PUBLICATIONS MATHÉMATIQUES DE L'I.H.É.S.

PAUL M. COHN

On the structure of the GL_2 of a ring

Publications mathématiques de l'I.H.É.S., tome 30 (1966), p. 5-53

<http://www.numdam.org/item?id=PMIHES_1966__30__5_0>

ON THE STRUCTURE OF THE $\mathbf{GL}_2$ OF A RING

by P. M. COHN

CONTENTS

1. Introduction.

The general linear group over a field, its subgroups and automorphisms, have been studied fairly extensively, even when the field is skew (cf. [8] and the references given there), but little is known about the general linear groups over an arbitrary ring. Now in the case of fields, the starting point is the observation that every invertible matrix is a product of elementary matrices; this suggests that in studying $\mathbf{GL}_n(\mathrm{R})$ it is best to confine attention at first to rings which share this property, and we therefore define a *generalized Euclidean ring*, or *GE-ring* for short, as an integral domain (not necessarily commutative) such that for all n,

GE_n: Every invertible $n \times n$ matrix is a product of elementary $n \times n$ matrices.

Examples of GE-rings are (i) the classical Euclidean rings (cf. [17]), (ii) rings with a weak algorithm [5], in particular free associative algebras over a commutative field, and (iii) free products of GE-rings which are also semifirs (cf. [6] and § 3 below); this includes in particular the group algebras of free groups and free products of skew fields.

The present paper is a study of $\mathbf{GL}_2(R)$, for various types of rings, and in particular GE-rings; as in the case of fields, there is a basic difference between the case of dimension 2 and dimension gieater than 2, so that it is reasonable to begin by concentrating attention on the former. For most of our results we shall find it enough to assume GE_n for $n=2$ only; in fact we can even dispense with GE_2 by limiting ourselves to the subgroup of $\mathbf{GL}_2(R)$ generated by the elementary matrices, and the results will usually be stated in this form. This more general point of view is actually forced on us when we come to give examples of rings that are not GE-rings.

Compared with Euclidean rings, GE-rings have a more intrinsic definition, in that no norm function is involved. But for a closer study of $\mathbf{GL}_2(R)$ we shall find some sort of norm function on R of great use. Most of our results will apply to *discretely normed* rings, defined in § 5 and to *discretely ordered* rings defined in § 8. It is the presence of the norm (or the ordering) which usually enables us to decide whether a given ring satisfies GE_2. As an application we shall show that the ring of polynomials in any number of indeterminates with integer coefficients, and the ring of polynomials in at least two indeterminates with coefficients in a field, are not GE-rings (§ 5). Further it is shown that the ring of algebraic integers in an imaginary quadratic number field is a GE-ring if and only if the field is Euclidean with respect to the usual norm (§ 6). In particular [1] taking the integers in $\mathbf{Q}(\sqrt{-19})$ we obtain a principal ideal domain which is not a GE-ring (clearly this says rather more than the usual assertion that this ring is not Euclidean).

The main tool of the paper is an explicit presentation of $\mathbf{GL}_2(R)$ for suitable GE-rings, called *universal GE-rings* (§ 2). As examples of such rings we have fields or more generally local rings (§ 4); moreover, any GE-ring which is either discretely normed or discretely ordeied is a universal GE-ring. This presentation brings out in a particularly clear form the extent to which $\mathbf{GL}_2(R)$ is independent of the multiplicative structure of R, and which accounts for the crucial difference between $\mathbf{GL}_n(R)$ for $n=2$ and $n>2$. Thus a generalization of a ring homomorphism, the U-homomorphism, is introduced in § 11 and it is shown that any U-homomorphism between two universal GE-rings induces a homomorphism between their $\mathbf{GL}_2$-groups. As a rather striking illustration one has the result that for any free associative algebra A over a commutative field F on at most countably many free generators, $\mathbf{GL}_2(A)\cong\mathbf{GL}_2(F[x])$. The automorphisms of $\mathbf{GL}_2(F[x])$ constructed by Reiner in [16] can also be obtained very simply from this point of view.

In the other direction the isomorphisms between the $\mathbf{GL}_2$-groups of GE-rings are studied under the assumption that the rings have a degree function defined on them such that all the elements of degree zero are units. It is shown that any such isomorphism can be built up by taking a U-isomorphism (or a U-anti-isomorphism), following it by a cen-

[1] This answers the question raised by various authors (I. Reiner, J.-P. Serre) whether every Dedekind ring, or more particularly, every principal ideal domain, is a GE-ring.

tral homothety and an inner automorphism (just as in the case of fields, cf. [8], where the notion of U-(anti-)isomorphism reduces to the ordinary (anti-)isomorphism). The proof is based on the fact that in a ring of the type named, all dihedral subgroups of order eight are conjugate, in case the characteristic of the underlying field is $\neq 2$, while in characteristic 2 all subgroups of the type of the symmetric group of degree 3 are conjugate. This last fact rests on a surprisingly delicate argument (Lemma 12.3) and it would be of interest to have a simpler proof.

The presentation introduced in § 2 can also be used to study the commutator quotient structure of $\mathbf{GL}_2(\mathrm{R})$. This is done in § 9 where an analogue of $\mathbf{SL}_2(\mathrm{R})$ is defined, denoted by $\mathbf{E}_2(\mathrm{R})$ and it is shown that for a universal GE-ring, $\mathbf{GL}_2(\mathrm{R})/\mathbf{E}_2(\mathrm{R}) \cong \mathbf{U}(\mathrm{R})^a$, where $\mathbf{U}(\mathrm{R})$ is the group of units of R and for any group G, $\mathrm{G}^a = \mathrm{G}/\mathrm{G}'$ is G made abelian. Further, if R is a discretely normed ring in which $\mathbf{U}(\mathrm{R})$ is commutative, then $\mathbf{E}_2(\mathrm{R})^a \cong \mathrm{R}/\mathrm{M}$, where M is a certain additive subgroup of R determined by the units of R (for a precise statement see Theorem 9.3). A number of other applications and generalizations of known results are given in § 10.

2. Elementary matrices over an arbitrary ring.

For a Euclidean ring R it is well known that $\mathbf{GL}_2(\mathrm{R})$, the group of invertible 2×2 matrices over R, is generated by the elementary matrices

$$\begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} \quad \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

where α, β, $a \in \mathrm{R}$ and α, β are units. Our object is to study, for any ring R, the group generated by these matrices; we shall denote this group by $\mathbf{GE}_2(\mathrm{R})$. In particular, we shall be interested in rings R for which $\mathbf{GE}_2(\mathrm{R}) = \mathbf{GL}_2(\mathrm{R})$; such rings will be called *GE_2-rings*. Clearly they include the GE-rings mentioned in § 1.

Let R be any ring (always associative, with 1); generally we shall denote arbitrary elements of R by latin letters and reserve greek letters for invertible elements of R. The group of units of R is denoted by $\mathbf{U}(\mathrm{R})$ and $\mathbf{U}_0(\mathrm{R}) = \mathbf{U}(\mathrm{R}) \cup \{0\}$. Further we set

$$(2.1) \qquad [\alpha, \beta] = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} \qquad \mathrm{D}(\alpha) = [\alpha, \alpha^{-1}] \qquad \mathrm{E}(a) = \begin{pmatrix} a & 1 \\ -1 & 0 \end{pmatrix}$$

and denote by $\mathbf{D} = \mathbf{D}_2(\mathrm{R})$ the group of 2×2 invertible diagonal matrices over R and by $\mathbf{E} = \mathbf{E}_2(\mathrm{R})$ the group generated by all $\mathrm{E}(a)$, $a \in \mathrm{R}$. Thus $\mathbf{GE}_2(\mathrm{R})$ is the group generated by $\mathbf{D}$ and $\mathbf{E}$. If we write $\mathrm{B}_{ij}(a) = \mathrm{I} + ae_{ij}$, where e_{ij} are the usual matrix units, then

$$\mathrm{B}_{12}(a) = \mathrm{E}(-a)\mathrm{E}(0)^{-1}, \qquad \mathrm{B}_{21}(a) = \mathrm{E}(0)^{-1}\mathrm{E}(a),$$

and

$$\mathrm{E}(0) = \mathrm{B}_{12}(1)\mathrm{B}_{21}(-1)\mathrm{B}_{12}(1), \qquad \mathrm{E}(a) = \mathrm{B}_{12}(-a)\mathrm{E}(0);$$

these equations show that $\mathbf{E}_2(\mathrm{R})$ is just the group generated by all $\mathrm{B}_{ij}(a)$ $(a \in \mathrm{R}, i \neq j)$. However, we shall mainly keep to the generators $\mathrm{E}(a)$, since the defining relations are expressed rather more easily in terms of them. We have the following relations between the matrices (2.1):

$$\text{(2.2)}\qquad E(x)E(0)E(y) = -E(x+y),$$

$$\text{(2.3)}\qquad E(\alpha)E(\alpha^{-1})E(\alpha) = -D(\alpha) \qquad x, y \in R,$$

$$\text{(2.4)}\qquad E(x)[\alpha, \beta] = [\beta, \alpha]E(\beta^{-1}x\alpha), \qquad \alpha, \beta \in \mathbf{U}(R).$$

Their verification is elementary and is left to the reader. These three relations are basic for much of what follows, for as we shall see, there are large classes of rings in which the relations (2.2)-(2.4), together with the relations in $\mathbf{D}_2(R)$, form a complete set of defining relations for $\mathbf{GE}_2(R)$. A ring which has this property is said to be *universal* for $\mathbf{GE}_2$; of particular interest are the GE_2-rings which are universal for $\mathbf{GE}_2$, or the *universal GE_2-rings*, as we shall call them. Thus a universal GE_2-ring is characterized by the property that $\mathbf{GL}_2(R)$ is generated by the matrices (2.1), with (2.2-4) and the relations of $\mathbf{D}_2$ as a complete set of defining relations.

The first task is to derive a number of consequences of (2.2-4): Putting $x=y=0$ in (2.2) and $\alpha = \pm 1$ in (2.3), we find

$$\text{(2.5)}\qquad E(0)^2 = -I, \qquad E(1)^3 = -I, \qquad E(-1)^3 = I.$$

Putting $y=-x$ in (2.2) and using (2.5) we get

$$\text{(2.6)}\qquad E(x)^{-1} = E(0)E(-x)E(0);$$

explicitly we have

$$E(x)^{-1} = \begin{pmatrix} 0 & -1 \\ 1 & x \end{pmatrix}, \qquad E(x) + E(x)^{-1} = xI.$$

By (2.6) we have

$$\text{(2.7)}\qquad E(x)E(y)^{-1} = E(x-y)E(0)^{-1} = -E(x-y)E(0).$$

From (2.7) we easily obtain the following generalization of (2.2):

$$\text{(2.8)}\qquad E(x)E(y)^{-1}E(z) = E(x-y+z).$$

If α is any unit in R and $x, y \in R$, then by (2.3),

$$\begin{aligned} E(x)E(\alpha^{-1})E(y) &= -E(x)E(\alpha)^{-1}D(\alpha)E(\alpha)^{-1}E(y) \\ &= -E(x-\alpha)E(0)D(\alpha)E(0)E(y-\alpha) \\ &= E(x-\alpha)D(\alpha^{-1})E(y-\alpha), \end{aligned}$$

where we have used (2.3), (2.7), (2.4) and (2.5). Replacing α by α^{-1}, we find

$$\text{(2.9)}\qquad E(x)E(\alpha)E(y) = E(x-\alpha^{-1})D(\alpha)E(y-\alpha^{-1}).$$

Finally if α, β are any units, and $x, y \in R$, then

$$\begin{aligned} E(x)E(\alpha+1)E(\beta+1)E(y) &= E(x)E(\alpha)E(0)E(1)^2E(0)E(\beta)E(y) \\ &= -E(x)E(\alpha)E(-1)E(\beta)E(y) \\ &= -E(x-\alpha^{-1})D(\alpha)E(-1-\alpha^{-1}-\beta^{-1})D(\beta)E(y-\beta^{-1}), \end{aligned}$$

where we have used (2.6), (2.5) and (2.9) in turn. Hence

$$\textbf{(2.10)}\qquad E(x)E(\alpha+1)E(\beta+1)E(y) = \\ -E(x-\alpha^{-1})D(\alpha)E(-1-\alpha^{-1}-\beta^{-1})D(\beta)E(y-\beta^{-1}).$$

We conclude this section by showing that in any ring, $\mathbf{D}_2$ normalizes $\mathbf{E}_2$ and by obtaining a certain standard form for the elements of $\mathbf{GE}_2$ which will later be shown to be unique, for suitably restricted rings.

Proposition **(2.1)**. — *In any ring* R, $\mathbf{GE}_2(R) = \mathbf{DE} = \mathbf{ED}$ *and* $\mathbf{E}$ *is normal in* $\mathbf{GE}_2(R)$.

Proof. — By (2.4), (2.5) and (2.3),

$$\begin{aligned}[\alpha, \beta]^{-1}E(x)[\alpha, \beta] &= [\alpha^{-1}\beta, \beta^{-1}\alpha]E(\beta^{-1}x\alpha)\\ &= D(\alpha^{-1}\beta)E(\beta^{-1}x\alpha)\\ &= E(\alpha^{-1}\beta)E(\beta^{-1}\alpha)E(\alpha^{-1}\beta)E(0)^2E(\beta^{-1}x\alpha).\end{aligned}$$

Thus $\mathbf{E}^{\mathbf{D}} \subseteq \mathbf{E}$ and since $\mathbf{GE}_2(R)$ is generated by $\mathbf{D}$ and $\mathbf{E}$, we have $\mathbf{E} \lhd \mathbf{GE}_2(R)$ and the result follows.

We note that in general $\mathbf{D} \cap \mathbf{E} \neq 1$, so that we do not have a semi-direct product. In fact, by (2.5) and (2.3), $D(\alpha) \in \mathbf{D} \cap \mathbf{E}$ for all $\alpha \in \mathbf{U}(R)$; later (in § 9) we shall find conditions under which the subgroup generated by the $D(\alpha)$ is exactly $\mathbf{D} \cap \mathbf{E}$.

By definition, every element of $\mathbf{GE}_2(R)$ is a product of matrices $[\alpha, \beta]$, $E(x)$ and $E(x)^{-1}$. Now by (2.6) any factor $E(x)^{-1}$ can be replaced by a product of E's, so it follows from Proposition 2.1 that every element A of $\mathbf{GE}_2(R)$ has the form

$$\textbf{(2.11)}\qquad A = [\alpha, \beta]E(a_1)\dots E(a_r).$$

If for some i in $1 < i < r$, $a_i = 0$, then we can shorten (2.11) by using (2.2); if $a_i \in \mathbf{U}(R)$ we can use (2.9) and (2.4) to shorten (2.11). Thus after a finite number of steps we reach a form (2.11) which cannot be shortened in this way and then $a_i \notin \mathbf{U}_0(R)$ for $1 < i < r$. Moreover, when $r = 2$, we may assume by (2.5) that a_1, a_2 do not both vanish. Such an expression for A is said to be a *standard form*. We note that only (2.2-4) and their consequences have been used in obtaining the standard form.

Our results may now be summed up as follows:

Theorem **(2.2)**. — *Let* R *be any ring and denote by* $\mathbf{GE}_2(R)$ *the group generated by all matrices*

$$[\alpha, \beta] = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} \quad (\alpha, \beta \in \mathbf{U}(R)) \qquad \textit{and} \qquad E(a) = \begin{pmatrix} a & 1 \\ -1 & 0 \end{pmatrix} \quad (a \in R),$$

and set $D(\alpha) = [\alpha, \alpha^{-1}]$. *These generators satisfy the relations:*

$$\textbf{(2.2)}\qquad E(x)E(0)E(y) = -E(x+y),$$

$$\textbf{(2.3)}\qquad E(\alpha)E(\alpha^{-1})E(\alpha) = -D(\alpha),$$

$$\textbf{(2.4)}\qquad E(x)[\alpha, \beta] = [\beta, \alpha]E(\beta^{-1}x\alpha).$$

These relations, together with those in the group $\mathbf{D}(\mathrm{R})$ *generated by the* $[\alpha, \beta]$ *imply*

$$E(0)^2 = -I, \qquad E(1)^3 = -I, \qquad E(-1)^3 = I, \tag{2.5}$$

$$E(x)^{-1} = E(0)E(-x)E(0), \tag{2.6}$$

$$E(x)E(y)^{-1} = E(x-y)E(0)^{-1} = -E(x-y)E(0), \tag{2.7}$$

$$E(x)E(y)^{-1}E(z) = E(x-y+z), \tag{2.8}$$

$$E(x)E(\alpha^{-1})E(y) = E(x-\alpha)D(\alpha^{-1})E(y-\alpha), \tag{2.9}$$

(2.10)

$$E(x)E(\alpha+1)E(\beta+1)E(y) = -E(x-\alpha^{-1})D(\alpha)E(-1-\alpha^{-1}-\beta^{-1})D(\beta)E(\gamma-\beta^{-1}).$$

Moreover, they imply that the group $\mathbf{E}_2(\mathrm{R})$ *generated by all* $E(a)$ *is normal in* $\mathbf{GE}_2(\mathrm{R})$ *and that every element* A *of* $\mathbf{GE}_2(\mathrm{R})$ *can be expressed in standard form*

$$A = [\alpha, \beta]E(a_1) \dots E(a_r), \tag{2.11}$$

where $\alpha, \beta \in \mathbf{U}(\mathrm{R})$, $a_i \in \mathrm{R}$ *and such that* $a_i \notin \mathbf{U}_0(\mathrm{R})$ *for* $1 < i < r$ *and* a_1, a_2 *are not both zero in case* $r = 2$.

In some rings it is possible to shorten (2.11) still further by an application of (2.10), but for many rings such a reduction is impossible and (2.11) actually represents a normal form for the elements of $\mathbf{GE}_2(\mathrm{R})$ (cf. §§ 4 and 7). Such a ring is said to have a *unique standard form for* $\mathbf{GE}_2$. We shall also meet rings R for which the only relation in $\mathbf{GE}_2(\mathrm{R})$ of the form $W = I$, where W is a word in standard form, is the trivial relation $I = I$. Such a ring is said to be *quasi-free* for $\mathbf{GE}_2$. It is clear that of the properties:

(i) R has a unique standard form for $\mathbf{GE}_2$,

(ii) R is quasi-free for $\mathbf{GE}_2$,

(iii) R is universal for $\mathbf{GE}_2$,

each implies the next, but as examples to be given later show, these three classes are distinct.

3. Direct and free products of GE-rings.

We now turn to consider ring constructions which preserve the property of being a GE-ring. Here it is more convenient not to restrict the size of the matrices in any way.

For any ring R and any integer $n \geqslant 1$ we may define $\mathbf{GE}_n(\mathrm{R})$, $\mathbf{D}_n(\mathrm{R})$ and $\mathbf{E}_n(\mathrm{R})$ as the subgroups of $\mathbf{GL}_n(\mathrm{R})$ generated by all elementary matrices, all diagonal matrices and all $B_{ij}(a)$ $(a \in \mathrm{R}, i \neq j)$ respectively. As in the case $n = 2$, it is easily verified that $\mathbf{E}_n\mathbf{D}_n = \mathbf{D}_n\mathbf{E}_n = \mathbf{GE}_n(\mathrm{R})$, and hence in every GE-ring, every $A \in \mathbf{GL}_n(\mathrm{R})$ has form

$$A = DB^{(1)} \dots B^{(r)}, \tag{3.1}$$

where $D \in \mathbf{D}_n$ and $B^{(\rho)} = B_{ij}(a)$ for some $i \neq j$, $a \in \mathrm{R}$ (depending on ρ).

Let R be a direct product of a family of rings, say $\mathrm{R} = \prod \mathrm{R}_\lambda$, and denote the canonical projection $\mathrm{R} \to \mathrm{R}_\lambda$ by ε_λ. Then ε_λ is a homomorphism which induces a group homomorphism $\varepsilon_\lambda^* : \mathbf{GL}_n(\mathrm{R}) \to \mathbf{GL}_n(\mathrm{R}_\lambda)$. Composing these homomorphisms, we obtain a mapping

$$\varepsilon^* : \mathbf{GL}_n(\mathrm{R}) \to \Pi\mathbf{GL}_n(\mathrm{R}_\lambda),$$

which is easily seen to be an isomorphism. Now suppose that each R_λ is a GE-ring; then R need not be a GE-ring. We need only take a product of infinitely many factors R_λ such that in R_λ the integer r in (3.1) cannot be taken to be bounded, as λ varies [1]. Then it is easy to write down a matrix $\mathrm{A}\in\mathbf{GL}_n(\mathrm{R})$ which is not of the form (3.1). However, when the number of factors R_λ is finite and each is a GE-ring, then any $\mathrm{A}\in\mathbf{GL}_n(\mathrm{R})$ is again of the form (3.1) and we obtain

Theorem **(3.1)**. — *The direct product of a finite number of* GE-*rings is again a* GE-*ring.*

The result holds more generally for the direct sum (= weak direct product) of any number of GE-rings, since any matrix over the direct sum lies in the direct product of a finite number of factors.

