MATHEMATICS

SKEW POLYNOMIAL RINGS

BY

T. H. M. SMITS

(Communicated by Prof. J. Popken at the meeting of October 28, 1967)

Abstract

Skew polynomial rings are considered with a multiplication defined by

$$x \cdot a = a_1 x + a_2 x^2 + \ldots + a_r x^r, \qquad a_i \in K,$$

where K is a (skew) field and the a_i depend on $a \in K$. Under certain conditions the rings appear to be non-commutative principal ideal rings with a unique factorization.

§ 1. Introduction

Given a field [1] K, an endomorphism σ and a $(1, \sigma)$-derivation δ of K, we denote by $K[x; \sigma, \delta]$ the ring of skew polynomials in an indeterminate x subject to the commutation formula

$$x \cdot a = (a\sigma)x + a\delta \qquad (a \in K).$$

Every element of this ring can be written uniquely in the form $\sum_{i=0}^{n} k^{(i)} x^i$ ($k^{(i)} \in K$, polynomial of degree n) and the multiplication is completely determined by the commutation formula. This ring has been first described by Ore [8], who showed that the ring $K[x; \sigma, \delta]$ satisfies a right division algorithm, hence every left ideal of this ring is principal (left principal ideal domain, left PID for short). Because of the fact that every element (polynomial) has a finite prime factorization we conclude by [4, Theorem 5.5, Corollary 1] that this ring is a *unique factorization domain* (UFD for short). The notion of a UFD is called to mind by the following definitions:

1. Two elements a, b of a ring R are said to be *similar*, if $R/aR \cong R/bR$, as right R-modules; in an integral domain (ring with a unit-element and without zero-divisors) this implies $R/Ra \cong R/Rb$ (cf. Jacobson [7], p. 33 or Cohn [4], p. 314);
2. A UFD is an integral domain such that every nonunit has a factorization into primes; two different factorizations of the same element have the same number of prime factors and the factors are similar in pairs (cf. Cohn [4], p. 317).

[1] The term "field" will be used in the sense of "skew field", i.e. "not necessarily commutative division ring".

In this paper we consider the ring R of polynomials over the field K in a single indeterminate x with the commutation formula

$$x \cdot a = a_1 x + a_2 x^2 + \ldots + a_r x^r,\ a_i \in K \qquad (r = 2, 3, \ldots),$$

where the a_i depend on a. By the associative and distributive laws we obtain many relations between the mappings $\delta_i : a \to a_i$ $(i = 1, 2, \ldots, r)$. If we assume that $\delta_2, \delta_3, \ldots, \delta_r$ are right K-independent, then these relations can be simplified.

In § 6 we derive that R satisfies a right Euclidean algorithm, hence R is a left PID. Because of a finite prime factorization R is a UFD (Theorem 4, § 6). From the relations obtained by the associative law it can be derived that if α is the mapping $a \to a_1$, i.e. $a_1 = a\alpha$, then α is an endomorphism of K and $\delta_2 : a \to a_2$ is an (α^2, α)-derivation of K.

Assume further α is an automorphism of K with inverse β and put $a_2 = a\delta\alpha$, then δ is an α-derivation of K. In § 2 we derive

$$a_k = a\delta^{k-1}\alpha \qquad (k = 1, 2, \ldots),$$

hence from $0 = a_{r+1} = a\delta^r\alpha$ we observe that δ is a nilpotent α-derivation of K and the index of nilpotence of δ is r.

It is rather remarkable that $y = x^{-1}$ satisfies again the linear Ore-rule

$$ya = (a\beta)y - a\beta\delta \text{ (Theorem 2).}$$

Thus we can obtain $R = K[x]$ in the following way. Take any skew polynomial ring $K[y; \beta, -\beta\delta]$ with $\delta^r = 0$, adjoin $y^{-1} = x$ and take subring of polynomials in x. The rings $R = K[x]$ and $\bar{R} = K[y; \beta, -\beta\delta]$ have the same quotient field $Q = K(y; \beta, -\beta\delta)$. We can also obtain $K[x]$ as follows. Take the skew polynomial ring $L = K[t, \alpha^r, 0]$, then $R = L[x]$ with $x^r = t$ (R/L is generated by x).

