}$ and we show that z_{ij} is in the proper power of M depending upon the difference $j - i$. Thinking of the usual matrix multiplication, z_{ij} is a product of the row vector $(x_{i1},\cdots,x_{in})$ with the column vector $\begin{pmatrix} y_{ij} \\ \vdots \\ y_{nj} \end{pmatrix}$; therefore, with respect to powers of M, the product has the form,

$$(m^{t+2},\dots,\underset{i+r-n(t+1)}{m^{t+2}},m^{t+1},\dots,\underset{i+r-nt}{m^{t+1}},m^{t}\dots m^{t}) \begin{pmatrix} m^{s} \\ \vdots \\ m^{s} \\ m^{s+1} \\ \vdots \\ m^{s+1} \\ m^{s+2} \\ m \\ \vdots \\ m^{s+2} \end{pmatrix} \begin{matrix} \\ \\ \leftarrow j-s+nu \\ \\ \\ \\ \leftarrow j-s+n(u+1) \\ \\ \\ \\ \end{matrix}$$

If $j - i < r+s - n(t+u)$, then $i+r - nt > j - s + nu$ and the above product is in M^{t+1}. If $j-i < (r+s) - n(t+u+1)$, then $i+r - n(t+1) > j - s + nu$ and $i+r - nt > j - s + n(u+1)$ and the product is in M^{t+2}. Finally, if $j - i > r + s - n(t+u+2)$, then $i+r - n(t+1) > j-s + n(u+1)$ and the product is in M^{t+3}. Hence Z is in N_{r+s} and the proof is complete.

For $1 \le r \le n_\beta$, let S_r denote the set $\{I + X_r : X_r \text{ is in } N_r\}$. Note $S_1 = S$ and if $s \ge r$, $S_s \subseteq S_r$.

Lemma 3.4. For $1 \le r \le n_\beta$, S_r is a normal subgroup of S.

Proof: Let $I + X_r$ and $I + Y_r$ be elements of S_r with X_r and Y_r in N_r. Then

$$(I + X_r)(I + Y_r) = I + X_r + Y_r + X_rY_r.$$

By Lemma 3.3 X_rY_r is in N_r and thus the product is in S_r. Then S_r is a subgroup. To consider normality we let $I + X_1$ be in S_1 with $(I + X_1)^{-1} = I + Y_1$ and Y_1, X_1 in N_1. Then

$$\begin{aligned}(I + X_1)^{-1}(I + X_r)(I + X_1) &= \\ &= (I + X_1)^{-1}(I + X_1 + X_r + X_rX_1) \\ &= I + X_r + X_rX_1 + Y_1X_r + Y_1X_rX_1 .\end{aligned}$$

Again by Lemma 3.3 the last sum is in S_r and therefore S_r is normal.

We thus have a normal series in S,

$$S = S_1 > S_2 > S_3 > \cdots > S_{n\beta-1} > S_{n_\beta} = \{I\}.$$

We shall show by restricting to subgroups of elements having determinant 1 that,

$$S_1 > S_2|_{\det 1} > S_3|_{\det} > \cdots > S_{n_{\beta-1}}|_{\det 1} > S_{n_\beta} = \{I\}$$

is the lower central series for S.

Theorem 3.4. Let $t = \left[\frac{r}{n}\right]$ and $1 < r < n_{\beta-1}$.

(i) If $r \not\equiv 0 \pmod n$ then S_r is generated by S_{r+1} and $\{B_{ij}(\lambda) : j - r = r - nt$ and λ in M^t or $j - i = r - n(t+1)$ and $\lambda \in M^{t+1}\}$.

(ii) If $r \equiv 0 \pmod n$, S_r is generated by S_{r+1} and $\{D_k(1 - \lambda) \mid \lambda \in M^t,\ k = 1,\cdots,n\}$.

Proof: Clearly S_r contains the asserted generating sets in both (i) and (ii). Moreover the only way an element of S_r can fail to be in S_{r+1} is by having the $(i,i+r-nt)$ or the $(i,i+r-n(t+1))$ entries in M^t or M^{t+1} respectively, or in the diagonal case (ii), the main diagonal entries in M^t. Thus in (i) the elementary transvections and S_{r+1} generate S_r.

Corollary: (i) If $r \not\equiv 0 \pmod n$ then $S_r|_{\det 1}$ is generated by $S_{r+1}|_{\det 1}$ and the elementary transvections in part (i) of the theorem.

(ii) If $r \equiv 0 \pmod n$, $S_r = S_{nt}$ is generated by S_{nt+1} and $\{D_k\left(\frac{1}{1-M^t}\right)D_{k+1}(1 - M^t) : k = 1,\cdots,n-1\}$.

Proof: Part (i) is exactly as in the theorem. In the $r = nt$ case we merely replace $D_k(1-M^t)$ by $D_k\left(\frac{1}{1-M^t}\right) D_{k+1}(1-M^t)$ to keep the determinant equal to one.

Theorem 3.5. Let $1 \le r \le n_{\beta-1}$. Then

$[S_r,S_1] = [S_r|_{\det 1},S_1] = S_{r+1}|_{\det 1}$. Furthermore if $1 \le r, s \le n_{\beta-1}$ and $r + s \not\equiv 0 \pmod{n}$, then $[S_r|_{\det 1},S_s|_{\det 1}] = S_{r+s}|_{\det 1}$.

Proof: We begin by showing that for any $1 < r, s < n\beta$

$$[S_r,S_s] < S_{r+s}|_{\det 1}.$$

Let $I + X_r$ be in S_r and $I + Y_s$ be S_s where X_r is in N_r and Y_s is in N_s. Since M is nilpotent, the following are finite:

$$(I + X_r)^{-1} = I - X_r + X_r^2 + \cdots$$

$$(I + Y_r)^{-1} = I - Y_s + Y_s^2 + \cdots \quad .$$

Thus we have

$$[I + X_r,I + Y_s] = (I + X_r)^{-1}(I + Y_s)^{-1}(I + Y_s + X_r + X_rY_s)$$

$$= (I + X_r)^{-1}(I + X_r - Y_sX_r + X_rY_s - 0_{r+s})$$

$$= (I + 0_{2r+s} + 0_{r+s}),$$

where 0_q denotes "something in N_q". The last term is in $S_{r+s}|_{\det 1}$ and thus $[S_r,S_s] \le S_{r+s}|_{\det 1}$.

Conversely we show that generators of $S_{r+1}|_{\det 1}$ are in $[S_r|_{\det 1},S_1]$. Since $[S_{n_{\beta-1}},S_1] = \{I\} = S_n$, we may use induction and thus it is only necessary to exhibit that the generators of $S_{r+1}|_{\det 1}$ which are not in $S_{r+2}|_{\det 1}$ are commutators of the proper form. If $r+1 \not\equiv 0(n)$ and $t = \left[\frac{r}{n}\right]$, then

$$B_{i,i+(r+1)-n(t)}(M^t) = [B_{i,i+r-nt}(M^t),B_{i+r-nt,i+(r+1)-n(t)}(1)].$$

By noting t may be replaced throughout by t + 1, we have the necessary

generators. If $r + 1 \equiv 0(n)$, we have

$$D_k \quad \frac{1}{1-M^{t+1}} D_{k+1}(1-M^{t+1})$$

$$= B_{k+1}\left(\frac{-M^{t+1}}{1-M^{t+1}}\right)\left[B_{k,k+1}(1), B_{k+1,k}(M^{t+1})\right] B_{k+1,k}\left(\frac{M^{2(t+1)}}{1-M^{t+1}}\right).$$

The first and last factors on the above are in S_{r+2}. Thus generators of $S_{r+1}\big|_{\det 1}$ are in $[S_r\big|_{\det 1}, S_1]$. Hence,

$$S_{r+1}\big|_{\det 1} \leq [S_r\big|_{\det 1}, S_1] \leq [S_r, S_1] \leq S_{r+1}\big|_{\det 1}.$$

Thus all three are equal which is the first part of the theorem.