We next suppose that we are dealing with K-algebras, K being a commutative ring with 1, and ask whether the GE-property is preserved under tensor products. The answer is easily seen to be " no ", even when K is a field, since K[x] is a GE-ring, for any indeterminate x over K, whereas the polynomial ring in two indeterminates, $\mathrm{K}[x,y] = \mathrm{K}[x]\otimes\mathrm{K}[y]$, is not (see § 7). In fact the analogy with the Euclidean algorithm suggests that we consider, not the tensor product but the free product (cf. [5]). Of course it is now more natural to replace the K-algebras by K-rings, where K is any ring with 1. It will be recalled that a K-ring is a ring R with a canonical homomorphism $\theta : \mathrm{K} \to \mathrm{R}$. In case θ is injective, R is said to be a *strict* K-ring: thus a strict K-ring is just a ring in which K is embedded in a canonical way. In considering free products of rings we shall assume that all the factors are strict K-rings for a fixed K, which may thus be considered as a common subring of all the factors.

To answer the question whether the GE-property is preserved by free products, would require a formidable calculation; we therefore confine ourselves to establishing the result in a special case (which turns out to be sufficient for many applications). Namely, we shall assume in addition that our rings are locally free ideal rings, i.e. semifirs [2]. The GE-rings which are semifirs are characterized in the following

Proposition **(3.2)**. — *For any ring* R, *the following assertions are equivalent:*

(i) R *is a* GE-*ring and a semifir,*

(ii) *for any* $n\geqslant 1$, *given* $a_1, \ldots, a_n, b_1, \ldots, b_n\in\mathrm{R}$, *if*

$$\Sigma a_i b_i = 0, \qquad b_1, \ldots, b_n \textit{ not all zero,} \tag{3.2}$$

then there exists $\mathrm{C}\in\mathbf{E}_n(\mathrm{R})$ *such that* $(a_1, \ldots, a_n)\mathrm{C}$ *has at least one zero coordinate.*

[1] We shall see below (in § 5) that there even exist GE_2-rings in which r cannot be taken bounded.

[2] Cf. [6], where they are called local firs. This name has now been abandoned as it may give rise to confusion with local rings. Thus a *semifir* is an integral domain (not necessarily commutative) in which all finitely generated right ideals are free and any two bases of a given free module have the same cardinal.

A ring satisfying either (and hence both) of (i), (ii) will be described as a GE-*semifir*, for brevity. We note that H. Bass in [3] takes (ii) as the definition of a generalized Euclidean ring.

Proof. — Suppose (i) holds and (3.2) is given, then by Theorem 2.6 of [6] there exists $A \in \mathbf{GL}_n(R)$ such that $(a_1, \ldots, a_n)A$ has a zero coordinate. Since R is a GE-ring, $A^{-1} = DE$, where $D \in \mathbf{D}_n$, $E \in \mathbf{E}_n$. Hence $(a_1, \ldots, a_n)E^{-1}D^{-1}$ has a zero coordinate, and so does $(a_1, \ldots, a_n)E^{-1}$. Thus R satisfies (ii).

Conversely, assume (ii); then R is a semifir, again by Theorem 2.6 of [6], so it remains to prove that GE_n holds and here we may clearly assume that $n > 1$. Let A be an invertible $n \times n$ matrix; if the first row has only one non-zero element then this must be a unit. By permuting the columns of A (which corresponds to right multiplication by a matrix in $\mathbf{E}_n$) we may bring this unit to the (1,1)-position and obtain

$$A = \begin{pmatrix} \alpha & 0 \quad 0 \quad \ldots \quad 0 \\ * & A_1 \end{pmatrix}$$

where $A_1 \in \mathbf{GL}_{n-1}(R)$. Now the result follows by induction on n. If A has precisely $r(>1)$ non-zero elements in the first row, we may assume that these elements come in the first r places. Writing $A = (a_{ij})$, $A^{-1} = (\breve{a}_{ij})$, we have, for $j = 2, \ldots, n$,

$$\sum_{i=1}^{r} a_{1i}\breve{a}_{ij} = 0. \tag{3.3}$$

Since A^{-1} is invertible, either $\breve{a}_{1j} \neq 0$ for some $j \geqslant 2$ or $a_{2j} \neq 0$ for some $j \geqslant 2$. Choosing j accordingly, we may assume that of the a_{ij} actually occurring in (3.3), not all vanish. By (ii) there exists $B_1 \in \mathbf{E}_r(R)$ such that $(a_{11}, \ldots, a_{1r})B_1$ has a zero coordinate. Hence

$$A\begin{pmatrix} B_1 & 0 \\ 0 & I_{n-r} \end{pmatrix}$$

has fewer than r non-zero elements in the first row and after at most $n - 1$ steps we reach the case $r = 1$ treated before. This completes the proof.

In connexion with this proposition we remark that a GE-ring need not be a semifir. For by Theorem 3.1, a GE-ring need not even be an integral domain. To obtain an example of a GE-ring which is an integral domain but not a semifir, take any commutative local domain (i.e. an integral domain with a single maximal ideal) which is not a valuation ring. This is not a Bezout ring and hence not a semifir, but as we shall see later, any local ring is a GE-ring (§ 4 below; see also [12]). In the other direction we note that a semifir need not be a GE-ring; an example of a commutative principal ideal domain which is not a GE-ring will be given in § 6.

We now show that the free product of any family of GE-semifirs over a skew field, is again a GE-semifir. Let R be any strict K-ring; by a suitable identification we may regard K as a subring of R. If K is a direct summand of R, as right K-module,

$$R = K \oplus N,$$

then R is called an *augmented* K-ring, with *augmentation module* N. It was shown in [6] that for any ring K, the free product of any family of augmented K-rings exists and is again an augmented K-ring. We shall use the terminology and notation for free products introduced in [6]; in particular (H^n) is the usual filtration of the free product (by height) and two elements a, b of the free product are said to *interact* if $h(ab) < h(a) + h(b)$. In any ring R, two n-tuples $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ are said to be GE-*equivalent* if there is a matrix $P = (p_{ij}) \in \mathbf{GE}_n(R)$ such that $b_j = \Sigma a_i p_{ij}$.

Lemma (3.3). — *Let* K *be a field (possibly skew) and* (R_λ) *a family of* GE-*semifirs. Denote the free product of the* R_λ *(over* K*) by* P *and let* (H^n) *be the filtration by height of* P. *Let* a_k, b_k $(k = 1, \ldots, r)$ *be any elements of* P *such that* $h(a_k b_k) = n$ *and*

$$\Sigma a_k b_k \equiv 0 \qquad (\text{mod } H^{n-1}). \tag{3.4}$$

Further, assume that the a_k *are ordered by decreasing height, say* $h(a_i) = m$ *for* $i \leqslant s$ *and* $h(a_j) < m$ *for* $j > s$. *Then either*

a) $(a_1, \ldots, a_s)$ *is* GE-*equivalent over* K *to an s-tuple* $(a'_1, \ldots, a'_s)$ *such that* $h(a'_i) \leqslant m$ *with strict inequality for at least one i, or*

b) *the elements* a_i $(i \leqslant s)$ *which interact with* b_i *in a given factor* R_λ *are* GE-*equivalent over* R_λ *to a tuple of elements of height* $\leqslant m$, *with strict inequality in at least one place, or*

c) *for each* a_i $(i \leqslant s)$ *there exist elements* $x_{ij} \in P$ *such that*

$$h(a_i - \Sigma a_j x_{ji}) < m, \qquad h(a_j x_{ji}) \leqslant m \qquad (j > s).$$

The proof of this lemma will be omitted as it is very similar to that of Lemma 4.1 of [6], the only difference being that instead of unimodular equivalence we assume and prove GE-equivalence throughout [1]. As in [6], Theorem 4.2 we now deduce

Theorem (3.4). — *Let* K *be a field (possibly skew) and* (R_λ) *a family of* GE-*semifirs. Then the free product of the* R_λ *over* K *is again a* GE-*semifir.*

In particular, all the examples of semifirs given in [6] are actually GE-semifirs. For all arise as free products whose factors are either fields or rings of the form $k[t]$ or $k[t, t^{-1}]$ and these are all clearly GE-rings. In particular, it follows that the group algebra over k of a free group (and the semigroup algebra of a free semigroup) are GE-semifirs, a fact used by H. Bass in [3].

[1] Note that the phrase " set including an element of height less than m " occurring twice in Lemma 4.1 of [6] should each time be replaced by: " tuple of elements of height $\leqslant m$, with strict inequality in at least one place ". Further, K should be a skew field. In fact this is what was proved (and used) there.

4. $\mathbf{GL}_2(\mathbf{R})$ for a local ring.

It is well known that any field (not necessarily commutative) is a GE-ring (cf. e.g. Dieudonné [8]) and this result was generalized to local rings by Klingenberg [12]. In the present section we shall prove that a local ring is in fact a universal GE_2-ring; however it does not have a unique standard form and in fact is not even quasi-free for $\mathbf{GE}_2$ (unless it is a field).

To investigate the uniqueness of the form

$$[\alpha, \beta]\mathrm{E}(a_1)\ldots\mathrm{E}(a_r)=\mathrm{I} \qquad (\alpha, \beta\in\mathbf{U}(\mathrm{R}), \quad a_i\notin\mathbf{U}_0(\mathrm{R}) \qquad 1<i<r), \tag{4.1}$$

for any ring R, let us look at low values of r. Clearly, when $r=0$, we must have $\alpha=\beta=1$. The case $r=1$ is impossible, as we see by comparing (1,2)-elements. When $r=2$, we have $\alpha a_1=\beta a_2=0$, hence $a_1=a_2=0$ and $\alpha=\beta=-1$. For $r=3$ a comparison of (2,2)-elements gives $\beta a_2=1$, hence a_2 is a unit, which contradicts the conditions in (4.1). Thus in any ring R the relation (4.1) (with the stated conditions) is possible only if $r\geqslant 4$, apart from trivial cases.

Next let $r=4$ and for convenience write the relation as

$$\mathrm{E}(a_1)\mathrm{E}(a_2)\mathrm{E}(a_3)\mathrm{E}(a_4)=[\alpha, \beta].$$

A comparison of terms shows that this holds if and only if

$$\begin{aligned} a_1a_2a_3a_4-a_1a_2-a_3a_4-a_1a_4+1&=\alpha,\\ a_1a_2a_3-a_1-a_3&=0,\\ a_2a_3a_4-a_2-a_4&=0,\\ a_2a_3-1&=-\beta. \end{aligned}$$

We shall now show that in any local ring R which is not a field, these relations can be satisfied by elements $\alpha, \beta\in\mathbf{U}(\mathrm{R})$, $a_2, a_3\notin\mathbf{U}_0(\mathrm{R})$. For by hypothesis, $\mathbf{U}_0(\mathrm{R})\neq\mathrm{R}$; choose any elements a_2, a_3 not in $\mathbf{U}_0(\mathrm{R})$, then $\beta=1-a_2a_3\in\mathbf{U}(\mathrm{R})$ and a_1, a_4 are given by the equations

$$a_1\beta+a_3=0, \qquad \beta a_4+a_2=0,$$

which determine a_1, a_4 uniquely as elements of R; moreover it is clear that a_1, a_4 so defined do not lie in $\mathbf{U}_0(\mathrm{R})$. The fourth relation is of the form $\alpha=1+$nonunit, and it determines $\alpha\in\mathbf{U}(\mathrm{R})$. Thus for $r=4$, (4.1) can be satisfied in any local ring not a field, hence such a ring is not quasi-free for $\mathbf{GE}_2$.

Let R be a ring in which any one-sided inverse is two-sided and which satisfies the following condition: for any $n>1$, given $a_1, \ldots, a_n\in\mathrm{R}$ such that $\Sigma a_i\mathrm{R}=\mathrm{R}$, there exist $b_1, \ldots, b_{n-1}\in\mathrm{R}$ such that $\Sigma(a_i+a_nb_i)\mathrm{R}=\mathrm{R}$. This is expressed by saying that $n=1$ is a stable range for R; any such ring is easily seen to be a GE-ring (cf. H. Bass [2], who proves this and also shows that these rings include any ring which modulo its Jacobson radical satisfies the minimum condition on left ideals). The case $n=2$ of the above condition states that for any $a, b\in R$ such that $aR+bR=R$, there exists $c\in R$ such that $a+bc\in \mathbf{U}(R)$; clearly this holds in any local ring. Assuming this condition we can easily show that any $A\in \mathbf{GL}_2(R)$ can be written in the form (2.11) with $r\leqslant 3$. For by hypothesis, for any $B\in \mathbf{GL}_2(R)$ we can find $u\in R$ such that $BE(u)^{-1}$ has a unit in the (1,2)-place and hence, for suitable $v\in R$, $BE(u)^{-1}E(v)^{-1}$ has 0 in the (1,2)-place. This matrix is therefore of the form

$$\begin{pmatrix}\alpha & 0\\ c & \beta\end{pmatrix}=-[\alpha, \beta]E(0)E(\beta^{-1}c).$$

Applying this argument with $E(0)A$ in place of B, we find that

$$E(0)A=-[\alpha, \beta]E(0)E(\beta^{-1}c)E(v)E(u), \qquad \text{i.e.}$$
$$A=-[\beta, \alpha]E(\beta^{-1}c)E(v)E(u),$$

which is of the required form. This shows in particular that any local ring is a GE_2-ring. We now come to

Theorem **(4.1)**. — *Any local ring is a universal* GE_2*-ring.*

After what has been said it only remains to show that any relation (4.1) in a local ring is a consequence of the defining relations (2.2-4) and the relations in $\mathbf{D}_2$. Let

$$\textbf{(4.2)} \qquad [\alpha, \beta]E(a_1)\dots E(a_r)=I, \qquad \alpha, \beta\in \mathbf{U}(R), \qquad a_i\notin \mathbf{U}_0(R) \qquad \text{for } 1<i<r.$$

If $i\geqslant 4$, then $a_2, a_3\notin \mathbf{U}_0(R)$ and hence $a_2-1, a_3-1\in \mathbf{U}(R)$, say $a_2=\gamma+1$, $a_3=\delta+1$. By (2.10), we can reduce (4.2) to

$$-[\alpha, \beta]E(a_1-\gamma^{-1})D(\gamma)E(-1-\gamma^{-1}-\delta^{-1})D(\delta)E(a_4-\delta^{-1})E(a_5)\dots E(a_r)=1.$$

Using (2.4) to pull $D(\gamma)$ and $D(\delta)$ through to the left and then (2.9) (as in the proof of Theorem 2.2) we obtain an expression of the form (4.2) but with r replaced by a smaller value. This process can be continued as long as $r\geqslant 4$; so we finally reduce (4.2) to the case $r\leqslant 3$ and only (2.2-4) and their consequences have been used in the reduction. For $r\leqslant 3$ we saw that the only relation (4.2) is $-E(0)^2=I$ and by (2.5) this is also a consequence of (2.2-4). Hence (4.2) itself is a consequence of (2.2-4) (and relations in $\mathbf{D}$), as we wished to show.

5. Discretely normed rings.

We now look for more general conditions on our ring R to ensure that the form (2.11) for the elements of $\mathbf{GE}_2(R)$, given by Proposition 2.2 is unique. The most natural way of imposing such conditions is by means of a norm; this is of course closely related to the question of the existence of a Euclidean algorithm, or a weak algorithm (cf. [5]) in R and hence to the question whether R is a Euclidean ring. In fact we shall use this method to construct rings which are not Euclidean in a rather strong sense: they are not GE-rings.

Definition. — *A* norm *on* R *is a mapping* $|\ |$ *from* R *to the nonnegative real numbers such that*

N. 1. $|x|=0$ *if and only if* $x=0$,
N. 2. $|x+y|\leqslant|x|+|y|$,
N. 3. $|xy|=|x|\,|y|$.

We note that by N. 3, R must be an integral domain (not necessarily commutative). We shall say that $|\ |$ is a *discrete norm* or that R is *discretely normed*, if further,

N. 4. $|x|\geqslant 1$ for all $x\neq 0$, with equality only if $x\in\mathbf{U}(\mathrm{R})$,
N. 5. there exists no $x\in\mathrm{R}$ such that $1<|x|<2$.

If α is any unit, then by N. 4, $|\alpha|\geqslant 1$, $|\alpha^{-1}|\geqslant 1$, while by N. 3,

$$|\alpha|.|\alpha^{-1}|=|\alpha\alpha^{-1}|=1,$$

hence $|\alpha|=1$ for all $\alpha\in\mathbf{U}(\mathrm{R})$. Now N. 5 shows that $|x|\geqslant 2$ for any $x\notin\mathbf{U}_0(\mathrm{R})$. Conversely, the two conditions

N. 4′ $|\alpha|=1$ for all $\alpha\in\mathbf{U}(\mathrm{R})$,
N. 5′ $|x|\geqslant 2$ for all $x\notin\mathbf{U}_0(\mathrm{R})$

imply N. 4 and N. 5.

In order to derive further consequences of N. 1-5 we need an expression for the product $\mathrm{E}(a_1)\ldots\mathrm{E}(a_r)$; this is closely related to the chain of equations in the Euclidean algorithm, cf. § 8. Let $t_1, t_2, \ldots$ be any noncommuting indeterminates and define a sequence of polynomials in the t's with integer coefficients recursively by the equations

$$\textbf{(5.1)}\qquad \begin{cases} e_{-1}=0, \quad e_0=1, \\ e_n(t_1, \ldots, t_n)=e_{n-1}(t_1, \ldots, t_{n-1})t_n-e_{n-2}(t_1, \ldots, t_{n-2}). \end{cases}$$

We note that for $n\geqslant 0$, the suffix of e_n just indicates the number of arguments and so may be omitted when the arguments are given explicitly; we shall do so in what follows and only write the suffix when the arguments are omitted. We assert that

$$\textbf{(5.2)}\qquad \mathrm{E}(a_1)\ldots\mathrm{E}(a_r)=\begin{pmatrix} e(a_1, \ldots, a_r) & e(a_1, \ldots, a_{r-1}) \\ -e(a_2, \ldots, a_r) & -e(a_2, \ldots, a_{r-1}) \end{pmatrix}.$$

This is clear for $r=1$ and in the general case follows by induction, since, writing $e_i=e(a_1, \ldots, a_i)$, $e'_i=e(a_2, \ldots, a_{i+1})$, we have

$$\begin{pmatrix} e_{r-1} & e_{r-2} \\ -e'_{r-2} & -e'_{r-3} \end{pmatrix}\begin{pmatrix} a_r & 1 \\ -1 & 0 \end{pmatrix}=\begin{pmatrix} e_r & e_{r-1} \\ -e'_{r-1} & -e'_{r-2} \end{pmatrix}.$$

From the symmetry of (5.2) it is easy to see that the e's may also be defined by $e_{-1}=0$, $e_0=1$,

$$e_n(t_1, \ldots, t_n)=t_1e_{n-1}(t_2, \ldots, t_n)-e_{n-2}(t_3, \ldots, t_n).$$

The following lemma is basic for all uniqueness questions in $\mathbf{GE}_2(\mathrm{R})$ where R is discretely normed.

Lemma **(5.1).** — *Let* $a_1, \ldots, a_r$ $(r>1)$ *be any elements of a discretely normed ring* R *such that* $a_i \notin \mathbf{U}_0(\mathrm{R})$ *for* $i>1$. *Then*

$$|e(a_1, \ldots, a_r)| \geqslant |e(a_1, \ldots, a_{r-1})| \tag{5.3}$$

and

$$e(a_1, \ldots, a_r) \neq 0. \tag{5.4}$$

If, moreover, the norm on R *is such that*

$$\text{for any } u\in\mathrm{R},\ |u|=1 \text{ and } |1+u|=2 \text{ imply } u=1, \tag{5.5}$$

then the inequality in (5.3) *is strict, unless* $|2|=2$ *and* $a_1=\alpha\in\mathbf{U}(\mathrm{R})$, $a_2=a_4=\ldots=2\alpha^{-1}$ *and* $a_3=a_5=\ldots=2\alpha$. *In this case,*

$$\mathrm{E}(a_1)\ldots\mathrm{E}(a_r) \quad \textit{equals} \quad \begin{pmatrix} 1 & \alpha \\ * & * \end{pmatrix} \quad \textit{or} \quad \begin{pmatrix} \alpha & 1 \\ * & * \end{pmatrix}$$

according as r is even or odd.

Proof. — Write $e_i=e(a_1, \ldots, a_i)$ for short. Then for $i>1$, $|a_i|\geqslant 2$, hence

$$|e_i|=|e_{i-1}a_i-e_{i-2}|\geqslant 2|e_{i-1}|-|e_{i-2}|, \qquad \text{i.e.}$$

$$|e_i|-|e_{i-1}|\geqslant |e_{i-1}|-|e_{i-2}| \qquad (i>1). \tag{5.6}$$

Suppose first that $a_1=0$; then $e_1=a_1=0$, $e_2=a_1a_2-1=-1$ and hence by (5.6),

$$|e_r|-|e_{r-1}|\geqslant |e_2|-|e_1|=1.$$

This proves (5.3) (with strict inequality) and hence (5.4).