In the sections 3, 4 and 5 we investigate the polynomial units of $R = K[x]$. First we give examples of splitting up units as products of units of lower degree. In § 5 we derive in a few lines the general decomposition theorem of units. Every unit can be written as a product of units of degree $\leqq r-1$, for $r \neq 2$ even as a product of units of degree $\leqq r-2$. The group G of units is multiplicatively generated by r isomorphic fields of units K_i of the form $x^i K y^i$ $(i = 0, 1, 2, \ldots, r-1)$. For $r = 2$ the ring R is the free product of K and $K_1 = xKx^{-1}$ (§ 5, Theorem 3). If $r \geqq 3$, then the ring R is a proper homomorphic image of the free product of the fields K and K_1 over the constant field C (Theorem 3).

Some of the results of the present paper were included in a doctoral dissertation presented to the Technological University Delft in 1967 and carried under the instructions of Professor H. J. A. Duparc. The author wishes to thank the Professors P. M. Cohn (Queen Mary College, London) and H. J. A. Duparc (Technological University Delft) for their many helpful suggestions and inspiration.

§ 2. Multiplication

Let K denote an arbitrary commutative or non-commutative field with an arbitrary characteristic. The objects of our investigations are polynomials in a formal variable x

$$F(x)=a^{(0)}x^n+a^{(1)}x^{n-1}+\ldots+a^{(n)}, \tag{1}$$

with $a^{(\nu)}$ belonging to K. Let

$$G(x)=b^{(0)}x^m+b^{(1)}x^{m-1}+\ldots+b^{(m)} \tag{2}$$

be a second polynomial of the same kind. Sum and difference $F(x) \pm G(x)$ are defined as usual i.e. by the polynomial one obtains from (1) and (2) by adding or subtracting corresponding coefficients. The polynomial $cF(x)$, where c is an arbitrary element of K, is the polynomial obtained from $F(x)$ by multiplying all coefficients on the left with c. The polynomials (1) therefore form an additive Abelian group with K as domain of multipliers.

We shall now define multiplication for the additive group formed by the polynomials (1), so that the group becomes a ring. We assume that the multiplication of polynomials is associative and both-sided distributive.

It is clear that, due to the distributive property, it suffices to define the product $x\cdot a$. We assume the commutation formula

$$x\cdot a=a_rx^r+a_{r-1}x^{r-1}+\ldots+a_2x^2+a_1x \qquad (r=1,\,2,\,\ldots), \tag{3}$$

where $a_r, a_{r-1}, \ldots, a_1$ are elements of K depending on a. From (3) one easily obtains

$$x(a+b)=xa+xb=(a_r+b_r)x^r+(a_{r-1}+b_{r-1})x^{r-1}+\ldots+(a_1+b_1)x.$$

This leads to the following relations

$$(a+b)_i=a_i+b_i, \qquad i=1,\,2,\,\ldots,\,r, \tag{4}$$

hence the mapping $\delta_i: a \to a_i$ $(i=1, 2, \ldots, r)$ is an endomorphism of the additive group of the field K.

In the special case of a skew polynomial ring $K[x;\,\sigma,\,0]$ we have $\delta_i=0$ $(i\geqq 2)$ and $\delta_1=\sigma$.

Now the special properties of the mappings δ_i will be discussed further.

The principal formula (3) yields

$$\begin{aligned} x^2a&=(xa_r)x^r+(xa_{r-1})x^{r-1}+\ldots+(xa_1)x=\\ &=a_{11}x^2+(a_{12}+a_{21})x^3+(a_{31}+a_{22}+a_{13})x^4+\ldots+a_{rr}x^{2r}=\\ &=a_{(2,2)}x^2+a_{(2,3)}x^3+a_{(2,4)}x^4+\ldots+a_{(2,2r)}x^{2r}, \end{aligned}$$

where

$$a_{(2,l)}=\sum_{i+j=l} a_{ij}.$$

Similarly

$$a_{(k,l)}=\sum_{\substack{i_1+i_2+\ldots+i_k=l\\ i_j=1,\ldots,r}} a_{i_1 i_2 \ldots i_k}.$$