To complete the proof we need to show that if $r + s \not\equiv 0 \pmod n$, any generator of $S_{r+s}\big|_{\det 1}$ is in $[S_r\big|_{\det 1}, S_s\big|_{\det 1}]$. If $u = \left[\frac{s}{n}\right]$ and $w = \left[\frac{r+s}{n}\right]$, then $w = t + u$ or $w = t + u + 1$. Just as above,

$$B_{i,i+(r+s)-n(t+u)}(M^{t+u}) =$$

$$[B_{i,i+r-nt}(M^t), B_{i+r-nt,i+(r+s)-n(t+u)}(M^u)]$$

and the commutator is in $[S_r\big|_{\det 1}, S_s\big|_{\det 1}]$. Since t may be replaced by $t + 1$ or u by $u + 1$, we have that any generator of $S_{r+s}\big|_{\det 1}$ is a commutator of the asserted form. Combining with the initial results we have,

$$S_{r+s}\big|_{\det 1} \leq [S_r\big|_{\det 1}, S_s\big|_{\det 1}] \leq [S_r, S_s] \leq S_{r+s}\big|_{\det 1}.$$

Corollary 1: The lower central series in S is

$$\{S = S_1, \{S_r|_{\det 1} : 2 \leq r \leq n\beta\}\}.$$

Corollary 2: S is nilpotent of class $n\beta - 1$.

Corollary 3: If n is odd, then the derived series in S is $\{S, \{S_{2^k}|_{\det 1} : k \geq 1\}\}$.

Proof: $[S_1,S_1] = S_2|_{\det 1}$ and since n is odd we have $2^k \nmid 0(n)$. Thus $[S_{2^k}|_{\det 1}, S_{2^k}|_{\det 1}] = S_{2\cdot 2^k}|_{\det 1} = S_{2^k + 1}|_{\det 1}$.

In Weir's treatment of S over a finite field, the subgroup S_r corresponds to matrices having 1 along the diagonal and zeros elsewhere for $j - i < r$. He shows that $[S_r,S_s] = S_{r+s}$ for all r, s and thus determines the derived series in all cases. This however is not in general the case as we now show.

Observation 3.1: If $r = nt$ and $s = nu$, then $[S_r|_{\det 1},S_s|_{\det 1}] = [S_r,S_s] = S_{r+s+1}|_{\det 1}$.

Proof: From the last theorem we have that

$$S_{r+s+1}|_{\det 1} = [S_{r+1}|_{\det 1}, S_s|_{\det 1}] \leq [S_r|_{\det 1}, S_s|_{\det 1}].$$

However it is an easy computation to show that if X_r is in N_r and Y_s is in N_s, then $X_rY_s - Y_sX_r$ is in N_{r+s+1} when $r = nt$ and $s = nu$. Then from the calculation made in the beginning of the last proof we have that $[I + X_r,I + Y_s]$ is in S_{r+s+1}. Thus

$$[S_r,S_s] \leq S_{r+s+1}|_{\det 1} \leq [S_r|_{\det 1}, S_s|_{\det 1}].$$

Therefore,

$$[S_r, S_s] = [S_r|_{\det 1}, S_s|_{\det 1}] = S_{r+s+1}|_{\det 1}.$$

The derived series in S has not been determined when R is local but not a field.

We turn now to the case where R is a finite local ring and assume that $|R/M| = q = p^r$. It was shown in Chapter I that

$$|GL_n(R)| = |M|^{n^2} \cdot q^{\frac{n(n-1)}{2}} \prod_{i=1}^{n} (q^i - 1).$$

Theorem 3.6. If R is a finite local ring and $|R/M| = q = p^r$, p a prime, then S is a p - Sylow subgroup of $GL_n(R)$.

Proof: It suffices to show that S is a maximal p-subgroup of $GL_n(R)$ by counting $|S|$. There are $\frac{n^2-n}{2}$ positions above the diagonal. Each of these can be filled in $|R|$ ways. The remaining $n^2 - \frac{n^2-n}{2} = \frac{n^2+n}{2}$ positions can be filled in $|M|$ ways. Thus

$$|S| = |M|^{\frac{n^2+n}{2}} \cdot |R|^{\frac{n^2-n}{2}} = |M|^{n^2} \cdot |R/M|^{\frac{n^2-n}{2}}$$

$$= |M|^{n^2} q^{\frac{n(n-1)}{2}}.$$

Therefore S is a p-subgroup and must be a maximal one since p does not divide $\prod_{i=1}^{n} (q^i - 1)$ and $|GL_n(R)| = |S| \cdot \prod_{i=1}^{n} (q^i - 1)$.

Characteristic Subgroups

Let R be a finite local ring with maximal ideal M. We have shown at the end of Chapter II that $SL_n(R)$, $GC_n(R,M)$, and $SC_n(R,M)$ are characteristic subgroups of $GL_n(R)$. In this section, by elementary counting techniques, we extend the collection of characteristic subgroups

to include K_1 and $SC_n(R,M^i)$ for $0 \leq i \leq \beta$. The following two theorems occur in Gorenstein's, Finite Groups [14] (p. 16).

Theorem 3.7. Let H and K be subgroups of the group G. Suppose H is characteristic in K and K is characteristic in G, then H is characteristic in G.

Theorem 3.8. If H is a normal subgroup of G whose index and order are relatively prime, then H is characteristic in G.

The two theorems will show that K_1 is a characteristic subgroup.

Theorem 3.9. If R is a finite local ring then the kernel K_1 is a characteristic subgroup of $GL_n(R)$.

Proof: We have already counted K_1 and $GC_n(R,M)$,

$$|K_1| = |M|^{n^2}, \quad |GC_n(R,M)| = |R^*| \cdot |M|^{n^2-1}.$$

Furthermore, $|R^*| = |R| \cdot \left(1 - \frac{1}{|R/M|}\right)$ and hence $|R^*| = |M| \cdot (|R/M| - 1)$. Thus for the index of K_1 in $GC_n(R,M)$,

$$[GC_n(R,M):K_1] = \frac{|M|^{n^2}(|R/M| - 1)}{|M|^{n^2}}$$

$$= |R/M| - 1.$$

But the order of K_1, $|M|^{n^2}$, is a power of the prime characteristic of R/M. Therefore the order of K_1 and the index of K_1 are relatively prime. Hence by Theorem 3.8 K_1 is characteristic in $GC_n(R,M)$ and by Theorem 3.7 is therefore characteristic in $GL_n(R)$.

Theorem 3.10. If R is a finite local ring, then the subgroups $SC_n(R,M^i)$ are characteristic in $GL_n(R)$ for $1 \leq i \leq \beta$.

Proof: We have already shown that $SL_n(R)$ and $SC_n(R,M)$ are

characteristic. For $1 < i$, $SC_n(R,M^i)$ is characteristic because it appears in the lower central series of the characteristic subgroup K_1. That is,

$$SC_n(R,M^2) = [K_1,K_1]$$

and

$$SC_n(R,M^i) = [K_1,SC_n(R,M^{i-1})], \quad i \geq 3.$$

Thus, as commutators of characteristic subgroups, the groups $SC_n(R,M^i)$ are characteristic in K_1 and hence in $GL_n(R)$.

We conclude the chapter with the following observation:

Observation 3.2. Automorphisms of $GL_n(R)$ induce automorphisms of $GL_n(R/M)$ through the natural surjection h_M.

Proof: Let Λ be an automorphism of $GL_n(R)$. Define,

$$\overline{\Lambda} : GL_n(R/M) \to GL_n(R/M)$$

by $\overline{\Lambda}(h_M(A)) = h_M(\Lambda(A))$. The map $\overline{\Lambda}$ is a well-defined automorphism because the kernel of h_M, K_1, is characteristic.

The automorphisms of $GL_n(R/M)$, when R/M is a finite field, have been described in [11] by J. Dieudonné. The above observation allows many of the results known for finite fields to be lifted back to R. However there seems to be no simple or direct way to show K_1 is a characteristic subgroup in the case of an infinite local ring and the automorphisms are somewhat more elusive.

CHAPTER IV

AUTOMORPHISMS

In this chapter we determine the automorphisms of the matrix ring over a local ring and the automorphisms of the general linear group over a local ring. For the latter we require that the characteristic of R/M be other than 2 and that the matrices have dimension n by n, $n \geq 3$.

Automorphisms of $M_n(R)$

If S is a ring let Aut(S) denote the group of ring automorphisms of S, that is, the group of all $\Lambda: S \to S$ such that Λ is a ring isomorphism. Let R be a local ring with maximal ideal M and let $M_n(R)$ denote the ring of all n by n matrices over R. Note that Aut(R) acts as a transformation group on $M_n(R)$ where if A is in $M_n(R)$ and α is in Aut(R), then A^α is the matrix determined by applying α to each of the entries. We shall show that

$$\mathrm{Aut}(M_n(R)) \cong GL_n(R) \dot{\times} \mathrm{Aut}(R)$$

where $\dot{\times}$ denotes the semi-direct product, that is, $GL_n(R) \dot{\times} \mathrm{Aut}(R)$ consists of order pairs with composition defined by,

$$(A,\alpha)(B,\beta) = (AB^\alpha, \alpha\beta).$$

We call a set $\{F_{ij}\}_{i,j=1}^{n}$ of elements of $M_n(R)$ a complete set of matrix units if

(i) $F_{ij} = F_{km}$ implies $(i,j) = (k,m)$

(ii) $F_{ij}F_{km} = \delta_{jk}F_{im}$

(iii) $\sum_{i=1}^{n} F_{ii} = I.$

Note that the collection $\{E_{ij}\}_{i,j=1}^{n}$, where E_{ij} has zeros everywhere except in the (i,j) position where a 1 appears, is a complete set of matrix units. $\{E_{ij}\}_{i,j=1}^{n}$ is called the standard set of matrix units.