Now let $a_1\neq 0$, then $|a_1|\geqslant 1$ and

$$|e_1|-|e_0|=|a_1|-1\geqslant 0. \tag{5.7}$$

Combining this with (5.6) we see that $|e_i|-|e_{i-1}|\geqslant 0$ for $i>1$, and hence

$$|e_r|\geqslant |e_{r-1}|\geqslant \ldots \geqslant |e_1|\geqslant 1. \tag{5.8}$$

This proves (5.3) and (5.4). If further, $|a_1|>1$, then we have strict inequality in (5.7) and hence in (5.3). Likewise, if $|a_i|>2$ for some $i>1$, then we have strict inequality in the corresponding formula (5.6) and hence in (5.3). There remains the case $|a_1|=1$, $|a_i|=2$ $(i>1)$; in particular, a_1 is then a unit. Suppose first that $a_1=1$ and assume that (5.5) holds. If we have equality in (5.3), then $|a_1a_2-1|=1$, so $a_1a_2=1+b$, where $|b|=1$, $|1+b|=|a_1a_2|=2$, hence $b=1$ and $a_2=a_1a_2=2$. Now

$$\mathrm{E}(1)\mathrm{E}(2)^k=\begin{pmatrix} 1 & 1 \\ -k-1 & -k \end{pmatrix}$$

and by induction it follows that all the a_i $(i>1)$ must be 2. In general, when $a_1=\alpha$ is any unit, we have

$$\begin{aligned}[\alpha^{-1}, 1]\mathrm{A} &= [\alpha^{-1}, 1]\mathrm{E}(\alpha)\mathrm{E}(a_2)\ldots\mathrm{E}(a_r)\\ &= \mathrm{E}(1)\mathrm{E}(a_2\alpha)\mathrm{E}(\alpha^{-1}a_3)\mathrm{E}(a_4\alpha)\ldots\end{aligned}$$

with a factor $D=[1, \alpha^{-1}]$ (for r odd) or $[\alpha^{-1}, 1]$ (for r even) at the right-hand end, as follows by repeated application of (2.4). Now if the elements of the first row of A have the same norm, then the same holds for $[\alpha^{-1}, 1]AD^{-1}$, so we are reduced to the case $a_1=1$ and the result follows.

Theorem **(5.2)**. — *Any discretely normed ring is quasi-free for* $\mathbf{GE}_2$, *and hence is universal for* $\mathbf{GE}_2$.

Proof. — By Theorem 2.2, any relation in $\mathbf{GE}_2(R)$ can be brought to standard form

$$(5.9) \qquad [\alpha, \beta]E(a_1)\ldots E(a_r)=I,$$

where $a_i \notin \mathbf{U}_0(R)$ for $1<i<r$. Comparing the (1,2)-elements in this equation we see that

$$e(a_1, \ldots, a_{r-1})=0,$$

and this contradicts Lemma 5.1 if $r-1>1$, i.e. $r>2$. But when $r\leqslant 2$, only the trivial cases noted in § 4 are possible. This completes the proof.

Corollary. — *A discretely normed ring which is a* GE_2*-ring is a universal* GE_2*-ring.*

When R is a discretely normed ring, it is easy to obtain an explicit form for the involutions in $\mathbf{GE}_2(R)$, i.e. the elements of order two. This will be useful later on, in the study of automorphisms of $\mathbf{GE}_2(R)$ and also in showing that for certain Dedekind rings the class number must be 1. We begin by determining the centre of $\mathbf{GE}_2(R)$, where R may be any ring:

Proposition **(5.3)**. — *For any ring* $R(\neq 0)$ *the centralizer of* $\mathbf{E}_2(R)$ *in* $\mathbf{GL}_2(R)$ *consists of all matrices* λI, *where* λ *runs over the central units of* R.

Proof. — Let $A=\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{GL}_2(R)$ centralize $\mathbf{E}_2(R)$; since A commutes with E(0) we must have $a=d$, $b+c=0$; since it commutes with E(1) we have $b=0$, i.e. $A=aI$ and now the commutativity with $E(x)$ $(x \in R)$ shows that a lies in the centre of R. This completes the proof.

We next determine how the standard form simplifies in the case of elements of finite order.

Proposition **(5.4)**. — *Let* R *be a ring which is quasi-free for* $\mathbf{GE}_2$; *then any matrix in* $\mathbf{GE}_2(R)$ *which is of finite order modulo the centre of* $\mathbf{GE}_2(R)$ *is conjugate* (*under* $\mathbf{E}_2(R)$) *to an element of the form*

$$(5.10) \qquad [\alpha, \beta]E(a)$$

or

$$(5.11) \qquad [\alpha, \beta]E(0)E(b),$$

and in case (5.10) *holds, we have* $a \in \mathbf{U}_0(R)$.

Proof. — Let $A \in \mathbf{GE}_2(R)$ be of finite order modulo the centre of $\mathbf{GE}_2(R)$ and write

$$A = [\alpha, \beta]E(a_1) \dots E(a_r). \tag{5.12}$$

Then any conjugate of A is again of finite order mod the centre. If $r \geqslant 3$ in (5.12) and $a_1 \in \mathbf{U}_0(R)$, then A is conjugate to

$$[\beta, \alpha]E(a')E(a_1) \dots E(a_{r-1}),$$

for some $a' \in R$, and now the length can be reduced. Likewise, if $a_r \in \mathbf{U}_0(R)$, we can transfer the factor $E(a_1)$ to the right-hand end and then make a reduction. Thus we obtain a form (5.12) for a conjugate of A, in which either $r \leqslant 2$ or $a_i \notin \mathbf{U}_0(R)$ for all i. Since $A^n = \lambda I$, we have

$$[\alpha, \beta]E(a_1) \dots E(a_r)[\alpha, \beta]E(a_1) \dots \dots \dots E(a_r) = \lambda I.$$

By pulling all the diagonal factors through to the left and then transferring them to the right-hand side, we obtain a relation

$$E(b_1) \dots E(b_{rn}) = \text{diagonal matrix},$$

where each b_j is associated to some a_i. By hypothesis this can only hold if some $b_j \in \mathbf{U}_0(R)$, hence some $a_i \in \mathbf{U}_0(R)$ and so in particular, $r \leqslant 2$. For $r = 1$, we obtain the form (5.10), where necessarily $a_1 \in \mathbf{U}_0(R)$. If $r = 2$, we may without loss of generality assume that $a_1 \in \mathbf{U}_0(R)$. Suppose first that $a_1 = \gamma \in \mathbf{U}(R)$. Then transforming A by $E(x)$, we have

$$\begin{aligned} E(x)AE(x)^{-1} &= E(x)[\alpha, \beta]E(\gamma)E(b)E(0)E(-x)E(0) \\ &= -[\beta, \alpha]E(\beta^{-1}x\alpha - \gamma^{-1})D(\gamma)E(b - \gamma^{-1} - x)E(0). \end{aligned}$$

Choose $x = b - \gamma^{-1}$, then the right-hand side becomes

$$[\alpha', \beta']E(c)$$

and this is again of the form (5.10). In the alternative case, $a_1 = 0$ and we obtain (5.11). Finally if in (5.12) $r = 0$, then this is of the form (5.11), with $b = 0$.

We remark that (5.10) is conjugate, under $\mathbf{GE}_2(R)$, to a matrix of the form

$$[1, \beta]E(a);$$

for we can reduce α to 1 by transforming by $[\alpha, 1]$.

It is now easy to derive a standard form for the involutions in $\mathbf{GE}_2(R)$, when R is quasi-free for $\mathbf{GE}_2$.

Theorem **(5.5).** — *Let* R *be an integral domain which is quasi-free for* $\mathbf{GE}_2$, *then any involution* $\neq -I$ *in* $\mathbf{GE}_2(R)$ *is conjugate to one of the form*

$$\pm\begin{pmatrix} 1 & 0 \\ h & -1 \end{pmatrix} \tag{5.13}$$

where h runs over a transversal of 2R *in* R *(qua additive group).*

In particular, if $2R = R$, *then every non-central involution is conjugate to* $[1, -1]$.

Proof. — By Prop. 5.4 and the remark following it, any involution $A \neq -I$ is conjugate to one of the forms

$$\begin{pmatrix} a & 1 \\ -\beta & 0 \end{pmatrix} \qquad \begin{pmatrix} \alpha & 0 \\ \beta b & \beta \end{pmatrix}.$$

In the first case, by squaring, we see that $[\beta, \beta]E(\beta^{-1}a)E(a) = I$, hence $a = 0$, $\beta = -1$ and we find $A = [1, -1]E(0)$. Transforming by $[1, -1]E(1)E(0)^{-1}$ we reach (5.13) with $h = 1$. In the second case squaring shows that $\alpha^2 = \beta^2 = 1$, hence $\alpha, \beta = \pm 1$ (because R is an integral domain) and $\beta(\alpha + \beta)b = 0$. Thus either $-\beta = \alpha = \pm 1$ and we have the form (5.13) or $\alpha = \beta = \pm 1$ and $b = 0$. But this means that $A = \pm I$ and these possibilities were excluded. Thus every involution can be brought to the form (5.13). Now transforming (5.13) by $B_{21}(c)$ we obtain

$$\pm \begin{pmatrix} 1 & 0 \\ h - 2c & -1 \end{pmatrix}$$

and by a suitable choice of c we can ensure that $h_1 = h - 2c$ belongs to the given transversal of $2R$ in R. In particular when $2R = R$ this means that every involution is conjugate to $\pm [1, -1]$; since $[-1, 1]$ is transformed to $[1, -1]$ by $E(0)$, we need only one of these forms and the proof is complete.

A ring is said to be *2-torsion free* if $2x = 0$ implies $x = 0$. With this definition we note the following application:

Proposition **(5.6)**. — *Let* R *be a 2-torsion free integral domain which is a quasi-free* GE_2*-ring. Then every projective module on two generators is free.*

Proof. — Any projective module M on two generators is a direct summand of R^2 and may be characterized by its projection, an idempotent 2×2 matrix, E say. The matrix $P = I - 2E$ has square I; if $P = \pm I$, then $E = 0$ or I, and M is clearly free. Otherwise P is an involution $\neq -I$ and by Theorem 5.5, after applying a suitable inner automorphism, we have

$$P = I - 2E = \pm \begin{pmatrix} 1 & 0 \\ h & 1 \end{pmatrix}.$$

In particular $h \equiv 0 \pmod{2R}$, hence by a further inner automorphism we may arrange that $P = \pm [1, -1]$; then $E = [1, 0]$ or $[0, 1]$ and M is again free.

In a Dedekind domain each ideal is projective and can be generated by two elements; moreover, the free ideals are just the principal ideals, hence we obtain the

Corollary. — *Let* R *be a 2-torsion free discretely normed Dedekind ring. If* R *is a* GE_2*-ring, then it must be a principal ideal domain.*

In the next section we shall see how the result may be applied to purely imaginary quadratic number fields. We conclude this section by giving some examples of discretely normed rings.

1. The rational integers, with the usual absolute value. More generally, consider the ring of algebraic integers in any imaginary quadratic number field. If this is

embedded in the complex numbers, the usual absolute value is a norm, i.e. it satisfies N. 1-3. N. 4 is also satisfied because the integers form a lattice which can only have finitely many points in the unit circle; if any were strictly inside the unit circle, their powers would give an infinite set. Further, if an element is on the unit circle, so is its conjugate and hence the element is a unit. Thus N. 4 always holds; N. 5 holds with a few exceptions (see § 6). Whenever it is satisfied we have a discretely normed ring.

2. Let R be any ring such that $\mathbf{U}_0(R)$ is a field, k say, and suppose that we have a *degree-function* on R, i.e. to each $a \in R$ there corresponds a non-negative integer or $-\infty$, written $d(a)$, such that

D. 1. $d(a) = -\infty$ if and only if $a=0$,
D. 2. $d(a)=0$ if and only if $a \in \mathbf{U}(R)$,
D. 3. $d(a-b) \leqslant \max\{d(a), d(b)\}$,
D. 4. $d(ab) = d(a) + d(b)$.

Then R becomes a discretely normed ring if we put $|a| = 2^{d(a)}$. For example, the polynomial rings over k, in any number of indeterminates, and the free associative algebras (polynomial rings in non-commuting indeterminates) are of this form.

3. If R is any discretely normed ring, then $R[x]$, the polynomial ring in a single indeterminate over R, may be discretely normed by the rule

$$|\Sigma a_i x^i| = \Sigma |a_i| 2^i.$$

4. The free product of fields does not at first sight admit a norm, but it does have a filtration which is very nearly a norm, and many of the results proved here for discretely normed rings can be carried over for free products of fields. We shall not enter into the details here (cf. [5]).

In a discretely normed ring it is often possible to decide whether the ring is a GE_2-ring. If it happens to be Euclidean with respect to the norm, it is clearly a GE_2-ring. In the next two sections we shall show that under certain conditions a discretely normed ring cannot be a GE_2-ring; this means in particular that no norm function can be defined for which the ring is Euclidean.

6. Rings of algebraic integers and algebraic functions.

We have just seen that the ring of integers in an imaginary quadratic number field is discretely normed, provided that N. 5 holds. This amounts to the condition that no integer a in the ring satisfies

$$1 < N(a) < 4.$$

It is easily verified that this holds for the integers in $\mathbf{Q}(\sqrt{-d})$, where d is positive and squarefree, except when

$$d = 1, 2, 3, 7, 11. \tag{6.1}$$

As it happens, these are just the fields in which there is a Euclidean algorithm with respect to the usual norm $|a| = N(a)^{1/2}$ (cf. [9], ch. 14). For the remaining cases we have

Theorem **(6.1)**. — *Let d be a squarefree positive integer and* I *the ring of integers in* $\mathbf{Q}(\sqrt{-d})$. *Then* I *is not a* GE_2*-ring, unless d has one of the values* (6.1).

Proof. — We shall assume that I is a GE_2-ring for some d which is not among the values (6.1) and derive a contradiction. By what has been said, I is discretely normed by $|a| = N(a)^{1/2}$, and since I is also a 2-torsion free Dedekind ring, it is a principal ideal domain, by Proposition 5.6 Corollary. By Theorem 2.2, every element $A \in \mathbf{GL}_2(I)$ has the form

$$\textbf{(6.2)} \qquad A = [\alpha, \beta]E(q_1) \dots E(q_r), \qquad \text{where } |q_i| \geqslant 2 \text{ for } 1 < i < r.$$

Since $|\ |$ is the ordinary distance in the Euclidean plane, condition (5.5) of Lemma 5.1 is always satisfied. Moreover, the only units in I are ± 1; this is easily checked because the low values (6.1) of d have been excluded. Let us denote the first row of A by (a, b); then three cases are possible, according to the value of q_r.

(i) $|q_r| \geqslant 2$. Then by Lemma 5.1, $|a| > |b|$, unless $a = \pm b$ (recall that ± 1 are the only units in I).

(ii) $|q_r| = 0$ and hence $q_r = 0$. Applying Lemma 5.1 to $AE(0)^{-1}$, we see that $|a| < |b|$, unless $a = \pm b$.

(iii) $|q_r| = 1$ and hence $q_r = \pm 1$. In this case we find that

$$\textbf{(6.3)} \qquad |a \pm b| < |b|$$

for at least one choice of sign, unless $a \pm b = b$. Again this follows by considering $AE(q_r)^{-1}$.

Now suppose that $|a| = |b|$, but $a \neq \pm b$. Then the first two alternatives do not apply and $a \neq 0$, $2b$, so that (6.3) must hold for at least one choice of sign. Thus if we can find a and b, forming the first row of an invertible matrix, such that

$$\textbf{(6.4)} \qquad |a| = |b|, \qquad |a \pm b| \geqslant |b|,$$

we have a contradiction. Now in a principal ideal domain any pair of elements without a common factor forms the first row of an invertible matrix, and any element of the field of fractions can be written as a quotient of two elements without a common factor. So to satisfy (6.4) we only have to find an element $\alpha \in \mathbf{Q}(\sqrt{-d})$ such that

$$\textbf{(6.5)} \qquad |\alpha| = 1, \qquad |\alpha \pm 1| \geqslant 1;$$

for on writing $\alpha = a/b$ in reduced form, we obtain a solution of (6.4). Now the equation $|\alpha| = 1$ defines a circle in R^2 (relative to oblique coordinates) on which the rational points are clearly dense. These correspond to points $\alpha \in \mathbf{Q}(\sqrt{-d})$ and by going sufficiently far from the real axis we can ensure that the second relation in (6.5) is also satisfied (for both signs). This gives us the required pair of values (a, b) and it shows that I is not a GE_2-ring.

The first value of d not listed in (6.1) for which I is a principal ideal domain is $d=19$ (cf. [9], p. 213) and in this case an explicit matrix in $\mathbf{GL}_2(\mathrm{I})$ but not in $\mathbf{GE}_2(\mathrm{I})$ is

$$\begin{pmatrix} 3-\theta & 2+\theta \\ -3-2\theta & 5-2\theta \end{pmatrix},$$

where $\theta^2-\theta+5=0$. By going through the steps of the proof of Proposition 5.6, Corollary, we can construct a similar example for $\mathbf{Q}(\sqrt{-5})$.

In the language of valuation theory we can say that for a ring of integers in an algebraic number field to be discretely normed it is necessary for the field to have a single place at infinity. When this condition holds we can prove an analogue of Theorem 6.1 for function fields [1].

Theorem **(6.2)**. — *Let* K *be a field of functions of a single variable with k as field of constants and assume that k has characteristic not two. Let* $\mathfrak{p}$ *be any place of* K/k *of degree* $f_{\mathfrak{p}}>1$ *and denote by* I *the ring of elements of* K *which are integral everywhere except possibly at* $\mathfrak{p}$. *Then* I *is not a* GE_2*-ring.*

Proof. — If $\nu_{\mathfrak{p}}$ denotes the exponential valuation at $\mathfrak{p}$ then the product formula shows that

$$\nu_{\mathfrak{p}}(x)\leqslant 0 \qquad \text{for all } x\in\mathrm{I},\ x\neq 0,$$

with equality only if $x\in k$. If we define

$$\delta(x)=-\nu_{\mathfrak{p}}(x),$$

then δ is easily verified to be a degree function on I. Hence I is discretely normed and since it is 2-torsion free and Dedekind, we see that if I is a GE_2-ring, it must be a principal ideal domain. Further, every $\mathrm{A}\in\mathbf{GL}_2(\mathrm{I})$ has the form

$$\mathrm{A}=\begin{pmatrix} a & b \\ c & d \end{pmatrix}=[\alpha,\beta]\mathrm{E}(q_1)\ldots\mathrm{E}(q_r) \qquad \delta(q_i)>0 \qquad (1<i<r).$$

As in the proof of Theorem 6.1 there are three possibilities:

(i) $\delta(q_r)>0$: then $\delta(a)>\delta(b)$ or $a=\lambda b$ $(\lambda\in k)$,

(ii) $q_r=0$: then $\delta(a)<\delta(b)$ or $a=\lambda b$ $(\lambda\in k)$,

(iii) $\delta(q_r)=0$: then $q_r=\lambda\in k$ and $\delta(a-\lambda b)<\delta(b)$.

To avoid these three possibilities we need only find a pair of elements a, b in I without common factor such that $\delta(a-\lambda b)=\delta(b)$ for all $\lambda\in k$. Let $k(\mathfrak{p})$ be the residue field at $\mathfrak{p}$ and $x\to\bar{x}$ the residue mapping. By hypothesis, $k(\mathfrak{p})\neq k$, so there exists $\eta\in k(\mathfrak{p})$, $\eta\notin k$; choose $y\in\mathrm{K}$ such that $\bar{y}=\eta$. Then $\delta(y-\lambda)=0$ for all $\lambda\in k$ and as in the proof of Theorem 6.1 we can write $y=a/b$, whence $\delta(a-\lambda b)=\delta(b)$, so that none of (i)-(iii) can hold. Hence I cannot be a GE_2-ring.

When K is of genus zero, the condition of Theorem 6.2 is necessary and sufficient, for in that case $f_{\mathfrak{p}}=1$ implies that K is rational and I is just a polynomial ring over k.

[1] I am indebted to J. V. Armitage for a helpful discussion of this point.

7. k-rings with a degree function.

We now turn to look at the second example of discretely normed rings given in § 5. Let k be a field (possibly skew), then a k-ring with a degree function was defined in § 5; we stress that an element of degree zero must by definition lie in k. In § 5 we showed that such a ring is discretely normed. We shall now show that for these rings we can strengthen Theorem 5.2 by proving that the standard form is unique.

Theorem **(7.1)**. — *Let* R *be a k-ring with a degree function (k a field); then* R *has a unique standard form for* $\mathbf{GE}_2$.

Proof. — By Theorem 2.2 every $A\in\mathbf{GE}_2(R)$ can be brought to the form

$$\text{(7.1)}\qquad A=[\alpha,\beta]E(a_1)\dots E(a_r)\qquad (a_i\in R,\ \alpha,\beta\in\mathbf{U}(R)),$$

where $a_i\notin k$ $(1<i<r)$ and a_1, a_2 are not both zero in case $r=2$. Suppose that we also have

$$\text{(7.2)}\qquad A=[\gamma,\delta]E(b_1)\dots E(b_s)\qquad b_j\notin k\ (1<j<s);$$

we must show that $\alpha=\gamma$, $\beta=\delta$ and $r=s$, $a_i=b_i$. We may assume $r\geqslant s$ and then use induction on r; further, we may assume that $\gamma=\delta=1$, by replacing A by $[\gamma,\delta]^{-1}A$ if necessary. If $a_r=b_s$ we can cancel the last factor and use induction, so we may assume that $a_r\neq b_s$. Then

$$\begin{aligned}I&=\pm[\alpha,\beta]E(a_1)\dots E(a_r)E(0)E(-b_s)\dots E(-b_1)E(0)\\&=\pm[\alpha,\beta]E(a_1)\dots E(a_r-b_s)E(-b_{s-1})\dots E(-b_1)E(0).\end{aligned}$$

If we put $a_r-b_s=c$ and multiply by E(0) we get

$$\text{(7.3)}\qquad E(0)=DE(a_1)\dots E(a_{r-1})E(c)E(-b_{s-1})\dots E(-b_1),$$

where D is some diagonal matrix. Comparing elements in the (2,2)-place we find

$$\text{(7.4)}\qquad e(a_2,\dots,a_{r-1},c,-b_{s-1},\dots,-b_2)=0.$$

We now apply Lemma 5.1 and note that condition (5.5) is satisfied because for R the hypothesis is vacuous, and $|2|\leqslant 1$; therefore the inequality in (5.3) is always strict. Assume first that $s\geqslant 3$, then also $r\geqslant 3$ and we reach a contradiction unless $c\in k$. But then c must be a unit because it is not zero, and (7.3) takes the form

$$E(0)=DE(a_1)\dots E(a_{r-1}-c^{-1})D(c)E(-b_{s-1}-c^{-1})E(-b_{s-2})\dots E(-b_1).$$

By hypothesis, a_{r-1}, $b_{s-1}\notin k$, hence $a_{r-1}-c^{-1}$, $b_{s-1}-c^{-1}\notin k$. If we pull D(c) through to the left (by (2.4)) and equate the (2,2)-elements, we obtain

$$e(a'_2,\dots,a'_{r-2},a'_{r-1}-c,-b_{s-1}-c^{-1},-b_{s-2},\dots,-b_2)=0,$$

where a'_i is associated to a_i and hence none of the arguments lie in k; but this contradicts Lemma 5.1.

When $s=2$ we have

$$\begin{pmatrix} b_1b_2-1 & b_1 \\ -b_2 & -1 \end{pmatrix} = [\alpha, \beta]\mathrm{E}(a_1)\ldots\mathrm{E}(a_r),$$

hence $e(a_2, \ldots, a_{r-1}) = -\beta^{-1}$. By Lemma 5.1, $r-2=0$ and in this case the uniqueness follows easily.

When $s=1$,

$$\mathrm{E}(b) = [\alpha, \beta]\mathrm{E}(a_1)\ldots\mathrm{E}(a_r),$$

hence $e(a_2, \ldots, a_{r-1}) = 0$. By Lemma 5.1, $r-2=-1$, so $r=1$ and the uniqueness is then clear. The same argument applies when $s=0$, and this completes the proof.

We remark that the result does not hold for all discretely normed rings. E.g., for the ring $\mathbf{Z}$ of rational integers,

$$\mathrm{E}(2)\mathrm{E}(-2)\mathrm{E}(2) = -\mathrm{E}(3)\mathrm{E}(2)\mathrm{E}(3)\left[=\begin{pmatrix} -12 & -5 \\ 5 & 2 \end{pmatrix}\right].$$

In the next section we shall introduce a different generating set for $\mathbf{GL}_2(\mathbf{Z})$ in terms of which a unique normal form is possible.

We now give some more specific examples of rings with a degree-function which are GE-rings and some which are not. Naturally, any Euclidean ring (with respect to its degree function) is a GE-ring, and for this to hold the ring need not be commutative, e.g. the ring $k[x]$ of polynomials in a single indeterminate over a skew field is Euclidean. More generally, we have the rings with a weak algorithm described in [5]. We shall not recall the definition here but merely note that they are always k-rings with a degree function, for some field k. A typical example is the tensor ring over an arbitrary k-bimodule; in particular, any free associative algebra over a commutative field is of this form. For these rings we have

Theorem **(7.2)**. — *Any ring with a weak algorithm is a* GE-*ring.*

Proof. — By Prop. 3.2 we need only show: for any $n \geqslant 1$, given $a_1, \ldots, a_n$, $b_1, \ldots, b_n \in \mathrm{R}$ such that

$$(7.5) \qquad \Sigma a_i b_i = 0 \qquad b_i \text{ not all zero,}$$

then there exists $\mathrm{C} \in \mathbf{E}_n(\mathrm{R})$ such that $(a_1, \ldots, a_n)\mathrm{C}$ has at least one zero coordinate. The proof is by induction on $\Sigma d(a_i)$. By suitably renumbering the a's and b's we may assume that $d(a_ib_i) = m$ for $i \leqslant s$ and $d(a_jb_j) < m$ for $j > s$. Then (7.5) shows that $a_1, \ldots, a_s$ are right R-dependent, hence by the weak algorithm, an a_i of maximal degree, say a_1 is right R-dependent on the rest:

$$a_1 = \sum_2^s a_i c_i + a_1' \qquad d(a_1') < d(a_1), \qquad d(a_i c_i) \leqslant d(a_1).$$

This shows that the transformation from $(a_1, \ldots, a_n)$ to $(a_1', a_2, \ldots, a_n)$ is a product of elementary transformations and it diminishes the value of the sum $\Sigma d(a_i)$. Hence the result follows by induction.

Since a ring with a weak algorithm is discretely normed, such a ring is a universal GE_2-ring; it is also easily seen that $\mathbf{E}_n(R)$ is normal in $\mathbf{GL}_n(R)$.

Next, to give some examples of rings which are not GE-rings, we establish a necessary condition for a k-ring with a degree function to be a GE_2-ring. Let R be any k-ring with a degree function which is also a GE_2-ring. We saw that in this case (5.5) always holds and $|2|\leq 1$, therefore Lemma 5.1 shows that in the present case

$$d(e(a_1, \ldots, a_r)) > d(e(a_1, \ldots, a_{r-1}))$$

for any elements $a_1, \ldots, a_r$ $(r>1)$ such that $a_i \notin k$ for $i>1$. Let us call a pair of elements (a, b) a *regular row* if it can occur as the first row of an invertible matrix over R (when R is commutative this reduces to the notion of a unimodular row defined by Bass in [2]). Take any regular row (a, b) and let $A \in \mathbf{GL}_2(R)$ be a matrix in which it occurs as first row. Since R is a GE_2-ring, A can be written in the form (7.1) and as in the proof of Theorem 6.2, there are there possibilities: (i) if $q_r \notin k$ then $d(a) > d(b)$, (ii) if $q_r = 0$ then $d(a) < d(b)$, (iii) if $q_r = \alpha \in \mathbf{U}(R)$, then $d(b\alpha - a) < d(b)$. In terms of the notion of (right) R-dependence, which is defined in any ring with a degree function (cf. [4] or [5]), we can sum up the result as follows:

Proposition **(7.3)**. — *If* R *is a k-ring with a degree function which is also a* GE_2*-ring, then of any two elements of the same degree which form a regular row, each is* R*-dependent on the other.*

With the help of this result it is easy to show e.g. that the ring $k[x, y]$ of polynomials in two indeterminates over a field is not a GE_2-ring. As the degree function we take the total degree in x and y. Then $(1 + xy, x^2)$ is a regular row consisting of two elements of the same degree, neither of which is R-dependent on the other. This means that the matrix

$$\begin{pmatrix} 1 + xy & x^2 \\ -y^2 & 1 - xy \end{pmatrix}$$

which is clearly invertible, cannot be expressed as a product of elementary matrices. The same reasoning shows that more generally, the ring $k[x_1, \ldots, x_d]$ is not a GE_2-ring whenever $d > 1$.

8. Discretely ordered rings.

The results of §§ 5-7 show the usefulness of a norm for the study of $\mathbf{GE}_2(R)$. In some respects the same purpose is served by assuming that the ring is totally ordered [1]; here it is again necessary to make some discreteness assumption.

Definition. — *A* discretely ordered ring *is a ring* R *which is totally ordered, such that*

(8.1) *for any* $a \in R$, *if* $a > 0$, *then* $a \geq 1$.

[1] I am indebted to H. Bass for drawing my attention to this possibility.

In order to describe $\mathbf{GE}_2(R)$ in this case it is advantageous to supplement the diagonal matrices by matrices with positive coefficients and we therefore replace the generators $E(x)$ by

$$P(x)=\begin{pmatrix} x & 1 \\ 1 & 0 \end{pmatrix}. \tag{8.2}$$

Of course the P's could also have been used to develop the theory for discretely normed rings; but the E's there offered the advantage that their determinant was 1. If $\mathbf{P}_2(R)$ denotes the group generated by all the P's, then in particular, $[1,-1]=P(1)P(-1)P(1)\in\mathbf{P}_2(R)$, and since $[1,-1]$ normalizes $\mathbf{E}_2(R)$ and

$$P(x)=[1,-1]E(x), \qquad E(x)=[1,-1]P(x), \tag{8.3}$$

it follows that $\mathbf{E}_2(R)$ is a subgroup of index 2 in $\mathbf{P}_2(R)$. In order to describe $\mathbf{P}_2\cap\mathbf{D}_2$ we shall write

$$C(\alpha)=[\alpha,-\alpha^{-1}] \qquad (\alpha\in\mathbf{U}(R)). \tag{8.4}$$

With these notations we have the following analogues of (2.2-4):

$$P(x+y)=P(x)P(0)P(y), \tag{8.5}$$

$$P(\alpha)P(-\alpha^{-1})P(\alpha)=C(\alpha), \tag{8.6}$$

$$P(x)[\alpha,\beta]=[\beta,\alpha]P(\beta^{-1}x\alpha). \tag{8.7}$$

Now the relation (8.3) between $P(x)$ and $E(x)$ shows that the defining relations (8.5-7) are equivalent to the relations (2.2-4). Hence R is universal for $\mathbf{GE}_2$ if and only if $\mathbf{GE}_2$ has (8.5-7) and the relations in $\mathbf{D}_2$ as a complete set of defining relations. As in § 2 we obtain the following relations as a consequence of (8.5-7):

$$P(0)^2=I, \qquad [P(1)P(-1)]^3=-I, \tag{8.8}$$

$$P(x)^{-1}=P(0)P(-x)P(0)=\begin{pmatrix} 0 & 1 \\ 1 & -x \end{pmatrix}, \tag{8.9}$$

$$P(x)P(y)^{-1}=P(x-y)P(0), \tag{8.10}$$

$$P(x)P(y)^{-1}P(z)=P(x-y+z), \tag{8.11}$$

$$P(x)P(\alpha)P(y)=P(x+\alpha^{-1})C(\alpha)P(y+\alpha^{-1}). \tag{8.12}$$

In particular, putting $\alpha=1$ in the last relation, we find

$$P(x)P(1)P(y)=P(x+1)C(1)P(y+1)=C(-1)P(-x-1)P(y+1),$$

whence on replacing x, y by $x-1, y-1$ respectively, we find

$$P(-x)P(y)=C(-1)P(x-1)P(1)P(y-1). \tag{8.13}$$

Similarly,

$$P(x)P(-y)P(z)=-P(x-1)P(1)P(y-2)P(1)P(z-1). \tag{8.14}$$

Next we shall derive expressions for a product of P's, analogous to (5.2); they are essentially the continuant polynomials, cf. [18] and [5]. Let $t_1, t_2, \ldots$ be a sequence of noncommuting indeterminates and define the polynomials p_i recursively by

$$\textbf{(8.15)}\qquad \begin{cases} p_{-1}=0, & p_0=1, \\ p_n(t_1, \ldots, t_n)=p_{n-1}(t_1, \ldots, t_{n-1})t_n+p_{n-2}(t_1, \ldots, t_{n-2}). \end{cases}$$

As in the case of the e's defined in (5.1) we shall omit either the arguments or the suffixes from the p's. It is easily seen that

$$\textbf{(8.16)}\qquad \mathrm{P}(a_1)\ldots\mathrm{P}(a_r)=\begin{pmatrix} p(a_1, \ldots, a_r) & p(a_1, \ldots, a_{r-1}) \\ p(a_2, \ldots, a_r) & p(a_2, \ldots, a_{r-1}) \end{pmatrix}.$$

The symmetry of (8.16) leads to the following alternative definition for the p's:

$$\textbf{(8.17)}\qquad \begin{cases} p_{-1}=0, & p_0=1, \\ p_n(t_1, \ldots, t_n)=t_1p_{n-1}(t_2, \ldots, t_n)+p_{n-2}(t_3, \ldots, t_n). \end{cases}$$

Either definition shows that p_n may be described as the sum of $t_1t_2\ldots t_n$ and all terms obtained by omitting one or more pairs of adjacent factors t_it_{i+1}. In particular, this shows the truth of

Lemma **(8.1.)** — *In any ordered ring* R, *given* $a_1, \ldots, a_r \in \mathrm{R}$ *such that* $a_i>0$ *for* $1\leqslant i\leqslant r$ $(r>0)$, *then*

$$p(a_1, \ldots, a_r)>0.$$

The same conclusion holds if $a_1\geqslant 0$, *and* $a_i>0$ *for* $2\leqslant i\leqslant r$, *provided that* $r\geqslant 2$.

This lemma leads to the following analogue of Theorem 2.2 and Theorem 5.2, which is related to the uniqueness of the expansion of a rational number in a simple continued fraction (cf. [9], p. 135).

Theorem **(8.2.)** — *Let* R *be any discretely ordered ring. Then* R *is universal for* $\mathbf{GE}_2$; *moreover, any* $\mathrm{A}\in\mathbf{GE}_2(\mathrm{R})$ *is unique of the form*

$$\textbf{(8.18)}\qquad \mathrm{A}=[\alpha, \beta]\mathrm{P}(a_1)\ldots\mathrm{P}(a_r) \qquad a_i\in\mathrm{R},\ \alpha, \beta\in\mathbf{U}(\mathrm{R}),$$

subject to the conditions

$$\textbf{(8.19)}\qquad a_1\geqslant 0, \qquad a_i>0 \qquad (1<i<r),$$

$$\textbf{(8.20)}\qquad \textit{when } r=2,\ a_1, a_2 \textit{ are not both zero.}$$

Proof. — By definition, $\mathbf{GE}_2(\mathrm{R})$ is generated by the diagonal matrices and all $\mathrm{P}(a)$, $a\in\mathrm{R}$; using (8.9) and (8.7) we can bring any such product to the form (8.18). If $a_i=0$ for some i $(1<i<r)$ we can use (8.5) to simplify the expression (8.18); so we may assume that $a_i\neq 0$ $(1<i<r)$. To complete the proof we use induction on the number of sign changes in the sequence $(a_1, \ldots, a_{r-1}, 1)$. If this number is zero, it means that the conditions (8.19) are satisfied. In the contrary case let i be the last suffix for which $a_i<0$. Then $-a_i>0$, hence $-a_i-1\geqslant 0$ and

$$\begin{aligned}\mathrm{A}&=[\alpha, \beta]\mathrm{P}(a_1)\ldots\mathrm{P}(a_{i-1})\mathrm{C}(-1)\mathrm{P}(-a_i-1)\mathrm{P}(1)\mathrm{P}(a_{i+1}-1)\ldots\mathrm{P}(a_r)\\ &=\pm[\alpha, -\beta]\mathrm{P}(-a_1)\ldots\mathrm{P}(-a_{i-1})\mathrm{P}(-a_i-1)\mathrm{P}(1)\mathrm{P}(a_{i+1}-1)\ldots\mathrm{P}(a_r).\end{aligned}$$

If $a_i=-1$, we can reduce the number of factors P by (8.5).

If $a_{i+1} > 1$, then there is one less sign change than before and we can apply the induction hypothesis. If $a_{i+1} = 1$, we can combine $P(1)$ with $P(a_{i+2})$ by (8.5), unless $i+1=r$; in any case the number of sign changes is again diminished by 1. There remains the case $a_{i+1} < 1$; this can only happen when $a_{i+1} \leqslant 0$, hence $i+1=r$ and the same conclusion applies. By induction it follows that the conditions (8.19) can always be satisfied; moreover only (8.5-7) were used in the process. To prove that the form (8.18) is unique, let us first assume that A is diagonal. If $a_1 > 0$, then a comparison of (1,2)-elements shows that $p(a_1, \ldots, a_{r-1}) = 0$, which contradicts Lemma 8.1. If $a_1 = 0$, we have

$$P(a_2) \ldots P(a_r) = DP(0),$$

where D is a diagonal matrix, and comparing (2,2)-elements we find that

$$p(a_3, \ldots, a_{r-1}) = 0,$$

which again contradicts Lemma 8.1, unless $r \leqslant 2$. In the latter case we are only left with the possibility $a_1 = a_2 = 0$, which was excluded in (8.20). If $a_1 < 0$, then $r = 1$ and A cannot be diagonal. Thus (8.18) is unique when A is diagonal.

In general, let (8.18) hold and also

$$A = [\gamma, \delta] P(b_1) \ldots P(b_s) \qquad b_1 \geqslant 0, \; b_j > 0 \quad (1 < j < s).$$

If $a_r = b_s$ we can cancel a term and use induction on $\max(r, s)$, so we may assume that $a_r \neq b_s$, say $a_r > b_s$. Writing D_1, D_2, etc., for diagonal matrices whose exact value is immaterial, we find

$$\textbf{(8.21)} \qquad I = D_1 P(a_1) \ldots P(a_r - b_s) P(-b_{s-1}) \ldots P(-b_1) P(0),$$

hence

$$\begin{aligned} I &= D_2 P(-a_1) \ldots P(b_s - a_r) P(b_{s-1}) \ldots P(b_2) P(b_1 - 1) P(1) P(-1) \\ &= D_3 P(a_1) \ldots P(a_r - b_s - 1) P(1) P(b_{s-1} - 1) P(b_{s-2}) \ldots P(b_2) P(b_1 - 1) P(1) P(-1). \end{aligned}$$

Since $a_r > b_s$, we have $a_r - b_s - 1 \geqslant 0$; in case this is zero we can combine $P(a_{r-1})$ with $P(1)$. Then all arguments are strictly positive except possibly the first, last and $b_1 - 1$. Moreover, $a_1 \geqslant 0$, $b_1 \geqslant 0$. Assume for the moment that $b_1 > 0$, then $b_1 - 1 \geqslant 0$ and by combining $P(b_2)$ with $P(1)$ in case $b_1 - 1 = 0$, we obtain a formula in which all arguments except the first and last are > 0, and the first is $\geqslant 0$. The special case proved shows that this is only possible when there are just two arguments, both zero. Thus we have uniqueness in this case. If $b_1 = 0$, the argument is the same except that the last two factors in (8.21) are omitted. This completes the proof.

The most obvious example of a discretely ordered ring is the ring $\mathbf{Z}$ of rational integers. Unfortunately this cannot be extended to the ring of integers in a real algebraic number field, since such a ring is never discretely ordered. If R is any discretely ordered ring, then $R[x]$, the ring of polynomials in an indeterminate over R, is again discretely ordered, if we take as positive polynomials those polynomials which have a positive leading coefficient. E.g. $\mathbf{Z}[x]$ is discretely ordered in this way. We shall show below that this ring is not a GE_2-ring, but in order to do so we need a refinement of Lemma 8.1.

Lemma **(8.3)**. — *Let* R *be discretely ordered,* $a_1, \ldots, a_r \in \mathrm{R}$ $(r \geqslant 2)$, $a_1 \geqslant 0$, $a_i > 0$ $(1 < i < r)$*; then*

$$p(a_1, \ldots, a_r) > p(a_1, \ldots, a_{r-1}). \tag{8.22}$$

Proof. — Write $p_i = p(a_1, \ldots, a_i)$ and assume first that $a_1 \neq 0$; then by definition, $p_r = p_{r-1}a_r + p_{r-2}$, hence

$$p_r - p_{r-1} = p_{r-1}(a_r - 1) + p_{r-2}.$$

Now $a_r > 0$, hence $a_r \geqslant 1$ and $p_{r-1} > 0$, $p_{r-2} > 0$ by Lemma 8.1, hence $p_r > p_{r-1}$, i.e. (8.22) holds. If $a_1 = 0$, then by (8.17), $p_r = p(a_3, \ldots, a_r)$, $p_{r-1} = p(a_3, \ldots, a_{r-1})$ and the result still holds.

Now let $\mathrm{A} \in \mathbf{GE}_2(\mathrm{R})$, where R is any discretely ordered ring. Then by Theorem 8.2,

$$\mathrm{A} = [\alpha, \beta]\mathrm{P}(q_1) \ldots \mathrm{P}(q_r), \qquad q_1 \geqslant 0,\ q_i > 0 \quad (1 < i < r). \tag{8.23}$$

Let $\mathrm{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and consider

$$\mathrm{AP}(q_r)^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -q_r \end{pmatrix} = \begin{pmatrix} b & a - bq_r \\ d & c - dq_r \end{pmatrix}.$$

If we compare this expression with (8.23) we see by Lemma 8.1, that b and $a - bq_r$ must have the same sign; replacing A by $-\mathrm{A}$ if necessary, we may take this sign to be positive. Then $b > a - bq_r$; now there are three possibilities, according to the sign of q_r:

Lemma **(8.4)**. — *Let* R *be a discretely ordered ring, and let* $\mathrm{A} \in \mathbf{GE}_2(\mathrm{R})$ *be any matrix with first row* (a, b), *where* $b > 0$ *and assume that* A *has the form* (8.23) *with* $r \geqslant 2$. *Then*

(i) *if* $q_r > 0$, *then* $a > b > 0$,

(ii) *if* $q_r = 0$, *then* $b > a > 0$,

(iii) *if* $q_r = -c < 0$, *then* $b > a > -bc$.