Lemma 4.1. Let R be a local ring and $\{F_{ij}\}_{i,j=1}^{n}$ a complete set of matrix units in $M_n(R)$. Then there is a Q in $GL_n(R)$ such that

$$QF_{ij}Q^{-1} = E_{ij}, \quad 1 \leq i, j \leq n.$$

Proof: Let e_i denote the vector having 1 in the i-th position and zeros elsewhere. The set $\{e_i\}_{i=1}^{n}$ is a free R-basis for the n-dimensional free module $\sum_{i=1}^{n} R = R^{(n)}$. It suffices to show that there is a basis $\{b_i\}_{i=1}^{n}$ of $R^{(n)}$ and a change of basis matrix Q such that $Q(b_i) = e_i$ and

$$QF_{ij}Q^{-1} = E_{ij}.$$

Since $\sum_{i=1}^{n} F_{ii} = I$ we know that some F_{ii} has a unit in the (1,1) position and we may assume it is F_{11}. In that case, if

$$F_{11}(e_1) = \sum_{i=1}^{n} r_i e_i,$$

then r_1 is a unit of R. Now define $b_i = F_{i1}e_1$, $1 \leq i \leq n$. To see that $\{b_i\}_{i=1}^{n}$ is a free R-basis it suffices to show $b_1, \cdots, b_n$ are linearly independent. Assume $\sum_{i=n}^{n} s_i b_i = 0$. Then for any j,

$$0 = F_{ij} \sum_{i=1}^{n} s_i b_i = \sum_{i=1}^{n} s_i F_{ij} b_i$$

$$= \sum_{i=1}^{n} s_i F_{ij} F_{l1} e_l = \sum_{i=1}^{n} s_i \delta_{ij} F_{l1} e_l$$

$$= s_j F_{l1} e_l = \sum_{i=1}^{n} s_j r_i e_i.$$

But then $s_j r_1 = 0$ and since r_1 is a unit, s_j must be 0. Thus $\{b_i\}_{i=1}^{n}$ is a free R-basis for $R^{(n)}$.

Now define Q to be the matrix such that $Q(b_i) = e_i$. The matrix Q is invertible since it is a change of basis matrix. We now compute $QF_{ij}Q^{-1}$. First note that for any k, $E_{ij}(e_k) = \delta_{jk} e_i$. On the other hand,

$$QF_{ij}Q^{-1}(e_k) = QF_{ij}(b_k)$$

$$= QF_{ij}(F_{kl}e_l)$$

$$= Q(\delta_{jk}F_{il}e_l)$$

$$= Q(\delta_{jk}b_i)$$

$$= \delta_{jk}e_i.$$

Therefore $QF_{ij}Q^{-1} = E_{ij}$ and the proof is complete.

Theorem 4.1. Let R be a local ring. If Λ is an element of $\mathrm{Aut}(M_n(R))$ then there is a Q in $GL_n(R)$ and an α in $\mathrm{Aut}(R)$ such that

$$\Lambda A = Q^{-1}A^{\alpha}Q \quad \text{for all A in } M_n(R).$$

Proof: Let $F_{ij} = \Lambda E_{ij}$ and note that the collection $\{F_{ij}\}_{i,j=1}^{n}$ is a complete set of matrix units. By Lemma 4.1 there is a Q in $GL_n(R)$ such that

$$QF_{ij}Q^{-1} = E_{ij}.$$

Letting $\bar{\Lambda}$ be the automorphism $\bar{\Lambda} = Q\Lambda(\ )Q^{-1}$ we have that $\bar{\Lambda}E_{ij} = E_{ij}$. Let rI be a scalar matrix and note that $\bar{\Lambda}(rI)$ commutes with all E_{ij} and therefore must be scalar. Now define α: R → R by $\alpha(r) = s$ if and only if $\bar{\Lambda}(rI) = sI$. α is obviously a ring automorphism because $\bar{\Lambda}$ is a ring automorphism. Now we have $\bar{\Lambda}(rI\ E_{ij}) = (rI)^{\alpha}E_{ij}$, equivalently $\Lambda(rI\ E_{ij}) = Q^{-1}(rI)^{\alpha}E_{ij}Q$. Finally, any A in $M_n(R)$ has unique representation

$$A = \sum(a_{ij}I)E_{ij}$$

and hence

$$\begin{aligned}\Lambda A &= \sum \Lambda(a_{ij}I)E_{ij} \\ &= \sum Q^{-1}(a_{ij}I)^{\alpha}E_{ij}Q \\ &= Q^{-1}(\sum(a_{ij}I)^{\alpha}E_{ij})Q \\ &= Q^{-1}A^{\alpha}Q.\end{aligned}$$

Corollary 1. $\mathrm{Aut}(M_n(R)) \cong GL_n(R) \dot{x} \mathrm{Aut}(R)$

Proof: Define a map by

$$\Lambda \to (Q,\alpha)$$

where the Q and α are described in the theorem. We need only check the multiplication of pairs. Suppose

$$\Lambda_1 \to (P_1,\alpha_1)$$

$$\Lambda_2 \to (P_2,\alpha_2)\ ,$$

then

$$\begin{aligned}\Lambda_1\Lambda_2(A) &= (P_1,\alpha_1)(P_2,\alpha_2)(A)\\ &= (P_1,\alpha_1)(P_2^{-1}A^{\alpha_2}P_2)\\ &= P_1^{-1}(P_2^{-1}A^{\alpha_2}P_2)^{\alpha_1}P_1\\ &= P_1^{-1}P_2^{-1\alpha_1}A^{\alpha_2\alpha_1}P_2^{\alpha_1}P\\ &= (P_1P_2^{\alpha_1},\alpha_1\alpha_2)(A).\end{aligned}$$

Therefore the product is semidirect.

Corollary 2. Any automorphism fixing the center of $M_n(R)$ is an inner automorphism.

Proof: The center of $M_n(R)$ consists of all scalar matrices. If $\Lambda(rI) = rI$ then $\alpha = 1_R$.

Canonical Form of an Involution

We assume for the remainder of the chapter that R/M is a field of characteristic other than 2. The assumption is utilized in two settings:

$$r = -r \text{ implies } r = 0$$

and

$$r^2 = 1 \text{ implies } r = \pm 1.$$

An element A of $GL_n(R)$ is called an involution if $A^2 = I$. With each involution A we associate two submodules of $R^{(n)}$:

$$N(A) = \{X \text{ in } R^{(n)}: A(X) = -X\}$$

and

$$P(A) = \{X \text{ in } R^{(n)}: A(X) = X\}.$$

Obviously $P(A) \cap N(A) = 0$. If X is in $R^{(n)}$

$$X = \tfrac{1}{2}(X - A(X)) + \tfrac{1}{2}(X + A(X))$$

and therefore $R^{(n)} = N(A) \oplus P(A)$.

Theorem 4.2. If A is an involution in $GL_n(R)$, then there is a unique integer t such that A is similar to the matrix $-I^{(t)} \oplus I^{(n-t)}$ where $I^{(q)}$ denotes a q by q identity block.

Proof: Note R is a local ring and thus any module which is projective over R is free. Since $R^{(n)} = N(A) \oplus P(A)$ it follows that both N(A) and P(A) are projective and hence free. Letting $\{x_1,\cdots,x_t\}$ be a basis of N(A) and $\{x_{t+1},\cdots,x_n\}$ a basis of P(A), we define a change of basis matrix Q by $Q(e_i) = x_i$, where $\{e_1,\cdots,e_n\}$ is the standard basis for $R^{(n)}$. Then

$$Q^{-1}AQ = -I^{(t)} \oplus I^{(n-t)}.$$

The dimension t of N(A) is unique since N(A) is completely determined by A and free modules over local rings have well-defined dimension.

Definition 4.1. An involution which is similar to a matrix of the form $-I^{(t)} \oplus I^{(n-t)}$ is said to be an involution of type (t,n-t). The index t satisfies $0 \leq t \leq n$.