Proof. — A comparison of (1,2)-elements in (8.23) shows that

$$b = \alpha p(a_1, \ldots, a_{r-1}),$$

hence $\alpha > 0$ and so α may be ignored. If $q_r > 0$, we can apply Lemma 8.3 to the first row of A. If $q_r = 0$ we apply Lemma 8.3 to $\mathrm{AP}(q_r)^{-1}$ unless $r \leqslant 3$, when a separate argument is necessary; this may be left to the reader. If $q_r < 0$, $q_r = -c$, we have $b > a + bc > 0$ and hence the result follows.

As an application, let us take any discretely ordered ring R and consider R$[x]$, with the ordering by highest coefficient described earlier. The matrix

$$\mathrm{A} = \begin{pmatrix} 1 + 2x & 4 \\ -x^2 & 1 - 2x \end{pmatrix}$$

is invertible and $1+2x>4>0$, so if $A\in\mathbf{GE}_2$, we are in case (i) of Lemma 8.4 (clearly $r\geqslant 2$ in any representation (8.23) of A). But then $q_r>0$ and $b>a-bq_r>0$, i.e.

$$a>bq_r>0.$$

Let u be the leading coefficient of q_r; then $u>0$, hence $u\geqslant 1$ and so $4u>0, 4u>2$; this shows that q_r must be of degree zero in x. But for such q_r we have $a-bq_r>b$, which is a contradiction. Hence A cannot be of the form (8.23) and this shows that $R[x]$ cannot be a GE_2-ring. In particular, $\mathbf{Z}[x_1, \ldots, x_d]$ is not a GE_2-ring, for any number (>0) of indeterminates. Similarly, free associative rings (i.e. algebras over $\mathbf{Z}$) are not GE_2-rings.

9. The commutator quotient structure of $GE_2(R)$ and of $E_2(R)$.

For any group G we shall denote the derived group by G′, the commutator quotient group G/G′ by G^a (i.e. G abelianized) and write $x\to x^a$ for the natural homomorphism $G\to G^a$. We saw in § 2 that for any ring R, $\mathbf{E}_2(R)$ is normal in $\mathbf{GE}_2(R)$. When R is universal for GE_2, this quotient is abelian; we therefore begin by determining the quotient in this case.

Theorem (**9.1**). — *Let* R *be any ring universal for* $\mathbf{GE}_2$. *Then*

$$\mathbf{GE}_2(R)/\mathbf{E}_2(R)\cong\mathbf{U}(R)^a. \tag{9.1}$$

Proof. — We define a homomorphism $f:\mathbf{GE}_2(R)\to\mathbf{U}(R)^a$ by the rule

$$E(x)\to 1, \qquad [\alpha,\beta]\to(\alpha\beta)^a.$$

To show that this is well-defined we need only check that the defining relations (2.2-4) are preserved. But this is immediate for (2.2) and (2.3), and for (2.4) it follows from the fact that $(\alpha\beta)^a=(\beta\alpha)^a$. Clearly f maps $\mathbf{GE}_2(R)$ onto $\mathbf{U}(R)^a$, so to complete the proof we need only show that $\ker f=\mathbf{E}_2(R)$. By definition, $E(x)\in\ker f$, so the kernel contains $\mathbf{E}_2(R)$. Conversely, if A, given by (2.11) say, lies in the kernel, then $\alpha\beta\in\mathbf{U}(R)'$; multiplying by $D(\beta)=E(0)^2E(\beta)E(\beta^{-1})E(\beta)$, we may suppose that $\beta=1$, so it only remains to show that $[\gamma, 1]\in\mathbf{E}_2(R)$ for all $\gamma\in\mathbf{U}(R)'$, and this will follow if we prove that $[\alpha^{-1}\beta^{-1}\alpha\beta, 1]\in\mathbf{E}_2(R)$ for all $\alpha,\beta\in\mathbf{U}(R)$. Now we saw that $D(\alpha)\in\mathbf{E}_2(R)$ and

$$[\alpha^{-1}\beta^{-1}\alpha\beta, 1]=D(\alpha^{-1})D(\beta^{-1})D(\alpha\beta)\in\mathbf{E}_2(R). \tag{9.2}$$

Hence $\ker f=\mathbf{E}_2(R)$ and (9.1) follows.

Corollary 1. — *In any ring which is universal for* $\mathbf{GE}_2$, $\mathbf{D}_2(R)\cap\mathbf{E}_2(R)$ *is generated by all matrices* $D(\alpha)$, $\alpha\in\mathbf{U}(R)$.

For if we restrict the homomorphism $f:\mathbf{GE}_2(R)\to\mathbf{U}(R)^a$ constructed in the proof of the theorem to $\mathbf{D}_2(R)$, the kernel is just $\mathbf{D}_2(R)\cap\mathbf{E}_2(R)$. Now $[\alpha,\beta]$ maps to 1 if and only if $\alpha\beta\in\mathbf{U}(R)'$, but in this case the proof of the theorem shows that $[\alpha,\beta]$ can be written as a product of matrices $D(\gamma)$, $\gamma\in\mathbf{U}(R)$.

Cor. 1 shows that any element $A \in \mathbf{E}_2(R)$ can be written

$$A = D(\alpha_1)D(\alpha_2)\dots D(\alpha_k)E(a_1)\dots E(a_r).$$

In particular, any relation in $\mathbf{E}_2(R)$ can be brought to this form, using only (2.2), (2.3) and

$$(9.3) \qquad E(x)D(\alpha) = D(\alpha^{-1})E(\alpha x \alpha),$$

which is just the special case of (2.4) where $\alpha\beta = 1$. If R is assumed to be quasi-free for $\mathbf{GE}_2$, then by the remark just made, we can further reduce this relation to the form

$$(9.4) \qquad D(\alpha_1)\dots D(\alpha_k) = I.$$

Using (9.2) we can finally reduce (9.4) to the form

$$(9.5) \qquad [\gamma_1, 1]\dots[\gamma_l, 1] = I \quad \text{where} \quad \gamma_j \in \mathbf{U}(R)', \ \gamma_1\dots\gamma_l = 1.$$

Thus we obtain

Corollary 2. — In a ring R *which is quasi-free for* $\mathbf{GE}_2$, $\mathbf{E}_2(R)$ *is generated by all* $E(x)$ $(x \in R)$ *and if* $D(\alpha)$ $(\alpha \in \mathbf{U}(R))$ *and* $[\gamma, 1]$ $(\gamma \in \mathbf{U}(R)')$ *are defined by* (2.3) *and* (9.2) *respectively, then a complete set of defining relations in terms of these generators is given by*

$$E(x)E(0)E(y) = -E(x+y),$$
$$E(x)D(\alpha) = D(\alpha^{-1})E(\alpha x \alpha),$$
$$[\gamma_1, 1]\dots[\gamma_l, 1] = I \qquad (\gamma_j \in \mathbf{U}(R)', \ \gamma_1\dots\gamma_l = 1).$$

In the case of universal GE_2-rings, Theorem 9.1 may be thought of as a generalization (in the case $n = 2$) of Dieudonné's determinants over a skew field (cf. [7]). Since the right-hand side of (9.1) is abelian, we obtain

Corollary 3. — For any ring which is universal for $\mathbf{GE}_2$,

$$(9.6) \qquad \mathbf{E}_2(R) \supseteq \mathbf{GE}_2(R)'.$$

This corollary can still be improved, as follows:

Proposition (**9.2**). *— For any ring which is universal for* $\mathbf{GE}_2$ *and in which 1 can be written as the sum of two units,*

$$(9.7) \qquad \mathbf{GE}_2(R)' = \mathbf{E}_2(R),$$

and hence $\mathbf{GE}_2(R)^a \cong \mathbf{U}(R)^a$.

The result follows, once (9.7) is proved, and this is well known (cf. e.g. [2]). It is proved by noting that if $\alpha + \beta = 1$, the commutator $X^{-1}Y^{-1}XY$ of $X = [\alpha^{-1}, 1]$ and $Y = E(\beta^{-1}x)E(0)^{-1}$ is $E(x)E(0)^{-1}$; by transposition we obtain $E(0)^{-1}E(x)$ and hence $\mathbf{GE}_2(R)'$ contains

$$-E(0)E(-1)E(-1)E(0)E(0)E(x-1) = E(0)E(0)D(-1)E(x) = E(x).$$

In particular, when $\mathbf{U}_0(R)$ is a field k, we have the

Corollary. — In any k-ring with a degree function, (9.7) *holds provided that k has more than two elements.*

It is convenient to consider next the group $\mathbf{E}_2(\mathrm{R})^a$. Let R be any discretely normed ring; working mod $\mathbf{E}_2(\mathrm{R})'$ we have by (2.5) and (2.2),

$$\begin{aligned}\mathrm{E}(x)\mathrm{E}(y) &\equiv \mathrm{E}(x)\mathrm{E}(0)\mathrm{E}(y)\mathrm{E}(0)\mathrm{E}(-1)\mathrm{E}(0)\mathrm{E}(-1)\mathrm{E}(0)\mathrm{E}(-1)\\ &\equiv \mathrm{E}(x+y-3),\end{aligned}$$

hence

$$\mathrm{E}(x)\mathrm{E}(y) \equiv \mathrm{E}(x+y-3) \qquad (\mathrm{mod}\ \mathbf{E}_2(\mathrm{R})'). \tag{9.8}$$

This suggests defining a mapping $\mathbf{E}_2(\mathrm{R}) \to \mathrm{R}$ by putting

$$\mathrm{E}(x) \to x-3. \tag{9.9}$$

This mapping can be extended to a homomorphism of $\mathbf{E}_2(\mathrm{R})$ into the additive group of R provided that the defining relations of $\mathbf{E}_2(\mathrm{R})$ are preserved. Since $\mathrm{E}(0)^2=-\mathrm{I}$, we must have $-\mathrm{I}\to-6$ and in particular,

$$12=0 \qquad \text{in R.} \tag{9.10}$$

Next the definition of $\mathrm{D}(\alpha)$ shows that

$$\mathrm{D}(\alpha) \to 2\alpha+\alpha^{-1}-3, \tag{9.11}$$

and using this value in (9.3) we find that $x-3+2\alpha+\alpha^{-1}-3=2\alpha^{-1}+\alpha-3+\alpha x\alpha-3$, i.e. $\alpha x\alpha - x = \alpha-\alpha^{-1}$. If we replace x by $x+\alpha^{-1}$, this equation reduces to

$$\alpha x \alpha = x. \tag{9.12}$$

Thus in order to be able to define a homomorphism into the additive group of R we must have (9.10) and (9.12). However we can obtain a homomorphism in all cases if we divide out by the appropriate subgroup of R. In order not to complicate the result, let us assume that $\mathbf{U}(\mathrm{R})$ is abelian. Then we have

Theorem **(9.3)**. — *Let* R *be a ring which is quasi-free for* $\mathbf{GE}_2$ *and such that* $\mathbf{U}(\mathrm{R})$ *is abelian, and denote by* M *the additive subgroup of* R *generated by 12, all* $\alpha x\alpha - x$ $(x\in\mathrm{R},\ \alpha\in\mathbf{U}(\mathrm{R}))$ *and all* $3(\alpha+1)(\beta+1)$ $(\alpha,\beta\in\mathbf{U}(\mathrm{R}))$; *then*

$$\mathbf{E}_2(\mathrm{R})^a \cong \mathrm{R}/\mathrm{M}, \tag{9.13}$$

under the homomorphism defined by

$$\mathrm{E}(x) \to x-3 \qquad (\mathrm{mod}\ \mathrm{M}). \tag{9.14}$$

It should be noted that M is in general not an ideal of R but merely an additive subgroup.

Proof. — It is clear from the remarks preceding the theorem and the definition of M that the first two relations of Theorem 9.1, Cor. 2 are preserved. Moreover, the third relation reduces in the present case to

$$\mathrm{D}(\alpha^{-1})\mathrm{D}(\beta^{-1})\mathrm{D}(\alpha\beta)=\mathrm{I}. \tag{9.15}$$

By definition of M we have $\alpha x\alpha\equiv x$, hence $\alpha\equiv\alpha^{-1}$ (mod M). Thus

$$\mathrm{D}(\alpha)\rightarrow 2\alpha+\alpha^{-1}-3\equiv 3(\alpha-1)\equiv 3(\alpha^{-1}-1) \qquad (\text{mod } \mathrm{M}),$$

and the left-hand side of (9.15) is mapped to

$$\begin{aligned} 3(\alpha-1)+3(\beta-1)+3(\alpha\beta-1) &\equiv 3(\alpha+\beta+\alpha\beta-3)\\ &\equiv 3(\alpha+\beta+\alpha\beta+1)\\ &\equiv 3(\alpha+1)(\beta+1)\\ &\equiv 0 \qquad (\text{mod } \mathrm{M}). \end{aligned}$$

Thus all defining relations of $\mathbf{E}_2(\mathrm{R})$ are preserved and (9.14) defines indeed a homomorphism $g:\mathbf{E}_2(\mathrm{R})\rightarrow\mathrm{R}/\mathrm{M}$. Clearly g is onto, and since the right-hand side is abelian, ker $g\supseteq\mathbf{E}_2(\mathrm{R})'$; to complete the proof we must establish equality here. By (9.8) we have

$$\mathrm{E}(x+3)\mathrm{E}(y+3)\equiv\mathrm{E}(x+y+3) \qquad (\text{mod } \mathbf{E}_2(\mathrm{R})'), \tag{9.16}$$

and hence

$$\mathrm{E}(x_1+3)\ldots\mathrm{E}(x_n+3)\equiv\mathrm{E}(\Sigma x_i+3) \qquad (\text{mod } \mathbf{E}_2(\mathrm{R})'). \tag{9.17}$$

If the left-hand side maps to zero, so must the right-hand side, whence $\Sigma x_i\equiv 0$ (mod M). Now we have the identity

$$\mathrm{E}(x)\mathrm{E}(0)^{-1}\mathrm{E}(x)^{-1}\mathrm{E}(0)=\mathrm{E}(x)\mathrm{E}(0)^{-2}\mathrm{E}(-x)=-\mathrm{E}(x)\mathrm{E}(-x);$$

hence

$$-\mathrm{E}(x)\mathrm{E}(-x)\equiv\mathrm{I} \qquad (\text{mod } \mathbf{E}_2(\mathrm{R})').$$

It follows that

$$\begin{aligned} \mathrm{D}(\alpha^2)=-\mathrm{D}(\alpha)\mathrm{D}(-\alpha)&=\mathrm{E}(\alpha)\mathrm{E}(\alpha^{-1})\mathrm{E}(\alpha).\mathrm{E}(-\alpha)\mathrm{E}(-\alpha^{-1})\mathrm{E}(-\alpha)\\ &\equiv\mathrm{I} \qquad (\text{mod } \mathbf{E}_2(\mathrm{R})'). \end{aligned}$$

Next we have

$$\mathrm{E}(x)\mathrm{D}(\alpha)=\mathrm{D}(\alpha^{-1})\mathrm{E}(\alpha x\alpha), \qquad \text{hence}$$

$$\mathrm{E}(\alpha x\alpha)\mathrm{E}(x)^{-1}\equiv\mathrm{D}(\alpha^2)\equiv\mathrm{I} \qquad (\text{mod } \mathbf{E}_2(\mathrm{R})'). \tag{9.18}$$

If we put $x=-3, y=0$ in (9.16), we see that $\mathrm{E}(3)\in\mathbf{E}_2(\mathrm{R})'$, and so, by (2.8), we obtain

$$\mathrm{E}(\alpha x\alpha-x+3)=\mathrm{E}(\alpha x\alpha)\mathrm{E}(x)^{-1}\mathrm{E}(3)\in\mathbf{E}_2(\mathrm{R})'. \tag{9.19}$$

By (9.18) we also have $\mathrm{E}(\alpha x\alpha)\equiv\mathrm{E}(x)$, hence $\mathrm{E}(\alpha)\equiv\mathrm{E}(\alpha^{-1})$ (mod $\mathbf{E}_2(\mathrm{R})'$), and so $\mathrm{E}(3\alpha)\equiv\mathrm{E}(\alpha)\mathrm{E}(0)\mathrm{E}(\alpha)\mathrm{E}(0)\mathrm{E}(\alpha)\equiv\mathrm{E}(0)^2\mathrm{E}(\alpha)\mathrm{E}(\alpha^{-1})\mathrm{E}(\alpha)\equiv\mathrm{D}(\alpha)$, hence

$$\mathrm{E}(3\alpha)\equiv\mathrm{D}(\alpha) \qquad (\text{mod } \mathbf{E}_2(\mathrm{R})'). \tag{9.20}$$

In particular, $\mathrm{E}(-6+3)\equiv\mathrm{E}(-3)\equiv-\mathrm{I}$, whence

$$\mathrm{E}(12+3)\equiv\mathrm{I} \qquad (\text{mod } \mathbf{E}_2(\mathrm{R})'). \tag{9.21}$$

Further, $\mathrm{E}(3\alpha\beta)\equiv\mathrm{D}(\alpha\beta)\equiv\mathrm{D}(\alpha)\mathrm{D}(\beta)\equiv\mathrm{E}(3\alpha)\mathrm{E}(3\beta)$ and so

$$\mathrm{E}(3(\alpha\beta+\alpha+\beta+1)+3)\equiv\mathrm{E}(3\alpha\beta)\mathrm{E}(3\alpha)\mathrm{E}(3\beta)\mathrm{E}(15)\equiv\mathrm{D}(\alpha\beta)\mathrm{D}(\alpha)\mathrm{D}(\beta)\equiv\mathrm{I},$$

by (9.18) and (9.21), hence

(9.22) $$E(3(\alpha+1)(\beta+1)+3)\equiv I \qquad (\operatorname{mod} \mathbf{E}_2(R)').$$

Now (9.19), (9.21) and (9.22) show that for any generator x of M, $E(x+3)\in\mathbf{E}_2(R)'$. By (9.17) this holds for any $x\in M$. Hence, in particular, if the left-hand side of (9.17) is mapped to zero by g, then the right-hand side lies in $\mathbf{E}_2(R)'$ and therefore so does the left-hand side. This shows that ker $g\subseteq\mathbf{E}_2(R)'$ and it completes the proof of Theorem 9.3.

The isomorphism (9.13) can be made to play a similar role for $\mathbf{E}_2(R)$ as the determinant for $\mathbf{GL}_2(R)$; in this connexion we note that when R is commutative, so that a determinant is defined, then (in case R is discretely normed or discretely ordered) $\mathbf{E}_2(R)$ is just the subgroup of $\mathbf{GE}_2(R)$ consisting of matrices of determinant 1, because every matrix in $\mathbf{GE}_2(R)$ is congruent $(\operatorname{mod}\mathbf{E}_2(R))$ to a matrix of the form $[\alpha, 1]$. In particular, for a commutative (discretely normed) GE_2-ring R we have $\mathbf{E}_2(R)=\mathbf{SL}_2(R)$. As an example, take the ring $\mathbf{Z}$ of rational integers. Here $M=12\mathbf{Z}$ and Theorem 9.3 shows that $\mathbf{SL}_2(\mathbf{Z})^a$ is cyclic of order 12. Secondly, let I be the ring of Gaussian integers. The only units are ± 1, $\pm i$ and in this case M is generated by 12 and all $2x$ and all $3(\alpha+1)(\beta+1)$. This is just the ideal $2R$, so $\mathbf{SL}_2(I)^a$ is the direct product of two cycles of order two.

Suppose that R is a k-ring with a degree-function and moreover, that k is in the centre of R, i.e. that R is a *k-algebra*. Then M contains $x(\alpha^2-1)$ for all $x\in R$, $\alpha\in k$ $(\alpha\neq 0)$, and hence coincides with R, unless every non-zero element of k has its square equal to 1. In the latter case k cannot have more than 3 elements and then it is easily verified that $M=0$. Thus we have

Corollary 1. — Let R *be a k-algebra with a degree-function, then* $\mathbf{E}_2(R)^a=0$ *unless k is the field of 2 or 3 elements, in which case* $\mathbf{E}_2(R)^a\cong R$.

At the other extreme R has no units apart from ± 1; then it is easily seen that M is the additive subgroup generated by 12. This is true in particular when R is discretely ordered, and so we have

Corollary 2. — Let R *be a discretely ordered ring (or a discretely normed ring in which* ± 1 *are the only units), then*

$$\mathbf{E}_2(R)^a\cong R/12.$$

For example, if I is the ring of integers in an imaginary quadratic extension of $\mathbf{Q}$ which is not Euclidean, then $\mathbf{E}_2(I)^a$ is the direct product of an infinite cycle and a cycle of order 12.