Let E_n denote the collection of all matrices in $GL_n(R)$ having 1 or -1 in any combination on the main diagonal and zero's elsewhere. An element of E_n is an involution and any two elements of E_n commute. Note that if A is in E_n, A is of type (t,n-t) if and only if A has exactly t -1's on the main diagonal.

Theorem 4.3. If $\{A_i\}_{i=1}^{r}$ is a collection of pairwise commutative involutions, then there is a P in $GL_n(R)$ for which $P^{-1}A_iP$ is in E_n for all $i = 1,2,\cdots,r$.

Proof: The theorem is obvious in $GL_2(R)$ and, proceeding by induction, assume the theorem holds in $GL_k(R)$ for $2 \leq k < n$. We may assume

that A_1 has type (t,n-t) for some t satisfying $1 \leq t \leq n-1$ for otherwise the result would be trivial. By Theorem 4.2 there is a Q in $GL_n(R)$ such that

$$Q^{-1}A_1Q = -I^{(t)} \oplus I^{(n-t)}.$$

For each i, $1 \leq i \leq r$, the matrix $Q^{-1}A_iQ$ commutes with $Q^{-1}A_1Q$ and therefore has the form

$$Q^{-1}A_iQ = B_i^{(t)} \oplus B_i^{(n-t)}$$

where $B_i^{(q)}$ denotes a q by q matrix block. The sets $\{B_i^{(t)}\}_{i=1}^r$ and $\{B_i^{(n-t)}\}_{i=1}^r$ are collections of pairwise commutative involutions from $GL_t(R)$ and $GL_{n-t}(R)$, respectively. By induction there exist $Q^{(t)}$ in $GL_t(R)$ and $Q^{(n-t)}$ in $GL_{n-t}(R)$ such that

$$Q^{(t)-1}B_i^{(t)}Q^{(t)} \quad \text{is in} \quad E_t$$

and

$$Q^{(n-t)^{-1}}B_i^{(n-t)}Q^{(n-t)} \quad \text{is in} \quad E_{n-t}.$$

Therefore, for all i, $1 \leq i \leq r$,

$$(Q^{(t)^{-1}} \oplus Q^{(n-t)^{-1}})Q^{-1}A_iQ(Q^{(t)} \oplus Q^{(n-t)})$$

is in E_n and the proof is complete.

Corollary. There are at most $\binom{n}{t}$ elements in any collection of pairwise commutative involutions of type (t,n-t).

Let J_{ij} denote the diagonal matrix with - 1 in both the (i,i) and (j,j) positions and 1 is elsewhere. The matrix J_{ij} is an involution of type (2,n-2); the collection $\{J_{ij}: 1 \leq i, j \leq n, i \neq j\}$ forms a set of $\binom{n}{2}$ pairwise commutative elements. Note also that $J_{ij}J_{jk} = J_{ik}$.

Preservation of J_{ij}

We now begin the preparations necessary to determine the group $\mathrm{Aut}(GL_n(R))$. Assume henceforth that $n \geq 3$ and that the characteristic of R/M is other than 2. Initially it is shown that any automorphism of $GL_n(R)$ acts as an inner automorphism on the set $\{J_{ij}\}_{i,j=1}^{n}$.

Theorem 4.4. If Λ is an automorphism of $GL_n(R)$, then there is a P in $GL_n(R)$ such that

$$P\Lambda J_{ij}P^{-1} = J_{ij} \quad \text{for all } i \neq j.$$

Proof: The collection

$$\{\Lambda J_{ij} : 1 \leq i, j \leq n, i \neq j\}$$

is made up of pairwise commutative involutions. Furthermore since J_{ij} is similar to J_{rs} for any pairs (i,j) and (r,s), it follows that ΛJ_{ij} is similar to ΛJ_{rs} for any pairs (i,j) and (r,s). Therefore the involutions ΛJ_{ij}, $1 \leq i, j \leq n$ are all of the same type, say $(t,n-t)$. It will be shown that $t = 2$ by eliminating all other possibilities.

There are $\binom{n}{2}$ elements in the collection $\{\Lambda J_{ij}\}$; hence, $t \neq 0$ and $t \neq n$ by the corollary to Theorem 4.3. Also when $n > 3$,

$$\binom{n}{1} = \binom{n}{n-1} < \binom{n}{2}$$

and therefore $t \neq 1$ and $t \neq n - 1$. If $n = 3$ and $t = 1$, then ΛJ_{12} and ΛJ_{23} are of type $(1,2)$. Thus the equation,

$$\Lambda(J_{12} \cdot J_{23}) = \Lambda J_{13},$$

implies ΛJ_{13} is of type $(2,1)$ which is a contradiction. We have thus shown that if $n = 3$, then $t = 2$ and if $n > 3$, then $1 < t < n - 1$.

Now assume that $2 < t < n-2$. Since $\binom{n}{2} < \binom{n}{t}$ there is at least one involution, B, of type (t,n-t) which is not in the set $\{\Lambda J_{ij}\}$ and which commutes with each ΛJ_{ij}. But then $\Lambda^{-1}B$ commutes with all J_{ij} and has type (2,n-2) which is impossible since $\{J_{ij}\}_{i,j=1}^{n}$ is a maximal collection of pairwise commutative involutions of type (2,n-2). Thus the only remaining possibilities are $t = 2$ or $t = n-2$. If $n = 4$, then either possibility yields $t = 2$; therefore we may assume that $n > 4$. Assume further that, with the same similarity factor, ΛJ_{12} is similar to $-I^{(n-2)} \oplus I^{(2)}$ and ΛJ_{23} is similar to either $1 \oplus -I^{(n-2)} \oplus 1$ or $I^{(2)} \oplus -I^{(n-2)}$. Thus $\Lambda(J_{12} \cdot J_{23}) = \Lambda J_{23}$ is of type (2,n-2) or of type (4,n-4). The first possibility is always a contradiction and the second is also unless $n = 6$. Assume that $n = 6$, that $Q^{-1}\Lambda J_{12}Q = -I^{(4)} \oplus I^{(2)}$, and that $Q^{-1}\Lambda J_{ij}Q$ is an element of E_6 for all $i \neq j$, $1 \leq i, j \leq 6$. Among the eight elements,

$$\{\Lambda J_{1k}, \Lambda J_{2k} : k = 3,4,5,6\}$$

at least one, denoted by ΛJ_{ik}, has a 1 in the (5,5) or the (6,6) position. A permutation matrix that fixes $Q^{-1}\Lambda J_{12}Q$ will bring ΛJ_{ik} into the form $1 + -I^{(n-2)} \oplus 1$. Then

$$\Lambda(J_{12} \cdot J_{ik}) = \begin{cases} \Lambda J_{1k}, & \text{if } i = 2 \\ \Lambda J_{2k}, & \text{if } i = 1 \end{cases}$$

will be of type (2,n-2) which is a contradiction. Therefore the involutions $\{\Lambda J_{ij}\}_{i,j=1}^{n}$ are of type (2,n-2).

By Theorem 4.3 there is a P in $GL_n(R)$ such that

$$P\Lambda J_{ij}P^{-1} = J_{k_i,k_j}, \quad i \neq j, \quad i,j = 1,\cdots,n,$$

where $k_1,k_2,\cdots,k_n$ is a permutation of $1,2,\cdots,n$. Let Q be the permutation matrix satisfying,

$$Q\begin{pmatrix} x_{k_1} \\ \cdot \\ \cdot \\ \cdot \\ x_{k_n} \end{pmatrix} = \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix}.$$

It then follows that

$$Q(P\Lambda J_{ij}P^{-1})Q^{-1} = J_{ij}$$

which completes the proof.

The Image of $B_{ij}(1)$

Our purpose in this section is to determine the action of an automorphism Λ of $GL_n(R)$ on elementary transvections of the form $B_{ij}(1)$. It will be shown that there is a P in $GL_n(R)$ such that either

$$P\Lambda B_{ij}(1)P^{-1} = B_{ij}(1) \quad \text{for all } i,j$$

or

$$P\Lambda B_{ij}(1)P^{-1} = B_{ji}(-1) \quad \text{for all } i,j.$$

Many of the proofs are based on the idea that the form of a matrix A is known whenever a suitable set of matrices that commute with A is known. For example if A commutes with all $B_{ij}(1)$ then A must be a scalar matrix; if A commutes with J_{12} then A has the form,

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \oplus B^{(n-2)}$$

where $B^{(n-2)}$ is an (n-2) by (n-2) block. The latter follows since right multiplication by J_{12} multiplies the first two columns of A by - 1 and left multiplication multiplies the first two rows by - 1. Thus if $j > 2$,

$a_{1j} = -a_{1j}$ and hence $a_{1j} = 0$. Note the use of the overall assumption that the characteristic of R/M is not 2.

Lemma 4.2. If a, b, c, d are elements of the local ring R and

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^2 = -I, \quad \begin{bmatrix} a & -b \\ c & -d \end{bmatrix}^2 = I,$$

then $a = d = 0$ and $c = -b^{-1}$.

Proof: The two matrix equations yield $a^2 + bc = -1$, $d^2 + bc = -1$, $a^2 - bc = 1$ and $d^2 - bc = 1$. Therefore we have $2bc = -2$ and hence that b and c are units with $c = -b^{-1}$. The matrix equations also imply that

$$0 = b(a+d) = c(a+d) = b(d-a) = c(a-d).$$

Thus $a = d = -d$ implies that $0 = a = d$.

Let $S_{i,i+1}$ denote the permutation matrix of the form

$$S_{i,i+1} = I^{(i-1)} \oplus \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \oplus I^{(n-i-1)}, \quad i = 1,2,\cdots,n-1.$$

Observe that left multiplication by $S_{i,i+1}$ replaces the i-th row by the (i+1)-th row and the (i+1)-th row by -1 times the i-th row. Right multiplication by $S_{i,i+1}$ yields the corresponding column operation. The matrix $S_{i,i+1}$ is an element of $SL_n(R)$ and has inverse,

$$S_{i,i+1}^{-1} = I^{(i-1)} \oplus \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \oplus I^{(n-i-1)}, \quad i = 1,\cdots,n.$$

We shall use the fact that if the diagonal $[a_1,\cdots,a_n]$ commutes with $S_{i,i+1}$, then $a_i = a_{i+1}$. For the action of automorphisms of $GL_n(R)$ on $S_{i,i+1}$ we prove the following theorem.

Theorem 4.5. If Λ is an automorphism of $GL_n(R)$, then there is an e in R such that $e = \pm 1$ and a P in $GL_n(R)$ with

$$P\Lambda S_{i,i+1}P^{-1} = eI^{(i-1)} \oplus \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \oplus eI^{(n-i-1)}, \quad i = 1,2,\cdots,n-1.$$

Proof: By the previous theorem there is a Q such that

$$Q\Lambda J_{i,i+1}Q^{-1} = J_{i,i+1}, \quad i = 1,2,\cdots,n-1.$$

Define $\bar{\Lambda}$ by $\bar{\Lambda}(A) = Q\Lambda(A)Q^{-1}$. We determine the action of $\bar{\Lambda}$ on $S_{i,i+1}$. It is obvious that $J_{i,i+1}$ commutes with S_{12} for $i = 1$ or $i \geq 3$ and hence that $\bar{\Lambda}S_{12}$ commutes with $\bar{\Lambda}J_{i,i+1} = J_{i,i+1}$ for $i = 1$ or $i \geq 3$. Therefore if $n = 3$ or $n \geq 5$,

$$\bar{\Lambda}S_{12} = \begin{bmatrix} a^{(1)} & b^{(1)} \\ c^{(1)} & d^{(1)} \end{bmatrix} \oplus [a_3^{(1)},\cdots,a_n^{(1)}]$$

or if $n = 4$,

$$\bar{\Lambda}S_{12} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \oplus \begin{bmatrix} w & x \\ y & z \end{bmatrix}.$$

By considering the equations

$$(1) \qquad (\bar{\Lambda}S_{12})^2 = \bar{\Lambda}J_{12} = J_{12} \quad \text{and} \quad (\bar{\Lambda}S_{12}\bar{\Lambda}J_{23})^2 = I$$

it may be concluded that when $n = 4$

$$w^2 = z^2 = 1 \quad \text{and} \quad x = y = 0.$$

That is, for all values of n,

$$\bar{\Lambda}S_{12} = \begin{bmatrix} a^{(1)} & b^{(1)} \\ c^{(1)} & d^{(1)} \end{bmatrix} \oplus [a_3^{(1)},\cdots,a_n^{(1)}].$$

Using the equations in (1) again,

$$\begin{bmatrix} a^{(1)} & b^{(1)} \\ c^{(1)} & d^{(1)} \end{bmatrix}^2 = -I, \quad \begin{bmatrix} a^{(1)} & -b^{(1)} \\ c^{(1)} & -d^{(1)} \end{bmatrix}^2 = I$$

and $(a_i^{(1)})^2 = 1$. By Lemma 4.2 it follows that $c^{(1)} = -b^{(1)^{-1}}$ and that $a^{(1)} = d^{(1)} = 0$. Thus $\bar{\Lambda}S_{12}$ has the form,

$$\bar{\Lambda}S_{12} = \begin{bmatrix} 0 & b^{(1)} \\ -b^{(1)^{-1}} & 0 \end{bmatrix} \oplus [a_3^{(1)},\cdots,a_n^{(1)}]$$

where $a_i^{(1)} = \pm 1$. In exactly the same manner,

$$\bar{\Lambda}S_{i,i+1} = [a_1^{(i)},\cdots,a_{i-1}^{(i)}] \oplus \begin{bmatrix} 0 & b^{(i)} \\ -b^{(i)^{-1}} & 0 \end{bmatrix} \oplus [a_{i+1}^{(i)},\cdots,a_n^{(i)}]$$

where $a_j^{(i)} = \pm 1$.

Now a similarity factor is constructed that does not alter the $J_{i,i+1}$ but transforms the $S_{i,i+1}$ to the form asserted in the theorem. Let

$$P_1 = \left[\prod_{t=1}^{n-1} b^{(t)^{-1}}, \prod_{t=2}^{n-1} b^{(t)^{-1}},\cdots,b^{(n-1)^{-1}},1\right]$$

and note that $P_1 J_{i,i+1} P_1^{-1} = J_{i,i+1}$ since P_1 is a diagonal matrix. For $1 \leq i \leq n-1$,

$$P_1\bar{\Lambda}S_{i,i+1}P_1^{-1} = [a_1^{(i)},\cdots,a_{i-1}^{(i)}] + \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + [a_{i+1}^{(i)},\cdots,a_n^{(i)}].$$

To complete the proof of the theorem it must be shown that $a_j^{(i)} = a_k^{(i)}$ for all j and k and that $a_j^{(i)} = a_j^{(i+1)}$ for all j and k and that $a_j^{(i)} = a_j^{(i+1)}$ for all i. Observe that $S_{i,i+1}$ commutes with elements of the set

$$\{S_{j,j+1} : 1 \leq j \leq i-2,\ i+2 \leq j \leq n-1\}$$

and hence that $P_1\bar{\Lambda}S_{i,i+1}P_1^{-1}$ commutes with elements of

$$\{P_1\bar{\Lambda}S_{i,i+1}P_1^{-1} : 1 \leq j \leq i-2,\ i+2 \leq j \leq n-1\}.$$

It follows that $a_1^{(i)} = a_2^{(i)} = \cdots = a_{i-1}^{(i)}$ and that $a_{i+1}^{(i)} = a_{i+2}^{(i)} = \cdots = a_n^{(i)}$. Thus

$$P_1\bar{\Lambda}S_{i,i+1}P_1^{-1} = a^{(i)}I \oplus \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \oplus a^{(i)'}I.$$

The fact that $(S_{i-1,i}\ S_{i,i+1})^3 = I$ implies

$$P_1\bar{\Lambda}S_{i-1,i}\bar{\Lambda}S_{i,i+1}P_1^{-1}{}^3 = I.$$

The last equation yields,

$$a^{(i-1)'}a^{(i)} = 1,\ (a^{(i)}\cdot a^{(i-1)})^3 = 1, \text{ and } (a^{(i)'}a^{(i-1)'})^3 = 1.$$

Since all the a's are either 1 or - 1 it follows that

$$a^{(i-1)'} = a^{(i)},\ a^{(i)} = a^{(i-1)}, \text{ and } a^{(i)'} = a^{(i-1)'}.$$

Letting e denote the common value,

$$P_1\Lambda S_{i,i+1}P_1^{-1} = eI^{(i-1)} \oplus \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \oplus eI^{(n-i-1)}$$

for $i = 1,2,\cdots,n-1$ where $e = \pm 1$. Finally, since $\bar{\Lambda} = Q\Lambda(\ )Q^{-1}$,

$$(P_1Q)\Lambda S_{i,i+1}(P_1Q)^{-1} = eI^{(i-1)} \oplus \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \oplus eI^{(n-i-1)}.$$

Taking $P = P_1Q$, the proof is complete.