To obtain more precise information about $\mathbf{GE}_2(R)^a$ it is necessary to find out more about $\mathbf{E}_2(R)/\mathbf{GE}_2(R)'$ in the cases where Prop. 9.2 does not apply. Let us return to the proof of Theorem 9.3 but drop the assumption that $\mathbf{U}(R)$ is abelian. By (2.8),

$$\begin{aligned}E(\alpha(x-3))E(x-3)^{-1}E(3\alpha)&=E(\alpha(x-3)-(x-3)+3\alpha)\\&=E((\alpha-1)x+3).\end{aligned}$$

Now (9.20) shows that $E(3\alpha)\equiv D(\alpha)$ (mod $\mathbf{GE}_2(R)'$); the proof uses no units apart from α and so does not depend on the commutativity of $\mathbf{U}(R)$. Since we clearly also have

$$\begin{aligned} I &= E(x)^{-1}[\alpha, 1]^{-1}E(x)[\alpha, 1] \\ &\equiv E(x)^{-1}E(\alpha x)D(\alpha) \qquad (\text{mod } \mathbf{GE}_2(R)'), \end{aligned}$$

we deduce that

$$\textbf{(9.23)} \qquad E((\alpha-1)x+3)\equiv I \qquad (\text{mod } \mathbf{GE}_2(R)').$$

By symmetry we also have

$$\textbf{(9.24)} \qquad E(x(\alpha-1)+3)\equiv I \qquad (\text{mod } \mathbf{GE}_2(R)').$$

Let N be the ideal of R generated by all $\alpha-1$ $(\alpha\in\mathbf{U}(R))$. We note that $N\supseteq M$ since e. g., $\alpha x\alpha - x=(\alpha-1)x\alpha+x(\alpha-1)$ and $12=-6(-1-1)$. It is easily seen that the mapping $h: E(x)\rightarrow x-1$ (mod N) preseives the first two relations of Theorem 9.1, Cor. 2, while the third one becomes trivial (mod $\mathbf{GE}_2(R)'$); therefore we have a homomorphism

$$\mathbf{E}_2(R)/\mathbf{GE}_2(R)'\rightarrow R/N.$$

If the left-hand side of (9.17) lies in the kernel of this mapping it follows that $\Sigma x_i\in N$ and now (9.23), (9.24) show that the right-hand side of (9.17) must lie in $\mathbf{GE}_2(R)'$. This proves that the kernel is in fact equal to $\mathbf{GE}_2(R)'$ and we have

Theorem **(9.4)**. — *Let* R *be any ring which is quasi-free for* $\mathbf{GE}_2$ *and denote by* N *the ideal generated by all* $\alpha-1$ $(\alpha\in\mathbf{U}(R))$. *Then*

$$\mathbf{E}_2(R)/\mathbf{GE}_2(R)'\cong R/N.$$

E.g., when $R=\mathbf{Z}$, then $N=2\mathbf{Z}$ and we obtain the result of Hua and Reiner [11], that $\mathbf{SL}_2(Z)/\mathbf{GL}_2(Z)'$ is cyclic of order two. If I is the ring of Gaussian integers, $N=(1+i)I$ and $\mathbf{SL}_2(I)/\mathbf{GL}_2(I)'$ again is cyclic of order two.

Theorems 9.1 and 9.4 may be summarized as follows:

Theorem **(9.5)**. — *Let* R *be a ring which is quasi-free for* GE_2 *and denote by* N *the ideal generated by all* $\alpha-1$ $(\alpha\in\mathbf{U}(R))$. *Then there is a split exact sequence*

$$0\rightarrow R/N\rightarrow \mathbf{GE}_2(R)^a\rightarrow \mathbf{U}(R)^a\rightarrow 0.$$

For the mapping $\alpha^a\rightarrow[\alpha^a, 1]$ clearly induces a splitting.

10. Generating sets and free products in $\mathbf{GE}_2(R)$.

It is well known that $\mathbf{SL}_2(\mathbf{Z})$, the two-dimensional unimodular group, can modulo its centre be written as a free product of a 2-cycle and a 3-cycle. It seems to be unusual for $\mathbf{GE}_2(R)$ or $\mathbf{E}_2(R)$ to be expressible as a free product in a non-trivial way, but there are a number of results on large subgroups which have the form of free products (cf. e.g. Nagao [14]). The question is related to the problem of deciding when $\mathbf{GE}_2(R)$ is finitely generated; by means of the normal form for elements in $\mathbf{GE}_2(R)$ we shall be able to give an answer to these questions which includes many of the known results as special cases. We begin by discussing the problem of finite generation of $\mathbf{GE}_2(R)$.

Theorem (**10.1**). — *Let* R *be any ring; if* $\mathbf{U}(R)$ *is finitely generated (as multiplicative group) and* R *is finitely generated, as* $\mathbf{U}(R)$*-bimodule, then* $\mathbf{GE}_2(R)$ *is finitely generated.*

Proof. — Let $\mathbf{U}(R)$ be generated by $\gamma_1, \ldots, \gamma_h$ and R, as $\mathbf{U}(R)$-bimodule, by $c_1, \ldots, c_k$. We already know that $\mathbf{GE}_2(R)$ is generated by all $E(a)$, $a \in R$, and $[\alpha, \beta]$, $\alpha, \beta \in \mathbf{U}(R)$. By (2.2) and (2.4), $E(a)$ can be expressed in terms of $E(c_r)$, $E(0)$ and $[\alpha, \beta]$, and hence in terms of

$$(\mathbf{10.1}) \qquad E(0),\ E(c_r) \quad (r=1, \ldots, k), \qquad [\gamma_i, \gamma_j] \quad (i,j=1, \ldots, h).$$

This shows that $\mathbf{GE}_2(R)$ is finitely generated.

The following is a partial converse:

Theorem (**10.2**). — *Let* R *be a ring quasi-free for* $\mathbf{GE}_2$ *and such that* $\mathbf{GE}_2(R)$ *is finitely generated. Then so is* $\mathbf{U}(R)^a$, *and moreover,* R *is finitely generated, as* $\mathbf{U}(R)$*-bimodule.*

Proof. — Since $\mathbf{GE}_2(R)$ is finitely generated, so is $\mathbf{U}(R)^a$, by Theorem 9.5. Moreover, every generating set of $\mathbf{GE}_2(R)$ contains a finite subset which is still a generating set. We can therefore find elements $c_1, \ldots, c_k \in R$ and $\gamma_1, \ldots, \gamma_h \in \mathbf{U}(R)$ such that the elements (10.1) generate $\mathbf{GE}_2(R)$. Thus every element of $\mathbf{GE}_2(R)$ can be expressed in the form

$$(\mathbf{10.2}) \qquad [\alpha, \beta]E(a_1) \ldots E(a_r),$$

where $a_1, \ldots, a_r$ belong to the $\mathbf{U}(R)$-bimodule generated by $c_1, \ldots, c_k$. Clearly we may assume that $a_i \neq 0$ $(i=1, \ldots, r)$ and using (2.9) we may even take $a_i \notin \mathbf{U}_0(R)$, provided that we let the a_i range over the $\mathbf{U}(R)$-bimodule generated by $1, c_1, \ldots, c_k$. Now take any $b \in R$ and consider $E(b)$. If $E(b)$ is expressed in the form (10.2), we necessarily have $r \geqslant 1$, hence we obtain

$$-[\alpha, \beta]E(a_1) \ldots E(a_{r-1})E(a_r - b)E(0) = I.$$

If $a_r - b \notin \mathbf{U}_0(R)$, then the argument used to prove Theorem 5.2 shows that $r \geqslant 2$ and we obtain a contradiction. Therefore $a_r - b \in \mathbf{U}_0(R)$ and it follows that b lies in the $\mathbf{U}(R)$-bimodule generated by $1, c_1, \ldots, c_k$; but b was any element of R. This shows that R is finitely generated, as $\mathbf{U}(R)$-bimodule.

This theorem shows e.g. that $\mathbf{GL}_2(k[x])$, for any field k, is not finitely generated (cf. Nagao [14]). It seems likely that Theorem 10.2 holds for any ring which is universal for $\mathbf{GE}_2$.

Next we turn to the question of free subgroups and free products. Let R be any ring, and A a subgroup of the additive group of R. Then it is clear that the set

$$\mathbf{B}_{12}(A) = \{B_{12}(a) \mid a \in A\}$$

is a subgroup of $\mathbf{E}_2(\mathrm{R})$ isomorphic to A, and of course the same holds for its conjugate with respect to P(o), which will be denoted by $\mathbf{B}_{21}(\mathrm{A})$.

Theorem **(10.3)**. — *Let* R *be a ring which is quasi-free for* $\mathbf{GE}_2$ *and* A, B *any two subgroups of the additive group of* R *which do not meet* $\mathbf{U}(\mathrm{R})$. *Then the subgroup of* $\mathbf{GE}_2(\mathrm{R})$ *generated by* $\mathbf{B}_{12}(\mathrm{A})$ *and* $\mathbf{B}_{21}(\mathrm{B})$ *is equal to their free product.*

Proof. — Let F be the subgroup generated by $\mathbf{B}_{12}(\mathrm{A})$ and $\mathbf{B}_{21}(\mathrm{B})$. Any element of F has the form $f=g_1h_1g_2h_2\dots g_rh_r$, where $g_i\in\mathbf{B}_{12}(\mathrm{A})$, $h_i\in\mathbf{B}_{21}(\mathrm{B})$ and $g_i\neq 1$ for $i\neq 1$, $h_i\neq 1$ for $i\neq r$. Write

$$g_i=\mathrm{B}_{12}(a_i)=\mathrm{E}(a_i)\mathrm{E}(0)^{-1}, \qquad h_i=\mathrm{B}_{21}(b_i)=\mathrm{E}(0)^{-1}\mathrm{E}(b_i),$$

then f has the form

$$\pm\mathrm{E}(a_1)\mathrm{E}(b_1)\dots\mathrm{E}(a_r)\mathrm{E}(b_r), \tag{10.3}$$

where $a_i, b_i\neq 0$ except possibly a_1 or b_r; moreover, $a_i, b_i\notin\mathbf{U}(\mathrm{R})$, by hypothesis. In particular, the left-hand side of any relation in F can be brought to the form (10.3). But by hypothesis there are no non-trivial relations of the form (10.3) and it follows that F is a free product as asserted.

For example, if x, y are elements of R such that $nx, ny\notin\mathbf{U}_0(\mathrm{R})$ for all $n\neq 0$, then the subgroup generated by $\mathrm{B}_{12}(x)$ and $\mathrm{B}_{21}(y)$ is free on these generators. In particular this yields the well known result that for any indeterminate x over a field of characteristic zero, $\mathrm{B}_{12}(x)$ and $\mathrm{B}_{21}(x)$ generate a free group. Similarly, for any integer $m\geqslant 2$, $\mathrm{B}_{12}(m)$ and $\mathrm{B}_{21}(m)$ generate a free group.

11. The construction of homomorphisms between general linear groups.

Let $f:\mathrm{R}\to\mathrm{S}$ be any homomorphism [1] of rings; clearly this induces a homomorphism of $n\times n$ matrix rings $f_n:\mathrm{R}_n\to\mathrm{S}_n$ and since a unit of R_n maps to a unit of S_n, we obtain a homomorphism

$$f^*:\mathbf{GL}_n(\mathrm{R})\to\mathbf{GL}_n(\mathrm{S}). \tag{11.1}$$

Secondly, let $g:\mathrm{R}\to\mathrm{S}$ be an antihomomorphism; this induces in the same way an antihomomorphism of groups $g^*:\mathbf{GL}_n(\mathrm{R})\to\mathbf{GL}_n(\mathrm{S})$. If we follow this by inversion in $\mathbf{GL}_n(\mathrm{S})$ we again obtain a homomorphism of groups

$$\check{g}:\mathbf{GL}_n(\mathrm{R})\to\mathbf{GL}_n(\mathrm{S}). \tag{11.2}$$

In particular, when $\mathrm{S}=\mathrm{R}$, the automorphisms and antiautomorphisms of R give rise in this way to automorphisms of $\mathbf{GL}_n(\mathrm{R})$. Not all automorphisms arise in this way, however; if σ is any homomorphism of $\mathbf{GL}_n(\mathrm{R})$ into the group of central units of R, then the mapping $\mathrm{A}\to\mathrm{A}.\mathrm{A}^\sigma$ is an endomorphism of $\mathbf{GL}_n(\mathrm{R})$, which will be called

[1] It is understood that such a homomorphism maps the 1 of R to the 1 of S.

a *central homothety.* When we come to study automorphisms of $\mathbf{GL}_n(\mathrm{R})$ we shall find that for $n \geqslant 3$ and a fairly wide class of rings, every automorphism of $\mathbf{GL}_n(\mathrm{R})$ is obtained by combining f^* or $\check{g}$ with an inner automorphism and a central homothety. For $n=2$ the situation is rather different; to describe it we need a

Definition. — *Let* R, S *be any rings. A* U-homomorphism $f:\mathrm{R}\to\mathrm{S}$ *is a homomorphism* $x\to x'$ *of the additive group of* R *into the additive group of* S *such that*

$$1'=1, \tag{11.3}$$

and

$$(\alpha a\beta)'=\alpha' a'\beta' \qquad \textit{for all } a\in\mathrm{R},\ \alpha,\ \beta\in\mathbf{U}(\mathrm{R}). \tag{11.4}$$

If g is a homomorphism from the additive group of R to that of S which satisfies (11.3) and instead of (11.4) satisfies

$$(\alpha a\beta)'=\beta' a'\alpha' \qquad \text{for all } a\in\mathrm{R},\ \alpha,\ \beta\in\mathbf{U}(\mathrm{R}), \tag{11.5}$$

then g is called a U-*antihomomorphism.*

Clearly a U-homomorphism between fields is just an ordinary homomorphism, and likewise for antihomomorphisms. More generally, we have

Proposition (**11.1**). — *Let* R *be a ring which is generated, qua ring, by its units. Then a* U-*homomorphism of* R *into an arbitrary ring is a homomorphism and a* U-*antihomomorphism is an antihomomorphism.*

Proof. — Let $x\to x'$ be a U-homomorphism and write

$$\mathrm{R}_0=\{x\in\mathrm{R}\mid (xa)'=x'a' \text{ for all } a\in\mathrm{R}\}$$

then by (11.4), R_0 contains $\mathbf{U}(\mathrm{R})$ and if $x, y\in\mathrm{R}_0$ then for any $a\in\mathrm{R}$,

$$\begin{aligned}[(x-y)a]'&=(xa-ya)'=(xa)'-(ya)'=x'a'-y'a'\\ &=(x'-y')a'=(x-y)'a',\end{aligned}$$

hence $x-y\in\mathrm{R}_0$. Further, $(xy)'=x'y'$ and for any $a\in\mathrm{R}$,

$$(xya)'=x'(ya)'=x'y'a'=(xy)'a',$$

which shows that $xy\in\mathrm{R}_0$. Together with (11.3) this shows that R_0 is a subring of R and since it contains the units of R, by (11.4), it must be the whole of R. But this means that $f: x\to x'$ is a homomorphism. The proof for antihomomorphisms is similar.

The form of the defining relations (2.2-4) now yields almost immediately

Theorem (**11.2**). — *Let* R *be any ring universal for* $\mathbf{GE}_2$, S *any ring and let* $f:\mathrm{R}\to\mathrm{S}$ *be any* U-*homomorphism* $x\to x'$. *Then* f *induces a homomorphism*

$$f^*:\mathbf{GE}_2(\mathrm{R})\to\mathbf{GE}_2(\mathrm{S}) \tag{11.6}$$

by the rule

$$\mathrm{E}(x)\to\mathrm{E}(x'), \qquad [\alpha,\beta]\to[\alpha',\beta']. \tag{11.7}$$

If $g : \mathrm{R} \to \mathrm{S}$ *is a* U-*antihomomorphism, then* g *induces an antihomomorphism* g^* *defined in the same way and hence induces a homomorphism*

$$(\mathbf{11.8}) \qquad \check{g} : \mathbf{GE}_2(\mathrm{R}) \to \mathbf{GE}_2(\mathrm{S})$$

by the rule

$$(\mathbf{11.9}) \qquad \mathrm{E}(x) \to \mathrm{E}(x')^{-1}, \qquad [\alpha, \beta] \to [\alpha', \beta']^{-1}.$$

Proof. — We need only check that the defining relations (2.2-4) and the relations in $\mathbf{D}_2(\mathrm{R})$ are satisfied and this is clearly the case.

When f is a homomorphism, the effect of f^* in (11.6) can be described more simply by the rule

$$(\mathbf{11.10}) \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \to \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}$$

but it should be noted that for arbitrary U-homomorphisms this is not so. Although there is still a group homomorphism f^* (given by (11.7)) it does not agree with the mapping (11.10), which in general will not be a homomorphism. A similar remark applies to U-antihomomorphisms.

Corollary. — *Any* U-*homomorphism of a universal* $\mathbf{GE}_2$-*ring* R *defines a homomorphism* (11.6) *of* $\mathbf{GL}_2(\mathrm{R})$ *and any* U-*antihomomorphism defines a homomorphism* (11.8) *of* $\mathbf{GL}_2(\mathrm{R})$.

When the units in R (together with 0) form a field k, a U-homomorphism of R is just a k-bimodule homomorphism such that 1 → 1. Thus it induces a monomorphism from k to the image ring, so a U-homomorphism in this case is essentially a k-semilinear mapping. Below we list some special cases of Theorem 11.2.

(i) Let R, S be k-rings and assume that R is a universal GE_2-ring with $\mathbf{U}_0(\mathrm{R}) = k$. Then any k-semilinear mapping of R into S induces a homomorphism of $\mathbf{GL}_2(\mathrm{R})$ into $\mathbf{GL}_2(\mathrm{S})$.

(ii) If R is a k-ring with a weak algorithm, then any k-semilinear mapping to a k-ring S induces a homomorphism of $\mathbf{GL}_2(\mathrm{R})$.

(iii) Any k-linear mapping of a k-ring R with a weak algorithm (into itself) induces an endomorphism of $\mathbf{GL}_2(\mathrm{R})$.

In the case where R is the ring of polynomials over a commutative field in a single indeterminate, this construction of automorphisms of $\mathbf{GL}_2(\mathrm{R})$ is due to Reiner [15], [16].

As an application of the above results, consider the free associative algebra over a commutative field k, on a free generating set which is at most countable. As k-bimodule this is just a vector space of countable dimension over k and so is isomorphic to $k[x]$, the ring of polynomials in a single indeterminate. Hence the $\mathbf{GL}_2$ over the free associative algebra is isomorphic to $\mathbf{GL}_2(k[x])$.

In case ±1 are the only units in R, the statement of Theorem 11.2 can be slightly simplified.

Theorem (**11.3**). — *Let* R *be any ring which is universal for* $\mathbf{GE}_2$ *and in which* ±1 *are the only units, and let* S *be any ring. Then any additive homomorphism from* R *to* S *which maps*

$1 \to 1$ *induces a homomorphism of* $\mathbf{GE}_2(R)$ *into* $\mathbf{GE}_2(S)$ *by the rule* $E(x) \to E(x')$, *and another one by the rule* $E(x) \to E(x')^{-1}$.

Corollary. — *Let* R *be any ring universal for* $\mathbf{GE}_2$, *in which* ± 1 *are the only units, then there is an automorphism of* $\mathbf{GE}_2(R)$ *defined by*

$$E(x) \to E(x)^{-1}.$$

The statement of the various special cases (in particular the case where R is a universal GE_2-ring) may be left to the reader. The result may in particular be applied to the case of discretely ordered rings, where the hypothesis relative to the units in R is automatically satisfied.

12. The analysis of isomorphisms of general linear groups.

We now consider the converse problem: Given an isomorphism between $\mathbf{GL}_n(R)$ and $\mathbf{GL}_n(S)$, when is this induced by a mapping $f: R \to S$? The aim will be to show that f can be taken to be an (anti)isomorphism or in case $n=2$, a U-(anti)isomorphism. We shall limit ourselves to k-rings with a degree-function and to begin with we assume $n=2$. Just as for fields, the case of characteristic 2 has to be treated separately. On the other hand, we do not need to restrict ourselves to GE_2-rings, but instead consider rings in which projective modules are free.

Lemma **(12.1)**. — *Let* R *be a k-ring with a degree function, where k is a field of characteristic not two, and assume that either of the following conditions is satisfied:*

(i) R *is a* GE_2*-ring,*

(ii) *every projective (right)* R*-module on two generators is free,*

then any pair of anticommuting involutions in $\mathbf{GL}_2(R)$ *can be transformed simultaneously by an inner automorphism to the forms* $[1, -1]$ *and* $P(0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ *respectively.*

Proof. — Let A, B be the given involutions. If (i) holds, then by Theorem 5.5, A is conjugate to $[1, -1]$, because A is non-central. If (ii) holds, consider $E = 1/2(I+A)$; clearly E is an idempotent, and since $A \neq \pm I$, it follows that $E \neq 0, I$. If we regard E as acting on R^2, it defines a direct decomposition and by assumption (ii), both the kernel and image of E are free R-modules; taking a suitably adapted basis, we obtain E in diagonal form: $E = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, and in this coordinate system, $A = [1, -1]$. Thus A has been transformed to the form $[1, -1]$, assuming only (i) or (ii). Let $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then $A^{-1}BA = \begin{pmatrix} a & -b \\ -c & d \end{pmatrix}$ and by hypothesis this equals $-B$, hence $a = d = 0$. Moreover, $B^2 = I$, therefore $bc = 1$, i.e. $b, c \in \mathbf{U}(R)$ and transforming B by $[b, 1]$, which leaves A unchanged, we obtain $[c, 1]B[b, 1] = P(0)$.