Theorem 4.6. If Λ is an automorphism of $GL_n(R)$, then there is a P in $GL_n(R)$ such that either

$$P\Lambda B_{ij}(1)P^{-1} = B_{ij}(1) \quad \text{for all } (i,j),$$

or

$$P\Lambda B_{ij}(1)P^{-1} = B_{ji}(-1) \quad \text{for all } (i,j).$$

Proof: By the previous theorems there is a P such that

$$P\Lambda S_{i,i+1}P^{-1} = eI^{(i-1)} \oplus \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \oplus eI^{(n-i-1)}$$

and

$$P\Lambda J_{i,i+1}P^{-1} = J_{i,i+1}.$$

Let $\bar{\Lambda}A = P\Lambda AP^{-1}$ for all A in $GL_n(R)$. It will be shown that

$$\bar{\Lambda}B_{ij}(1) = B_{ij}(1)$$

or that

$$\bar{\Lambda}B_{ij}(1) = B_{ji}(-1).$$

The matrix $\bar{\Lambda}B_{12}(1)$ must have the form,

$$\bar{\Lambda}B_{12}(1) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \oplus [a_3,\cdots,a_n]$$

when $n \neq 4$ because $B_{12}(1)$ commutes with J_{12} and $J_{i,i+1}$ for $i \geq 3$. Thus $\bar{\Lambda}B_{12}$ commutes with $\bar{\Lambda}J_{i,i+1} = J_{i,i+1}$ for all i, $i \neq 2$ and therefore $\bar{\Lambda}B_{12}$ has the above form. By the argument with $n = 4$ it may be concluded that

$$\bar{\Lambda}B_{12}(1) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \oplus \begin{bmatrix} w & x \\ y & z \end{bmatrix}.$$

We begin the case $n \neq 4$ by showing that $a_3 = a_4 = \cdots = a_n$. The element $B_{12}(1)$ commutes with $S_{i,i+1}$ for $i \geq 3$ hence $\bar{\Lambda}B_{12}(1)$ commutes with $\bar{\Lambda}S_{i,i+1}$ for $i \geq 3$. In particular $\bar{\Lambda}B_{12}(1)\cdot\bar{\Lambda}S_{34} = \bar{\Lambda}S_{34}\bar{\Lambda}B_{12}(1)$ implies that

$$e\begin{bmatrix} a & b \\ c & d \end{bmatrix} \oplus \begin{bmatrix} 0 & a_3 \\ -a_4 & 0 \end{bmatrix} \oplus e[a_5,\cdots,a_n] = e\begin{bmatrix} a & b \\ c & d \end{bmatrix} \oplus \begin{bmatrix} 0 & a_4 \\ -a_3 & 0 \end{bmatrix} \oplus e[a_5,\cdots,a_n].$$

Thus $a_3 = a_4$. In a similar fashion it is concluded that $a_3 = a_4 = \cdots = a_n$ and hence that

$$\bar{\Lambda}B_{12}(1) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \oplus a'I^{(n-2)}.$$

We now claim $a' = e$. Consider the equations,

$$\bar{\Lambda}(B_{12}(1)J_{23})^2 = I, \tag{2}$$

$$\bar{\Lambda}(S_{12}B_{12}(1))^3 = I, \tag{3}$$

$$\bar{\Lambda}(S_{23}^{-1}B_{12}(1)S_{23}) = \bar{\Lambda}B_{13}(1). \tag{4}$$

From (2) it follows that $(a')^2 = 1$. From (3) it follows that $(ea')^3 = 1$. Thus $ea' = 1$ and since $e = \pm 1$ and $a' = \pm 1$ it follows that $e = a'$.

The above three equations also determine the form of the matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$. From (2) we have that

$$\begin{bmatrix} a & -b \\ c & -d \end{bmatrix}^2 = I$$

and thus the equations $a^2 - bc = 1$, $d^2 - bc = 1$, $0 = b(d-a) = c(a-d)$. The matrix $B_{12}(1)$ commutes with $B_{13}(1)$ and thus from (4) we have

$$\bar{\Lambda}B_{12}(1)\bar{\Lambda}(S_{23}{}^{-1}B_{12}(1)S_{23}) = \bar{\Lambda}(S_{23}B_{12}(1)S_{23})\bar{\Lambda}B_{12}(1).$$

Comparing the upper-left 3 by 3 blocks for this equation yields,

$$\begin{bmatrix} a^2 & be & aeb \\ ac & de & bec \\ c & 0 & ed \end{bmatrix} = \begin{bmatrix} a^2 & ab & b \\ ec & ed & 0 \\ aec & bec & ed \end{bmatrix}.$$

Thus $bec = 0$ and $b(a-e) = c(a-e) = 0$. Since e is a unit, it follows that $bc = 0$ and from (2) that $a^2 = d^2 = 1$.

We now claim that either b or c is a unit. Equation (3) implies that

$$\begin{bmatrix} c & d \\ -a & -b \end{bmatrix}^3 = I,$$

that is,

$$\begin{bmatrix} c(c^2 - 2ad) + bad & d(c^2 + b^2) - a \\ - a(c^2 + b^2) + d & - b(b^2 - 2ad) - cad \end{bmatrix} = I.$$

Thus $c(c^2 - 2ad) + bad = 1$ and therefore it is impossible for c and b both to be non-units. Note that since $bc = 0$, if b is a unit, then $c = 0$ and if c is a unit, then $b = 0$. Also from (4) we have that $b(a - e) = c(a - e) = 0$ which now implies that $a = e$; from (2) we have $b(a - d) = c(a - d) = 0$ and hence $a = d$. Furthermore when $c = 0$, from (3), we have that $bad = 1$ and since $a = d = \pm 1$ it may be concluded that $b = 1$. Also from (3) if $b = 0$, then $- cad = 1$ and hence $c = - 1$. We have now shown that

$$\overline{\Lambda}B_{12}(1) = eI^{(n)} + E_{12}$$

or

$$\overline{\Lambda}B_{12}(1) = eI^{(n)} - E_{21}$$

using equation (4) we also have

$$\overline{\Lambda}B_{13}(1) = eI^{(n)} + eE_{13}$$

or if $\Lambda B_{12}(1)$ has the second form,

$$\overline{\Lambda}B_{13}(1) = eI^{(n)} - eE_{31}.$$

We now show that $e = 1$. Observe that $S_{12}^{-1}B_{13}(1)S_{12} = B_{23}(1)$ and hence

$$\overline{\Lambda}B_{23}(1) = eI^{(n)} + E_{23}$$

or

$$\overline{\Lambda}B_{23}(1) = eI^{(n)} - E_{32}.$$

The commutator relation,

$$[\bar{\Lambda}B_{12}(1),\bar{\Lambda}B_{23}(1)] = \bar{\Lambda}B_{13}(1)$$

yields that $e^4 = e$ and hence that $e = 1$.

To determine the images of the other elementary transvections assume that $\bar{\Lambda}B_{12}(1) = B_{12}(1)$ and note that $\bar{\Lambda}B_{21}(1) = \bar{\Lambda}(S_{12}^{-1}B_{12}(-1)S_{12}) = B_{21}(1)$. Assume inductively that

$$\bar{\Lambda}B_{1i}(1) = B_{1i}(1) \text{ and } \bar{\Lambda}B_{i1}(1) = B_{i1}(1).$$

Then since $S_{i,i+1}^{-1} B_{1i}(1)S_{i,i+1} = B_{1,i+1}(1)$ and $S_{i,i+1}^{-1} B_{i1}(1)S_{i,i+1} = B_{i+1,1}(1)$ it follows that

$$\bar{\Lambda}B_{1,i+1}(1) = B_{1,i+1}(1) \text{ and } \bar{\Lambda}B_{i+1,1}(1) = B_{i+1,1}(1).$$

Thus, if $\bar{\Lambda}B_{12}(1) = B_{12}(1)$, then $\bar{\Lambda}B_{1i}(1) = B_{1i}(1)$ and $\bar{\Lambda}B_{i1}(1) = B_{i1}(1)$ for $i = 2,\cdots,n$. Similarly it can be shown that if $\bar{\Lambda}B_{12}(1) = B_{21}(-1)$, then $\bar{\Lambda}B_{1i}(1) = B_{i1}(-1)$ and $\bar{\Lambda}B_{i1}(1) = B_{1i}(-1)$. Finally by using the Basic Identity

$$[B_{ij}(1),B_{jk}(1)] = B_{ik}(1), \quad i \neq j \neq k,$$

it follows that

$$\bar{\Lambda}B_{ij}(1) = B_{ij}(1) \quad \text{for all} \quad i,j,$$

or

$$\bar{\Lambda}B_{ij}(1) = B_{ji}(-1) \quad \text{for all} \quad i,j.$$

The proof is now complete for the case $n \neq 4$.

Now let $n = 4$. Since $\bar{\Lambda}B_{12}(1)$ commutes with $\bar{\Lambda}J_{12}$,

$$\bar{\Lambda}B_{12}(1) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \oplus \begin{bmatrix} w & x \\ y & z \end{bmatrix}.$$

We show that w = z and x = - y. The matrix $B_{12}(1)$ commutes with S_{34} and hence,

$$\begin{bmatrix} y & z \\ -w & -x \end{bmatrix} = \begin{bmatrix} -x & w \\ -z & y \end{bmatrix}.$$

Thus x = - y and z = w. Using equation (2),

$$\begin{bmatrix} a^2 - bc & -b(a-d) & 0 & 0 \\ c(a-d) & d^2 - bc & 0 & 0 \\ 0 & 0 & x^2 + w^2 & 0 \\ 0 & 0 & 0 & x^2 + w^2 \end{bmatrix} = I$$

which implies that $a^2 - bc = 1$, $d^2 - bc = 1$, $x^2 + w^2 = 1$, and $0 = c(a-d) = b(a-d)$. We claim that either b or c is a unit, that a = d = w, and that x = 0. That b or c is a unit follows, just as in the first case, by considering the (1,1) positions in equation (3). We have

$$c(c^2 - 2ad) + bad = 1,$$

which implies that it is impossible for both b and c to be non-units. It follows from $0 = c(a-d) = b(a-d)$ that a = d. To complete the argument use equation (4) and the fact that $\bar{\Lambda}B_{12}(1)$ commutes with $\bar{\Lambda}B_{13}(1)$ to conclude,

$$\begin{bmatrix} a^2 & bw & eba & -ebx \\ ac & aw & ebc & -eax \\ ecw & ex^2 & wa & wx \\ -ecx & ewx & -xa & w^2 \end{bmatrix} = \begin{bmatrix} a^2 & ab & ebw & ebx \\ wc & wa & ex^2 & -exw \\ eca & ecb & aw & ax \\ exc & exa & -ax & a^2 \end{bmatrix}.$$

From this matrix equation it follows that b(a - w) = 0, c(a - w) = 0, 2ebx = 0 and 2ecx = 0. Using the fact that either b or c is a unit it

follows that $a = w$ and $x = 0$. Finally, since $w^2 = a^2 = 1$, recall that $a^2 - bc = 1$ and hence $bc = 0$. Thus, if b is a unit then $c = 0$ and when c is a unit then $b = 0$. To summarize, we have shown $\bar{\Lambda}B_{12}(1)$ has one of the following forms:

$$\bar{\Lambda}B_{12}(1) = \begin{bmatrix} a & b \\ 0 & a \end{bmatrix} \oplus [a,a]$$

or

$$\bar{\Lambda}B_{12}(1) = \begin{bmatrix} a & 0 \\ c & a \end{bmatrix} \oplus [a,a].$$

The proof may be completed just as in the first case.

It has been shown that either

$$\bar{\Lambda}B_{ij}(1) = B_{ij}(1) \quad \text{for all} \quad i,j,$$

or

$$\bar{\Lambda}B_{ij}(1) = B_{ji}(-1) \quad \text{for all} \quad i,j.$$

Thus since $\bar{\Lambda}(\ ) = P\Lambda(\ )P^{-1}$,

$$P\Lambda B_{ij}(1)P^{-1} = B_{ij}(1) \quad \text{for all } i,j,$$

or

$$P\Lambda B_{ij}(1)P^{-1} = B_{ji}(-1) \quad \text{for all } i,j,$$

and the proof of the theorem is complete.

Note that $B_{ji}(-1)$ is the transpose of the inverse of $B_{ij}(1)$. Let A^t denote the transpose of the matrix A and recall the facts

$$(AB)^t = B^tA^t \quad \text{and} \quad ((AB)^{-1})^t = (A^{-1})^t(B^{-1})^t.$$

The composite operation of inverse-transpose will be denoted by,

$$(A^{-1})^t = A^*.$$

Theorem 4.7. If Λ is an automorphism $GL_n(R)$ then there is a ring

automorphism α: $R \to R$ and a P in $GL_n(R)$ such that

$$P\Lambda AP^{-1} = A^{\alpha} \quad \text{for all } A \text{ in } SL_n(R)$$

or

$$P\Lambda AP^{-1} = (A^{\alpha})^* \quad \text{for all } A \text{ in } SL_n(R).$$

Proof: $SL_n(R)$ is generated by elementary transvections so that it suffices to find an automorphism α and a matrix P such that either

$$P\Lambda B_{ij}(\lambda)P^{-1} = B_{ij}(\lambda^{\alpha}) \quad \text{for all } B_{ij}(\lambda)$$

or

$$P\Lambda B_{ij}(\lambda)P^{-1} = B_{ji}(-\lambda^{\alpha}) \quad \text{for all } B_{ij}(\lambda).$$

Observe that if $A = \Pi B_{ij}(\lambda)$ then

$$\Lambda A = \Pi\Lambda B_{ij}(\lambda) = \Pi P B_{ij}(\lambda^{\alpha})P^{-1}$$

$$= P\Pi B_{ij}(\lambda^{\alpha})P^{-1} = PA^{\alpha}P^{-1}.$$

Assume that $P\Lambda B_{ij}(1)P^{-1} = B_{ij}(1)$ and that $P\Lambda S_{i,i+1}P^{-1} = S_{i,i+1}$. Let λ be an element of R and, using the fact that $B_{12}(\lambda)$ commutes with $B_{12}(1)$ and $B_{ij}(1)$, $3 \le i,j \le n$, conclude that

$$P\Lambda B_{12}(\lambda)P^{-1} = \begin{bmatrix} a & b \\ 0 & a \end{bmatrix} \oplus a_1 I^{(n-2)}.$$

Now, using the fact, $S_{23}^{-1} B_{12}(\lambda) S_{23} = B_{13}(\lambda)$,

$$P\Lambda B_{13}(\lambda)P^{-1} = \begin{bmatrix} a & 0 & b \\ 0 & a_1 & 0 \\ 0 & 0 & a \end{bmatrix} \oplus a_1 I^{(n-3)}. \tag{5}$$

But from the commutator relation,

$$[B_{12}(\lambda), B_{23}(1)] = B_{13}(\lambda)$$

it follows that

$$P[\Lambda B_{12}(\lambda), \Lambda B_{23}(1)]P^{-1} = P\Lambda B_{13}(\lambda)P^{-1}$$

and therefore that

$$P\Lambda B_{13}(\lambda)P^{-1} = \begin{bmatrix} 1 & 0 & ba_1a^{-2} \\ 0 & 1 & 1-a_1a^{-1} \\ 0 & 0 & a^{-2} \end{bmatrix} \oplus I^{(n-3)}. \tag{6}$$

Comparing (5) and (6) yields $a = 1$ and $a_1 = 1$.

Now define $\alpha: R \to R$ by $\alpha(\lambda) = b$ if and only if

$$P\Lambda B_{12}(\lambda)P^{-1} = I^{(n)} + bE_{12}.$$

As in the proof of Theorem 4.6 by using the fundamental commutator relations we can generate the other elementary transvections and show that

$$P\Lambda B_{ij}(\lambda)P^{-1} = B_{ij}(\lambda^{\alpha}), \quad \text{for all } i,j.$$

Since $\Lambda B_{12}(\lambda_1)\cdot\Lambda B_{12}(\lambda_2) = \Lambda B_{12}(\lambda_1 + \lambda_2)$, it is obvious that the map α is additive. If $\lambda^{\alpha} = 0$, then $\Lambda B_{12}(\lambda) = I$ and thus $\lambda = 0$. Hence α is injective. To show α is multiplicative consider,

$$\begin{aligned} B_{13}[(\lambda_1\lambda_2)^{\alpha}] &= P\Lambda B_{13}(\lambda_1\lambda_2)P^{-1} \\ &= P[\Lambda B_{12}(\lambda_1), \Lambda B_{23}(\lambda_2)]P^{-1} \\ &= B_{13}(\lambda_1^{\alpha}\,\lambda_2^{\alpha}). \end{aligned}$$

It has now been shown that α is a monomorphism. To show α is surjective consider the image of $SL_n(R)$ under Λ. By Theorem 4.7, $\Lambda[SL_n(R)]$ is contained in $SL_n(R)$ and has order ideal (Chapter II) R. Furthermore since $\Lambda SL_n(R)$ is normal in $GL_n(R)$ it follows by Klingenburg's results that

$$\Lambda[SL_n(R)] = SL_n(R).$$

If r is any element of R there is some $B = \Pi B_{ij}(\lambda_{ij})$ such that

$$P\Lambda BP^{-1} = B_{12}(r).$$

That is,

$$\Pi B_{ij}(\lambda_{ij}^{\alpha}) = B_{12}(r)$$

and thus r is a finite sum of finite products of the λ_{ij}^{α}. Hence r is an α image and thus α is surjective. Thus α is an automorphism of R and the proof is complete.