Now let R be a k-ring and S a k'-ring, both with a degree-function, where k, k' are fields of characteristic $\neq 2$ and assume that one of R and S, say S, is either a $\mathbf{GE}_2$-ring or has its projective modules free. Let

$$(12.1) \qquad f : \mathbf{GL}_2(\mathrm{R}) \rightarrow \mathbf{GL}_2(\mathrm{S})$$

be an isomorphism, and for brevity write $\mathrm{D}_0 = [1, -1]$ (in R or S). Now over any integral domain, $-\mathrm{I}$ is characterized as the only central involution, hence $(-\mathrm{I})f = -\mathrm{I}$. Since D_0 and $\mathrm{P}(0)$ are a pair of anticommuting involutions in $\mathbf{GL}_2(\mathrm{R})$, their images under f are again anticommuting involutions in $\mathbf{GL}_2(\mathrm{S})$, and the same is true of $\eta(\mathrm{D}_0 f)$, $\eta(\mathrm{P}(0)f)$, where η is one of the numbers $1, -1$. Let us make a definite choice of η, then by composing f with an inner automorphism of $\mathbf{GL}_2(\mathrm{S})$ we may assume (by Lemma 12.1) that

$$(12.2) \qquad \mathrm{D}_0 f = \eta \mathrm{D}_0, \qquad \mathrm{P}(0)f = \eta \mathrm{P}(0).$$

By (12.2) the centralizer of D_0 in $\mathbf{GL}_2(\mathrm{R})$ is mapped into the centralizer of D_0 in $\mathbf{GL}_2(\mathrm{S})$. But this centralizer is easily seen to be the set of all diagonal matrices, therefore f maps the subgroup $\mathbf{D}_2(\mathrm{R})$ onto $\mathbf{D}_2(\mathrm{S})$. Our next problem is to characterize the triangular matrices $\mathrm{T}(h) = \begin{pmatrix} 1 & h \\ 0 & 1 \end{pmatrix}$. Writing $\mathrm{T} = \mathrm{T}(h)$, $\mathrm{T}^{\mathrm{D}} = \mathrm{D}^{-1}\mathrm{T}\mathrm{D}$, we clearly have

$$(12.3) \qquad (\mathrm{T}\mathrm{D}_0)^2 = \mathrm{I},$$

$$(12.4) \qquad \mathrm{T}^{\mathrm{D}}\mathrm{T} = \mathrm{T}\mathrm{T}^{\mathrm{D}} \qquad \text{for all } \mathrm{D} \in \mathbf{D}_2(\mathrm{R}).$$

It follows that the image $\mathrm{U} = \mathrm{T}f$ satisfies the same equations. Taking $\mathrm{U} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $\mathrm{D} = [1, \delta]$, we have $\mathrm{U}^{\mathrm{D}}\mathrm{U} = \mathrm{U}\mathrm{U}^{\mathrm{D}}$, and equating the (1, 1)-terms in this equation we obtain

$$(12.5) \qquad b(\delta - \delta^{-1})c = 0.$$

If we assume that k' has more than 3 elements, we can find $\delta \in k'$ such that $\delta^2 \neq 1$, and now (12.5) shows that $b = 0$ or $c = 0$. In the remaining case k' is the field of 3 elements (because the characteristic is $\neq 2$) and from (12.3) we obtain $(\mathrm{U}\mathrm{D}_0)^2 = \mathrm{I}$, whence by equating coefficients,

$$(12.6) \qquad ab = bd, \qquad ca = dc, \qquad a^2 - bc = d^2 - cb = 1.$$

Moreover, in this case, $\mathrm{T}^3 = \mathrm{I}$, hence $\mathrm{U}^3 = \mathrm{I}$; in detail

$$\mathrm{U}^2 = \begin{pmatrix} a^2 + bc & ab + bd \\ ca + dc & d^2 + cb \end{pmatrix} = \begin{pmatrix} 2a^2 - 1 & 2ab \\ 2dc & 2d^2 - 1 \end{pmatrix}$$

$$\mathrm{I} = \mathrm{U}^3 = \begin{pmatrix} 2a^3 - a + 2abc & * \\ * & 2d^3 - d + 2dcb \end{pmatrix},$$

thus $a(2a^2 - 1 + 2bc) = 1$, $d(2d^2 - 1 + 2cb) = 1$. This shows that $a, d \in \mathbf{U}(\mathrm{S}) \subseteq k'$, i.e. $a, d = \pm 1$, therefore $a^2 = d^2 = 1$ and the last equation (12.6) now shows that $bc = 0$. Again we deduce that either $b = 0$ or $c = 0$.

Let us assume that $b\neq 0$, say, then $c=0$, and by (12.6) which always holds, we have $a^2=d^2=1$, i.e. $(a+1)(a-1)=0$. Since S is an integral domain, $a=\pm 1$, and similarly $d=\pm 1$; moreover, the first equation (12.6) shows that $a=d$. This then shows that U is of the form

$$U=\pm\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}, \tag{12.7}$$

and clearly every such matrix satisfies the conditions (12.3) and (12.4) for T. Similarly if $b=0$ but $c\neq 0$ we obtain the form

$$U=\pm\begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix} \tag{12.8}$$

and this matrix again satisfies the conditions (12.3) and (12.4). Finally, if $b=c=0$, then all we can conclude is that $a=\pm 1$ and $d=\pm 1$, and thus we have $\pm I$ or $\pm D_0$. This exhausts all the possibilities.

Now f leaves D_0 fixed except for a scalar factor and it maps diagonal matrices to diagonal matrices, hence f maps any matrix $T(h)$ to a matrix of the form (12.7) or (12.8). Consider $T(1)$; if $T(1)f=T(k)$ say, where $k\in S$, then $k\neq 0$ because f is injective. But for any $h\in R$, $T(h)f$ commutes with $T(k)=T(1)f$, and it must therefore be again of the form (12.7), so in this case the subgroup $\mathbf{B}_{12}(R)$ is mapped into the subgroup $\pm\mathbf{B}_{12}(S)$ by f. The other possibility is that $T(1)=B_{21}(k)$; in this case we modify f by composing it with the inner automorphism defined by $E(0)=D_0P(0)$. This leaves (12.2) unaffected except to replace η by $-\eta$, but now $T(1)f=E(0)^{-1}B_{21}(k)E(0)=T(-k)$, so that this is essentially reduced to the previous case. More precisely, we have shown that by composing f with a suitable inner automorphism we can ensure that (12.2) holds (for $\eta=\pm 1$) and $\mathbf{B}_{12}(R)$ is mapped into $\pm\mathbf{B}_{12}(S)$.

As a consequence we can write

$$T(x)f=\varepsilon(x)T(x^\sigma), \tag{12.9}$$

where $x\rightarrow x^\sigma$ is a mapping of R into S and $x\rightarrow\varepsilon(x)$ is a mapping of R into $\{\pm 1\}$. Since $T(x)T(y)=T(x+y)$, we have

$$\begin{aligned}\varepsilon(x+y)T((x+y)^\sigma)&=\varepsilon(x)\varepsilon(y)T(x^\sigma)T(y^\sigma)\\&=\varepsilon(x)\varepsilon(y)T(x^\sigma+y^\sigma),\end{aligned}$$

hence

$$(x+y)^\sigma=x^\sigma+y^\sigma, \tag{12.10}$$

$$\varepsilon(x)\varepsilon(y)=\varepsilon(x+y). \tag{12.11}$$

Putting $y=x$ in (12.11), we find that $\varepsilon(2x)=\varepsilon(x)^2=1$, hence $\varepsilon(x)=1$ for all $x\in R$. We saw that $x^\sigma\neq 0$ for $x\neq 0$, hence (12.10) shows that the mapping σ is injective. Repeating the argument with f replaced by f^{-1} (and reading (12.2) in the opposite direction), we see that σ is an isomorphism of the additive groups of R and S. Next we look at $E(x)$:

$$E(x)=D_0T(x)P(0).$$

Applying f, we obtain $E(x)f=\eta^2D_0T(x^\sigma)P(0)$, hence

$$E(x)f=E(x^\sigma). \tag{12.12}$$

If we apply f to (2.3) and use (12.12) to simplify the result, we get

$$E((\alpha^{-1})^\sigma)E(\alpha^\sigma)E((\alpha^{-1})^\sigma)=-D(\alpha^{-1})f.$$

The right-hand side belongs to $\mathbf{D}_2(S)$, hence so does the left-hand side; but this can only happen if $\alpha^\sigma\in\mathbf{U}_0(S)$. We know that $\alpha^\sigma\neq 0$ and so we find

$$D(\alpha^{-1})f=-E((\alpha^{-1})^\sigma-(\alpha^\sigma)^{-1})D(\alpha^\sigma)E((\alpha^{-1})^\sigma-(\alpha^\sigma)^{-1}).$$

Equating the (1, 2)-terms we see that

$$(\alpha^{-1})^\sigma=(\alpha^\sigma)^{-1}, \tag{12.13}$$

so both sides of this equation may without risk of ambiguity be denoted by $\alpha^{-\sigma}$. Using (12.13) on the preceding equation we obtain $D(\alpha^{-1})f=D((\alpha^\sigma)^{-1})$, or replacing α by α^{-1} and using (12.13) again,

$$D(\alpha)f=D(\alpha^\sigma). \tag{12.14}$$

In particular, taking $\alpha=1$, we obtain

$$1^\sigma=1. \tag{12.15}$$

We can now use Hua's Theorem (cf. e.g. [1], p. 37) and conclude from (12.10), (12.13) and (12.15) that σ is either a homomorphism or an antihomomorphism of k into k'. The same argument applied to f^{-1} shows that σ is actually a bijection between k and k', and therefore an isomorphism or an anti-isomorphism.

If we apply f to (2.4) we obtain

$$D(\alpha^{-\sigma})E(\alpha^\sigma x^\sigma \alpha^\sigma)=E(x^\sigma)D(\alpha^\sigma)=D(\alpha^{-\sigma})E((\alpha x\alpha)^\sigma), \qquad \text{and hence}$$

$$(\alpha x\alpha)^\sigma=\alpha^\sigma x^\sigma \alpha^\sigma \qquad (x\in R,\ \alpha\in\mathbf{U}(R)). \tag{12.16}$$

This relation can actually be used in the proof of Hua's theorem; more generally, we can use it to show, in exactly the same way as in the proof of Hua's theorem, that we have either

$$(\alpha x\beta)^\sigma=\alpha^\sigma x^\sigma \beta^\sigma, \tag{12.17}$$

or

$$(\alpha x\beta)^\sigma=\beta^\sigma x^\sigma \alpha^\sigma, \tag{12.18}$$

for all $x\in R$ and $\alpha,\beta\in\mathbf{U}(R)$. Thus σ is either a U-isomorphism or a U-anti-isomorphism.

Next we consider the effect of f on diagonal matrices. Every diagonal matrix can be reduced to the form $[1, \alpha]$ by multiplying by a matrix $D(\beta)$, for suitable β. Now $D(\beta)f$ is given by (12.14); we may therefore restrict our attention to $[1, \alpha]$. We know that this is again mapped to a diagonal matrix by f, say

$$[1, \alpha]f = [\alpha^\lambda, \alpha^\mu]. \tag{12.19}$$

Since $P(0)[1, \alpha]P(0) = [\alpha, 1] = [\alpha, \alpha^{-1}][1, \alpha]$, we have

$$[\alpha^\sigma, \alpha^{-\sigma}][\alpha^\lambda, \alpha^\mu] = P(0)[\alpha^\lambda, \alpha^\mu]P(0) = [\alpha^\mu, \alpha^\lambda],$$

and hence

$$\alpha^\mu = \alpha^\sigma \alpha^\lambda. \tag{12.20}$$

Using this relation we obtain

$$[\alpha, \beta]f = ([\alpha, \alpha^{-1}][1, \alpha\beta])f = [\alpha^\sigma, \alpha^{-\sigma}][(\alpha\beta)^\lambda, (\alpha\beta)^\sigma(\alpha\beta)^\lambda], \qquad \text{i.e.}$$

$$[\alpha, \beta]f = [\alpha^\sigma, \alpha^{-\sigma}(\alpha\beta)^\sigma](\alpha\beta)^\lambda. \tag{12.21}$$

Suppose now that σ is a U-isomorphism; then (12.21) states

$$[\alpha, \beta]f = [\alpha^\sigma, \beta^\sigma](\alpha\beta)^\lambda.$$

If we apply this to the equation

$$[1, \alpha\beta] = [1, \alpha][1, \beta], \tag{12.22}$$

we obtain

$$[1, \alpha^\sigma\beta^\sigma](\alpha\beta)^\lambda = [1, \alpha^\sigma]\alpha^\lambda[1, \beta^\sigma]\beta^\lambda \qquad \text{i.e.}$$

$$(\alpha\beta)^\lambda = [1, \beta^{-\sigma}]\alpha^\lambda[1, \beta^\sigma]\beta^\lambda.$$

Equating the (1, 1)-terms in this equation we find

$$(\alpha\beta)^\lambda = \alpha^\lambda\beta^\lambda,$$

and it can be used to simplify the equation between the (2, 2)-terms to

$$\beta^\sigma\alpha^\lambda = \alpha^\lambda\beta^\sigma. \tag{12.23}$$

Thus λ is a homomorphism of $\mathbf{U}(R)$ into $\mathbf{U}(S)$, and since σ maps $\mathbf{U}(R)$ onto $\mathbf{U}(S)$, (12.23) shows that α^λ centralizes $\mathbf{U}(S)$. By (2.4)

$$\begin{aligned}[\beta^\sigma, \alpha^\sigma](\alpha\beta)^\lambda E((\beta^{-1}x\alpha)^\sigma) &= E(x^\sigma)[\alpha^\sigma, \beta^\sigma](\alpha\beta)^\lambda \\ &= [\beta^\sigma, \alpha^\sigma]E(\beta^{-\sigma}x^\sigma\alpha^\sigma)(\alpha\beta)^\lambda.\end{aligned}$$

This shows again that $(\alpha x\beta)^\sigma = \alpha^\sigma x^\sigma \beta^\sigma$ and taking $\beta = 1$, $x = y\alpha^{-1}$ we find $\alpha^\lambda E(y^\sigma) = E(y^\sigma)\alpha^\lambda$, which shows that α^λ lies in the centre of S. Thus λ is a central homothety, and (12.12), (12.21) show that f is just the isomorphism induced by the U-isomorphism σ, followed by the central homothety λ. Finally, a comparison of (12.21) and (12.2) shows that $\eta = (-1)^\lambda$.

Next assume that σ is a U-anti-isomorphism, then (12.21) becomes

$$[\alpha, \beta]f = [\alpha^\sigma, \alpha^{-\sigma}\beta^\sigma\alpha^\sigma](\alpha\beta)^\lambda.$$

If we express λ in terms of σ and μ by means of (12.20), we can write this relation as

$$[\alpha, \beta]f = [\beta^{-\sigma}, \alpha^{-\sigma}](\alpha\beta)^{\mu}.$$

Applying this to (12.22) we find that μ is a homomorphism of $\mathbf{U}(\mathrm{R})$ into the centre of $\mathbf{U}(\mathrm{S})$ and as before we can use (2.4) to show that α^{μ} in fact lies in the centre of S. Thus f is now the isomorphism induced by the U-anti-isomorphism σ, followed by the central homothety μ.

The result may be summed up as

Theorem (**12.2**). — *Let* R *be a k-ring and* S *a k'-ring, both with a degree-function, where k and k' are any fields of characteristic* $\neq 2$ *and* S *is either a* GE_2*-ring or all projective* S*-modules on two generators are free. Then every isomorphism between* $\mathbf{GL}_2(\mathrm{R})$ *and* $\mathbf{GL}_2(\mathrm{S})$ *is obtained by taking the isomorphism induced by a* U*-isomorphism or* U*-anti-isomorphism, followed by a central homothety and an inner automorphism.*

This includes the result of Reiner [16] when $\mathrm{S} = \mathrm{R} = k[x]$ and earlier results of Schreier-v.d. Waerden and Hua (cf. [8] and the references given there).

If R is any k-ring, then the characteristic of k is $\neq 2$ if and only if $\mathbf{GL}_2(\mathrm{R})$ contains a central involution. Hence if we are given an isomorphism

$$f : \mathbf{GL}_2(\mathrm{R}) \cong \mathbf{GL}_2(\mathrm{S}),$$

where R is a k-ring and S a k'-ring, then k, k' both have characteristic $\neq 2$ or both $= 2$. In order to deal with the latter case we need a Lemma analogous to Lemma 12.1. Whereas Lemma 12.1 shows that in characteristic not two (and under the given hypotheses) the subgroups of $\mathbf{GL}_2(\mathrm{R})$ of the type of the dihedral group of order eight form a single conjugacy-class, the next lemma establishes a corresponding fact for subgroups of the type of the symmetric group of degree 3, in the case of characteristic 2.

Lemma (**12.3**). — *Let* R *be a k-ring with a degree-function, where k is a field of characteristic 2, and assume further that* R *is a* GE_2*-ring. Then any pair of involutions in* $\mathbf{GL}_2(\mathrm{R})$ *whose product has order 3, can be transformed simultaneously by an inner automorphism to the forms* $\mathrm{T}(1)(=\mathrm{B}_{12}(1))$, $\mathrm{P}(0)$ *respectively.*

Proof. — Let $\mathrm{A}, \mathrm{B} \in \mathbf{GL}_2(\mathrm{R})$ be the given involutions and write $\mathrm{C} = \mathrm{AB}$. By Prop. 5.4 and the remark following it, C can be brought to one of the forms

$$[1, \beta]\mathrm{E}(a), \qquad [\alpha, \beta]\mathrm{E}(0)\mathrm{E}(b). \tag{12.24}$$

Suppose that C has the second of these forms; comparing terms in the equation $\mathrm{C}^3 = \mathrm{I}$, we see that $\alpha^3 = \beta^3 = 1$, $b\alpha^2 + \beta b\alpha + \beta^2 b = 0$. If we transform by $\mathrm{E}(\beta^2 b\alpha)^{-1}$, we therefore obtain

$$\begin{aligned}\mathrm{E}(\beta^2 b\alpha)[\alpha, \beta]\mathrm{E}(0)\mathrm{E}(b)\mathrm{E}(0)\mathrm{E}(\beta^2 b\alpha)\mathrm{E}(0) &= [\beta, \alpha]\mathrm{E}(\beta b\alpha^2)\mathrm{E}(0)\mathrm{E}(b)\mathrm{E}(0)\mathrm{E}(\beta^2 b\alpha)\mathrm{E}(0)\\ &= [\beta, \alpha]\mathrm{E}(b + \beta^2 b\alpha + \beta b\alpha^2)\mathrm{E}(0)\\ &= [\beta, \alpha].\end{aligned}$$

Thus C has been transformed to diagonal form. If C has the first form in (12.24), the equation $\mathrm{C}^3 = \mathrm{I}$ shows that $\beta = a^2$, $a^3 = 1$. When $a \neq 1$, it follows that $a^2 + a + 1 = 0$, and transforming by $\begin{pmatrix} 1 & 1 \\ 1 & \beta \end{pmatrix}$ we reduce C to the form $[\beta, 1]$. Thus C has been reduced to either of the following forms

$$(12.25) \qquad \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \qquad [\alpha, \beta].$$

We assert that in fact C can always be reduced to the first of these forms. Since C has order 3, we have $\alpha^3 = \beta^3 = 1$ and α, β are not both 1 in the second form (12.25). Suppose that $\alpha = 1$, say and let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then by equating (1, 1)-terms in the equations $A^2 = (AC)^2 = I$, we find $a^2 + bc = a^2 + b\beta c = 1$, hence $b(\beta - 1)c = 0$; therefore b or c must vanish, say $b = 0$. Then $a = d = 1$, and equating (2,2)-terms in $(AC)^2 = 1$ we find $\beta^2 = 1$, i.e. $\beta = 1$, which is a contradiction; the same reasoning applies if $c = 0$. Thus neither α nor β in (12.25) can be 1 and they therefore satisfy the equation

$$(12.26) \qquad x^2 + x + 1 = 0.$$

Assume first that this equation has a root in the centre of k, say ω. Then $(\alpha - \omega)(\alpha - \omega^2) = 0$, hence $\alpha = \omega$ or $\alpha = \omega^2$, and likewise for β. If $\beta = \alpha$, then C is a scalar matrix; leaving this case aside for the moment, so as not to interrupt the argument, we may assume that $\beta \neq \alpha$; it follows that $\beta = \alpha^{-1}$, and transforming C by $\begin{pmatrix} 1 & \beta \\ 1 & \alpha \end{pmatrix}$ we obtain the first form in (12.25) for C. If (12.26) has no roots in the centre Z of k, it must be irreducible over Z and so all its roots in k are conjugate (Herstein [10]). Hence there exists $\gamma \in k$ such that $\gamma^{-1}\beta\gamma = \alpha^{-1}$ and transforming $[\alpha, \beta]$ by $[1, \gamma]$ we are reduced to the previous case.