In the next theorem we determine the action of an automorphism on all of $GL_n(R)$.

Theorem 4.8. Let R be a local ring with the characteristic of R/M other than 2. Suppose $n \geq 3$ and Λ is an automorphism of $GL_n(R)$. Then there is a P in $GL_n(R)$, an automorphism α of R into R, and a group morphism χ of R^* into R^* such that

$$\Lambda A = \chi(\det A)P^{-1}A^{\alpha}P \quad \text{for all A in G,}$$

or

$$\Lambda A = \chi(\det A)P^{-1}(A^{\alpha})^{*}P \quad \text{for all A in G.}$$

Proof: If A is in $GL_n(R)$, then $A = D_n(r)\cdot B$ where det A = r and B is a product of elementary transvections. Since ΛB has been determined, it is only necessary to compute $\Lambda D_n(r)$. Here we assume that

$$\Lambda B_{ij}(\lambda) = P^{-1}B_{ij}(\lambda^{\alpha})^{*}P.$$

The other case is similar.

For any (i,j)

$$D_n(r)B_{ij}(1)D_n(r^{-1})$$

is an element of $SL_n(R)$ and thus,

$$P(\Lambda D_n(r)\Lambda B_{ij}(1)\Lambda D_n(r^{-1}))P^{-1}$$

$$= [D_n(r)B_{ij}(1)D_n(r^{-1})]^{\alpha^*}$$

$$= D_n(r^{\alpha^{-1}})B_{ji}(-1)D_n(r^{\alpha}).$$

Then,

$$D_n(r^{\alpha})P\Lambda D_n(r)P^{-1}B_{ji}(-1)$$

$$= B_{ji}(-1)D_n(r^{\alpha})P\Lambda D_n(r)P^{-1}.$$

Therefore $D_n(r^{\alpha})P\Lambda D_n(r)P^{-1}$ commutes with $B_{ji}(-1)$ for all $i \neq j$ and hence is a scalar matrix. That is,

$$\Lambda D_n(r) = (\text{scalar})P^{-1}D_n(r^{\alpha^{-1}})P.$$

Define $\chi: R^* \to R^*$ by

$$\chi(r) = \text{scalar associated with } \Lambda D_n(r).$$

Since $\Lambda[D_n(r_1)\cdot D_n(r_2)] = \Lambda D_n(r_1r_2)$, it follows that $\chi(r_1)\cdot\chi(r_2) = \chi(r_1r_2)$ and hence that χ is a group morphism. For any A in $GL_n(R)$,

$$\Lambda A = \Lambda(D_n(\det A)\Pi B_{ij}(\lambda_{ij}))$$

$$= \chi(\det A)P^{-1}D_n(r^{\alpha})^*P\cdot P^{-1}\Pi B_{ij}(\lambda_{ij}^{\alpha})^*P$$

$$= \chi(\det A)P^{-1}(A^{\alpha})^*P.$$

This completes the proof.

Characteristic Subgroups

We are now in a position to answer the questions posed in Chapter II

regarding characteristic subgroups. As stated previously the characteristic of $R/M \neq 2$ and $n \geq 3$.

Theorem 4.9. Let R be a local ring with maximal ideal M. Then for $i \geq 1$ the subgroups $GC_n(R,M^i)$ are characteristic in $GL_n(R)$.

Proof: Let Λ be an element of $\mathrm{Aut}(GL_n(R))$. Note that for any automorphism $\alpha: R \to R$ that $\alpha M = M$ and hence $\alpha(M^i) = M^i$. If $A = rI + N_i$ is an element of $GC_n(R,M^i)$, with N_i an n by n matrix with elements in M^i, then there are P in $GL_n(R)$, $\chi: R^* \to R^*$ and α in $\mathrm{Aut}(R)$ such that

$$\Lambda(rI + N_i) = \chi(\det A)P^{-1}((rI)^\alpha + N_i^\alpha)P.$$

By the above observation on ring automorphisms we have N_i^α has all entries from M^i. Since $GC_n(R,M^i)$ is normal it is clear that the right hand side of the above equation is in $GC_n(R,M^i)$ and hence $\Lambda(GC_n(R,M^i)) \subseteq GC_n(R,M^i)$. Thus $GC_n(R,M^i)$ is characteristic. For automorphisms of the second type observe that $GC_n(R,M^i)$ is invariant under the operation of inverse-transpose.

Corollary 1. For $i \geq 1$, $SC_n(R,M^i)$ is a characteristic subgroup of $GL_n(R)$.

Proof: Note $SC_n(R,M^i) = [GL_n(R),GC_n(R,M^i)]$.

BIBLIOGRAPHY

1. Artin, Emil Geometric Algebra. New York: Interscience Publishers, 1962.

2. Atiyah, Michael F., and MacDonald, Ian G. Introduction to Commutative Algebra. Reading, Mass.: Addison-Wesley, 1969.

3. Barbee, Willard Louise. "On the Representations of Groups of Order p^n by Matrices Over the Integers Taken Modulo a Power of p." Unpublished Ph.D. dissertation, The University of Michigan, 1968.

4. Bass, Hyman. Algebraic K-Theory. New York: W. A. Benjamin, 1968.

5. Brenner, Joel. "The Linear Homogeneous Group." Annals of Mathematics, 39 (1938), 137-153.

6. Brenner, Joel. "The Linear Homogeneous Group: II." Annals of Mathematics, 45 (1944), 100-109.

7. Brenner, Joel. "The Linear Homogeneous Group: III." Annals of Mathematics, 71 (1960), 210-223.

8. Chien, Yen Shih. "Linear Groups Over a Ring." Chinese Mathematics, 7 (1965), 163-179.

9. Cohn, P. M. On the Structure of the GL_2 of a Ring. Institut des Hautes Études Scientifiques Publications Mathématiques 30. Paris, 1966.

10. Dickson, Leonard Eugene. Linear Groups. New York: Dover Publications, 1958.

11. Dieudonné, Jean. On the Automorphisms of the Classical Groups. Memoirs 2. Providence: American Mathematical Society, 1951.

12. Dieudonné, Jean. Sur les groupes classiques. Pairs: Hermann, 1948.

13. Ganske, Gary L. "Finite Local Rings." Unpublished Ph.D. dissertation, University of Oklahoma, 1971.

14. Gorenstein, Daniel. Finite Groups. New York: Harper and Row, 1968.

15. Hall, Marshal Jr. The Theory of Groups. New York: Macmillan, 1959.

16. Hua, L. K. and Reiner, I. "Automorphisms of the Unimodular Group," Transactions of the American Mathematical Society, 71(1951), 331-348.

17. Kaplansky, Irving. "Projective Modules", Annals of Mathematics. 68(1958), 372-377.

18. Klingenberg, Wilhelm. "Linear Groups Over Local Rings," American Journal of Mathematics, 83(1961), 137-153.

19. Landin, Joseph and Reiner, Irving. "Automorphisms of the General Linear Group Over a Principal Ideal Domain," Annals of Mathematics, 65(1957), 519-526.

20. Nagata, Masayoshi. Local Rings, New York: Interscience Publishers, 1962.

21. O'Meara, O. T. "The Automorphisms of Linear Groups Over Any Integral Domain," Journal für die reine und angewandte Mathematik, 223(1966), 56-100.

22. Robertson, E. F. "On Certain Subgroups of GL(R)," Journal of Algebra, 15(1970), 293-300.

23. Rotman, Joseph J. Homological Algebra. New York: Van Nostrand, 1970.

24. Rotman, Joseph J. Theory of Groups: an Introduction. Boston: Allyn and Bacon, 1965.

25. Schreier, O. and van der Waerden, B. L. "Die Automorphismen der projektiven Gruppen," Abhandlungen aus dem mathematishen Seminar der Hamburgishen Universitat. 6(1928), 303-322.

26. Solazzi, R. E. "Isomorphisms of Congruence Groups," Bulletin of the American Mathematical Society. 71(1971), 164-168.

27. Suprenenko, D. Soluble and Nilpotent Linear Groups. Translations of Mathematical Monographs Vol. 19. Providence: American Mathematical Society, 1963.

28. Weir, A. J. "Sylow p-subgroups of the General Linear Group Over Finite Fields of Characteristic p." Proceedings of the American Mathematical Society. 6(1955), 454-463.