Thus C has now been reduced to the form $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ (with the exception noted). Now by hypothesis, $C = AB$, $C^{-1} = BA$, therefore

$$(12.27) \qquad CB = A = BC^{-1}.$$

Taking $B = \begin{pmatrix} u & v \\ w & z \end{pmatrix}$ and using the expression found for C we obtain by equating components in (12.27) and simplifying,

$$z = u, \qquad u + v + w = 0.$$

If we equate terms in $B^2 = I$, we find that $uv = vu$, $u^2 + vw = 1$. Moreover, $A = \begin{pmatrix} v & w \\ u & v \end{pmatrix}$; and by Theorem 5.5, A is conjugate to $T(h)$, for some $h \in R$. Thus there exists an invertible matrix $P = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ such that

$$(12.28) \qquad AP = PT(h).$$

Equating coefficients, we obtain the equations

$$\begin{aligned} va + wc &= a, & vb + wd &= ah + b, \\ ua + vc &= c, & ub + vd &= bh + d, \end{aligned}$$

or after some simplification,

$$(12.29)\qquad \begin{cases} (v+1)a=wc, & (v+1)b+wd=ah,\\ ua=(v+1)c, & ub+(v+1)d=bh.\end{cases}$$

Now (12.28) may be written as

$$(A+I)P=P\begin{pmatrix}0 & h\\ 0 & 0\end{pmatrix}.$$

The first row of this matrix equation reads

$$(12.30)\qquad (v+1,\ w)P=(0,\ ah),$$

hence

$$(12.31)\qquad (v+1,\ w)=(0,\ ah)P^{-1}.$$

From the first two equations (12.29), we have

$$(12.32)\qquad (v+1)c+w(a+c)=a,$$

and by (12.31), ah is a common left factor of $v+1$ and w, therefore $a=ahk$, which shows h to be a unit, $h=\eta$ say. Combining (12.32) and the third equation of (12.29) we obtain an equation which in matrix form can be written

$$(v+1,\ w)\begin{pmatrix} b+c\eta\\ d+(a+c)\eta\end{pmatrix}=0.$$

Using (12.30), we may write this as

$$(0,\ a\eta)P^{-1}\begin{pmatrix} b+c\eta\\ d+(a+c)\eta\end{pmatrix}=0.$$

If $a=0$, then $c\neq 0$ and the first two equations (12.29) show that $w=0$, $v=1$, and hence $u=1$. In this case transformation by C^{-1} achieves the desired reduction. If $a\neq 0$, then $a\eta\neq 0$, and from the last equation written it follows that

$$P^{-1}\begin{pmatrix} b+c\eta\\ d+(a+c)\eta\end{pmatrix}=\begin{pmatrix}k\\ 0\end{pmatrix}\quad \text{for some } k\in R.$$

Hence

$$\begin{pmatrix} b+c\eta\\ d+(a+c)\eta\end{pmatrix}=P\begin{pmatrix}k\\ 0\end{pmatrix};$$

thus $b=c\eta+ak$, $d=(a+c)\eta+ck$. Inserting these values in P, we find that

$$P=\begin{pmatrix}a & b\\ c & d\end{pmatrix}=\begin{pmatrix}a & ak+c\eta\\ c & ck+(a+c)\eta\end{pmatrix}=\begin{pmatrix}a & c\\ c & a+c\end{pmatrix}\begin{pmatrix}1 & k\\ 0 & \eta\end{pmatrix}.$$

This shows the matrix $P_1=\begin{pmatrix}a & c\\ c & a+c\end{pmatrix}$ to be invertible, evidently it commutes with $C=\begin{pmatrix}1 & 1\\ 1 & 0\end{pmatrix}$ and transforming A by P_1 we obtain T(1). The same transformation reduces B=AC to the form P(0), as we wished to show.

It only remains to show that the case $C=\omega I$, where ω is a root of (12.26) in the centre of k, cannot occur. Let us assume the contrary and write

$$R^+=\{x\in R\,|\,x\omega=\omega x\},\qquad R^-=\{x\in R\,|\,x\omega=\omega^2 x\}.$$

For brevity the elements in R^+ will be called *symmetric* and the elements in R^- *skew*. By hypothesis, every element of k is symmetric; further, every element of R can be expressed as the sum of a symmetric element and a skew element in just one way, for if

(12.33) $$x=x^+ + x^- \qquad (x^\varepsilon\in R,\ \varepsilon=\pm),$$

then

(12.34) $$\omega x=x^+\omega+x^-\omega^2,$$

and solving the equations (12.33-4) we find

$$x^+=x\omega^2+\omega x,\qquad x^-=x\omega+\omega x;$$

thus only one choice is possible for x^+, x^- and this choice clearly satisfies (12.33). By (12.27) we have

$$\omega B=B\omega^2,$$

hence all the elements of B are skew. Now B is invertible and hence of the form

$$[\alpha,\beta]E(q_1)\dots E(q_r).$$

Denoting the first row of B by (a, b), we have

(12.35) $$a=bq_r+a',$$

where $d(a')<d(b)$, by Lemma 5.1. If we equate the symmetric components here we obtain

$$b(q_r)^-+(a')^+=0.$$

But $d((a')^+)\leqslant d(a')<d(b)$, hence $(a')^+=(q_r)^-=0$ and so a' is skew and q_r is symmetric. Therefore $BE(q_r)^{-1}$ consists entirely of skew elements; by induction on r we conclude that $[\alpha,\beta]$ consists of skew elements, but this contradicts the fact that $k\subseteq R^+$. This completes the proof of Lemma 12.3.

With the help of this lemma it is now an easy matter to obtain an analogue of Theorem 12.2 for the case of characteristic 2. Let

$$f:\mathbf{GL}_2(R)\rightarrow\mathbf{GL}_2(S)$$

be an isomorphism, where R is a k-ring and S is a k'-ring, both with a degree-function, k and k' are fields of characteristic 2 and moreover S is a GE_2-ring. In $\mathbf{GL}_2(R)$ the matrices $T(1)$, $P(0)$ form a pair of involutions whose product has order 3, hence so do their images in $\mathbf{GL}_2(S)$, and by Lemma 12.3 we may therefore assume (by combining f with a suitable inner automorphism of $\mathbf{GL}_2(S)$),

(12.36) $$T(1)f=T(1),\qquad P(0)f=P(0).$$

The subgroup $\mathbf{B}_{12}(\mathrm{R})$ of $\mathbf{GL}_2(\mathrm{R})$ may be characterized as the maximal abelian subgroup of exponent 2 in the centralizer of $\mathrm{T}(1)$ in $\mathbf{GL}_2(\mathrm{R})$ [1]. It is therefore mapped to $\mathbf{B}_{12}(\mathrm{S})$, the maximal abelian subgroup of exponent 2 in the centralizer of $\mathrm{T}(1)$ in $\mathbf{GL}_2(\mathrm{S})$. Thus there is a mapping $\sigma : \mathrm{R} \rightarrow \mathrm{S}$ such that

$$\mathrm{T}(x)f = \mathrm{T}(x^\sigma) \qquad x \in \mathrm{R}.$$

By (2.2), we have

$$(x+y)^\sigma = x^\sigma + y^\sigma, \tag{12.37}$$

and by (12.36),

$$1^\sigma = 1. \tag{12.38}$$

Now a diagonal matrix D in $\mathbf{GL}_2(\mathrm{R})$ is characterized by the fact that both D and $\mathrm{P}(0)\mathrm{D}\mathrm{P}(0)$ normalize $\mathbf{B}_{12}(\mathrm{R})$. This shows that f maps diagonal matrices over R to diagonal matrices over S; we can now follow the proof of Theorem 12.2 exactly and finally obtain

Theorem **(12.4)**. — *Let* R *be a k-ring and* S *a k′-ring, both with a degree-function, where k, k′ are any fields of characteristic 2 and* S *is a* GE_2*-ring. Then every isomorphism between* $\mathbf{GL}_2(\mathrm{R})$ *and* $\mathbf{GL}_2(\mathrm{S})$ *is obtained by taking the isomorphism induced by a* U*-isomorphism or a* U*-anti-isomorphism, followed by a central homothety and an inner automorphism.*

In conclusion we briefly discuss the isomorphisms of $\mathbf{GL}_n(\mathrm{R})$. It turns out that by a method similar to that used in proving Theorem 12.2, we obtain

Theorem **(12.5)**. — *Let* R *be a k-ring and* S *a k′-ring, both with a degree-function, where k and k′ are fields of characteristic* $\neq 2$, *and assume further that every finitely generated projective* S*-module is free. Then every isomorphism between* $\mathbf{GL}_n(\mathrm{R})$ *and* $\mathbf{GL}_n(\mathrm{S})$ (*for* $n \geqslant 3$) *is obtained by taking an isomorphism or anti-isomorphism from* R *to* S, *followed by a central homothety and an inner automorphism.*

Proof. — $\mathbf{GL}_n(\mathrm{R})$ contains a system $\mathscr{J}$ of 2^n commuting involutions, namely $[\pm 1, \ldots, \pm 1]$; moreover, the symmetric group Σ of degree n acts on this set by permutations: there are $n+1$ orbits, of $\binom{n}{k}$ elements respectively $(k = 0, 1, \ldots, n)$. The isomorphism f transforms $\mathscr{J}$ into a set of 2^n commuting involutions in $\mathbf{GL}_n(\mathrm{S})$. Now if P is any involution in $\mathbf{GL}_n(\mathrm{S})$, then $\mathrm{E} = 1/2(\mathrm{I}+\mathrm{P})$ is an idempotent, so that we get a set of 2^n commuting idempotents in the matrix ring S_n. Since projective S-modules are free, any idempotent can be diagonalized, and likewise any commuting set of idempotents can be transformed simultaneously to diagonal form, with diagonal elements ± 1. By applying this transformation we thus bring the 2^n commuting idempotents to the form $\mathrm{E} = [e_1, \ldots, e_n]$, where $e_i = 0$ or 1. Now $\mathrm{P} = 2\mathrm{E} - \mathrm{I}$ is again an involution, so that we have a system of 2^n commuting involutions in diagonal form in $\mathbf{GL}_n(\mathrm{S})$. More precisely, by applying a suitable inner automorphism, we may assume that for any

[1] This subgroup may also be described as the set of all involutions in the centralizer of T(1), together with I.

involution $P=[\varepsilon_1, \ldots, \varepsilon_n]$, $\varepsilon_i=\pm 1$, Pf is again diagonal, with ± 1 on the diagonal. If k of the ε_i are $+1$ and the rest are -1, P is said to be of type $(k, n-k)$ or a $(k, n-k)$-involution. The subgroup of Σ centralizing a $(k, n-k)$-involution has the form $\Sigma_k \times \Sigma_{n-k}$; hence if P is of type $(k, n-k)$, then Pf is of type $(k, n-k)$ or $(n-k, k)$, i.e. either Pf or $-Pf$ is of type $(k, n-k)$. Consider the n $(1, n-1)$-involutions $P_1, \ldots, P_n$; they form an orbit under Σ, hence $P_1 f, \ldots, P_n f$ are conjugate and so they are all of type $(1, n-1)$ or $(n-1, 1)$. By combining f with a suitable inner automorphism we may thus assume that

(12.39)
$$P_i f = \eta P_i \qquad (\eta=\pm 1, i=1, \ldots, n).$$

Since every involution in $\mathscr{J}$ is a product of P_i's, we have, for any involution $P \in \mathscr{J}$,

$$Pf=\eta_P P \qquad \text{where } \eta_P=\eta^k \text{ if P has type } (k, n-k).$$

Let S_σ be the permutation matrix corresponding to the permutation $\sigma \in \Sigma$; then $S_\sigma^{-1} P_i S_\sigma = P_{i\sigma}$, hence applying f and noting (12.39), we obtain

$$(S_\sigma f)^{-1} P_i (S_\sigma f) = P_{i\sigma}.$$

Therefore $S_\sigma (S_\sigma f)^{-1}$ commutes with P_i for $i=1, \ldots, n$ and hence is diagonal, say

(12.40)
$$S_\sigma f = D_\sigma S_\sigma.$$

Consider the n-cycle $\rho=(1\ 2\ \ldots\ n)$ and let $D_\rho=[\alpha_1, \ldots, \alpha_n]$; then the equation $S^n=I$ shows that

(12.41)
$$\alpha_n \alpha_{n-1} \ldots \alpha_1 = 1.$$

If we transform the images under f by a fixed diagonal matrix $T=[\gamma_1, \ldots, \gamma_n]$, the $P \in \mathscr{J}$ remain unchanged and (12.40) takes on the form

$$S_\sigma f = D'_\sigma S_\sigma, \qquad \text{where } D'_\sigma = T^{-1} D_\sigma . S_\sigma T S_\sigma^{-1}.$$

In particular, we have

$$D'_\rho = [\gamma_1^{-1} \alpha_1 \gamma_n, \gamma_2^{-1} \alpha_2 \gamma_1, \ldots, \gamma_n^{-1} \alpha_n \gamma_{n-1}].$$

Our aim will be to choose $\zeta=\pm 1$ and γ_i $(i=1, \ldots, n-1)$ such that

(12.42)
$$\gamma_i^{-1} \alpha_i \gamma_{i-1} = \zeta \qquad (i=1, \ldots, n;\ \gamma_0=\gamma_1=1).$$

If this is to hold, then by (12.41) we must have

$$\gamma_1=\alpha_1\zeta, \qquad \gamma_2=\alpha_2\gamma_1\zeta=\alpha_2\alpha_1\zeta^2, \qquad \ldots, \qquad \gamma_{n-1}=\alpha_{n-1}\ldots\alpha_1\zeta^{n-1}, \qquad \gamma_n=\zeta^n=1;$$

conversely, these equations will ensure that (12.42) holds. Now these equations can always be solved if n is even, while for odd n they can be solved provided that $\zeta=1$. To sum up, we can always transform by an inner automorphism (induced by a diagonal matrix) such that for $\rho=(1\ 2\ \ldots\ n)$, (12.40) reduces to

(12.43)
$$S_\rho f = \chi(\rho) S_\rho,$$

where χ is a linear character of Σ.

Next consider a transposition, say $\tau=(12)$; if $D_\tau=[\beta_1, \ldots, \beta_n]$, the relation $S_\tau^2=I$ shows that $\beta_1\beta_2=1$, $\beta_i^2=1$ $(i\geqslant 3)$. If we put S_τ into diagonal form for a moment we see that it is a $(1, n-1)$-involution, and hence $S_\tau f$ is of type $(1, n-1)$ or $(n-1, 1)$; this means $\beta_3=\beta_4=\ldots=\beta_n$ $(=\pm 1)$. Thus

$$D_\tau=[\beta, \beta^{-1}, \delta, \ldots, \delta] \qquad (\delta=\pm 1).$$

Now the equation $(1\ 2\ \ldots\ n)=(n\ n-1)(n-1\ n-2)\ \ldots\ (3\ 2)(3\ 1)$ shows that

$$S_\rho=S_\tau^{\rho^{n-2}}S_\tau^{\rho^{n-3}}\ldots S_\tau,$$

where S^ρ denotes the transform of S by S_ρ. Applying f and equating diagonal terms, we find that

$$(12.44) \qquad \beta\delta^{n-2}=\chi(\rho), \qquad \beta^{n-1}=\chi(\rho).$$

If n is odd, $\chi(\rho)=1$ and these equations reduce to

$$\beta=\delta.$$

In this case $D_\tau=\delta I$ is a scalar matrix and since ρ, τ generate Σ, every D_σ is scalar, in fact,

$$(12.45) \qquad S_\sigma f=\chi(\sigma)S_\sigma,$$

where $\chi(\sigma)$ is a linear character of Σ, the identity or the alternating character according as $\delta=1$ or $=-1$. If n is even, the equations (12.44) reduce to

$$\beta=\chi(\rho)\ (=\chi(\tau)),$$

and in this case ζ is still at our disposal. We now choose $\zeta=\chi(\rho)$; then D_τ is again a scalar matrix and so (12.45) holds in this case too.

Now the diagonal matrices may be characterized as the set of matrices centralized by the P's; consider the set centralized by

$$(12.46) \qquad [1, 1, \varepsilon_3, \ldots, \varepsilon_n] \qquad \varepsilon_i=\pm 1.$$

Clearly this is the set of all $A\dotplus[d_3, \ldots, d_n]$, where $A\in\mathbf{GL}_2(R)$. Hence the set centralized by the elements (12.46) and the permutations not involving 1 or 2 is $A\dotplus\lambda I_{n-2}$ $(\lambda\in k)$. Thus f induces a mapping $A\dotplus I_{n-2}\rightarrow A\bar{f}\dotplus\lambda(A)I_{n-2}$, where $\bar{f}$ is an isomorphism $\mathbf{GL}_2(R)\rightarrow\mathbf{GL}_2(S)$. Transforming both sides by S_σ (for $\sigma\in\Sigma$) we see that a corresponding formula holds for other rows and columns. Now by Theorem 12.2, f is of the form $f_1f_2f_3$, where f_1 is induced by a U-isomorphism or a U-anti-isomorphism φ, say the former, f_2 is a central homothety and f_3 is an inner automorphism. Thus f_1 maps $B_{ij}(x)$ to $B_{ij}(x^\varphi)$, and since

$$(B_{12}(x), B_{23}(y))=B_{13}(xy)$$

(where $(A, B)=A^{-1}B^{-1}AB$), we have

$$(B_{12}(x^\varphi), B_{23}(y^\varphi))=B_{13}((xy)^\varphi), \qquad \text{i.e.}$$

$$(xy)^\varphi=x^\varphi y^\varphi.$$

Thus φ is in fact an isomorphism; similarly a U-anti-isomorphism is shows to be an anti-isomorphism. By taking commutators with elements of the form $I_2 \dotplus B$ we see that $\lambda(A)$ lies in the centre of S; dividing by $\lambda(A)$ we thus have a homomorphism

$$A \dotplus I_{n-2} \to \theta(A)A^{\varphi} \dotplus I_{n-2},$$

and by definition, $\theta(A) = I$ on involutions. Further,

$$[\alpha, 1, 1, \ldots, 1] \to [\theta(\alpha)\alpha^{\varphi}, \theta(\alpha), 1, \ldots, 1];$$

permuting the second and third rows and columns (which does not change the left-hand side) we see that $\theta([\alpha, 1]) = 1$. Therefore $\theta = 1$ and the result follows.

REFERENCES

[1] E. Artin, *Geometric Algebra* (New York, 1957).

[2] H. Bass, K-theory and stable algebra, *Publ. math. I.H.E.S.*, n° 22, Paris (1964).

[3] H. Bass, Projective modules over free groups are free, *J. of Algebra*, 1 (1964), 367-373.

[4] P. M. Cohn, On a generalization of the Euclidean algorithm, *Proc. Cambridge Phil. Soc.*, 57 (1961), 18-30.

[5] P. M. Cohn, Rings with a weak algorithm, *Trans. Amer. Math. Soc.*, 109 (1963), 332-356.

[6] P. M. Cohn, Free ideal rings, *J. of Algebra*, 1 (1964), 47-69.

[7] J. Dieudonné, Les déterminants sur un corps non commutatif, *Bull. Soc. Math. France*, 71 (1943), 27-45.

[8] J. Dieudonné, La géométrie des groupes classiques, *Ergeb. d. Math.*, n° 5 (Berlin, 1955).

[9] G. H. Hardy et E. M. Wright, *Introduction to the theory of numbers*, 2nd ed. (Oxford, 1945).

[10] I. N. Herstein, An elementary proof of a theorem of Jacobson, *Duke Math. J.*, 21 (1954), 45-48.

[11] L. K. Hua et I. Reiner, Automorphisms of the unimodular group, *Trans. Amer. Math. Soc.*, 71 (1951), 331-348.

[12] W. Klingenberg, Die Struktur der linearen Gruppen über einem nichtkommutativen lokalen Ring, *Archiv d. Math.*, 13 (1962), 73-81.

[13] J. Landin et I. Reiner, Automorphisms of the two-dimensional general linear group over a Euclidean ring, *Proc. Amer. Math. Soc.*, 9 (1958), 209-216.

[14] H. Nagao, On GL(2, K[x]), *J. Inst. Polytech. Osaka City Univ.*, Ser. A, 10 (1959), 117-121.

[15] I. Reiner, A theorem on continued fractions, *Proc. Amer. Math. Soc.*, 8 (1957), 1111-1113.

[16] I. Reiner, A new type of automorphism of the general linear group over a ring, *Ann. of Math.*, 66 (1957), 461-466.

[17] B. L. v. d. Waerden, *Moderne Algebra*, I (Berlin, 1937).

[18] J. M. H. Wedderburn, Non-commutative domains of integrity, *J. reine u. angew. Math.*, 167 (1932), 129-141.

Queen Mary College,
University of London.

Manuscrit reçu le 15 juillet 1965.