We obtain by induction on k:

$$x^k a = \sum_{i=k}^{kr} a_{(k,i)} x^i, \quad k = 1, 2, \ldots, r. \tag{5}$$

Because of

$$a_{(k,i)} = 0 \text{ for } i > kr \text{ and } i < k \tag{6}$$

formula (5) can be rewritten in the form

$$x^k a = \sum_{i=1}^{\infty} a_{(k,i)}\, x^i. \tag{7}$$

Important relations are found from the associative law $(xa)b = x(ab)$, all $a, b \in K$. One has by (7)

$$(xa)b = \sum_{i=1}^{\infty} a_i x^i b = \sum_{i,j=1}^{\infty} a_i b_{(i,j)}\, x^j,$$

and

$$x(ab) = \sum_{i=1}^{\infty} (ab)_i\, x^i = \sum_{i=1}^{r} (ab)_i\, x^i.$$

Equate coefficients:

$$(ab)_i = \sum_{k=1}^{r} a_k b_{(k,i)} \qquad (i = 1, 2, \ldots), \tag{8}$$

e.g.

$$(ab)_1 = a_1 b_1, \tag{9}$$

$$(ab)_2 = a_2 b_{11} + a_1 b_2, \tag{10}$$

$$(ab)_3 = a_3 b_{111} + a_2(b_{21} + b_{12}) + a_1 b_3, \tag{11}$$

$$(ab)_r = a_r b_{(r,r)} + a_{r-1} b_{(r-1,r)} + \ldots + a_1 b_r. \tag{12}$$

For $i > r$ the left-hand side of (8) vanishes, hence

$$\sum_{k=2}^{r} a_k\, b_{(k,i)} = 0 \qquad (i \geqq r+1). \tag{13}$$

To be able to continue the calculations we require two assumptions (**B**) and (C).

ASSUMPTION (B). If $\delta_i : a \to a_i$ then $\delta_2, \ldots, \delta_r$ are right K-independent, i.e.

$$(u\delta_2)c^{(2)} + (u\delta_3)c^{(3)} + \ldots + (u\delta_r)c^{(r)} = 0 \qquad (\text{all } u \in K), \tag{14}$$

implies $c^{(2)} = c^{(3)} = \ldots = c^{(r)} = 0$.

This assumption is satisfied for instance if for each l there exists an element $v^{(l)} \in K$ such that

$$v^{(l)}\delta^l \neq 0,\ v^{(l)}\delta_m = 0\ (m>l), \qquad l=2, 3, \dots, r.$$

In fact if (14) holds, let $c^{(i)}$ be the first non-zero coefficient and put $u=v^{(i)}$, then

$$0=(v^{(i)}\delta_i)c^{(i)}+0 \text{ (zero because } v^{(i)}\delta_j=0 \text{ for } j>i).$$

Now we have $v^{(i)}\delta_i \neq 0$, $c^{(i)} \neq 0$, hence a contradiction.

Because of assumption (B) relation (13) yields immediately

$$b_{(k,i)}=0 \text{ for } k=2, \dots, r; \qquad i=r+1,\ r+2, \dots . \tag{15}$$

Of course from

$$b_{(k,r+1)}=b_{(k,r+2)}=\dots=b_{(k,kr)}=0$$

it follows

$$b_{(k+1,2r+1)}=b_{(k+1,2r+2)}=\dots=b_{(k+1,kr+r)}=0,$$

so that the relations (15) are a consequence of the matrix relation

$$B^*=\{b_{(k,i)}\}=0 \qquad (k=2, \dots, r;\ i=r+1, \dots, 2r). \tag{16}$$

Now (5) becomes

$$x^k a = \sum_{i=k}^{r} a_{(k,i)}\, x^i, \qquad k=1, 2, \dots, r. \tag{17}$$

With vector $\mathbf{x}^T=(x, x^2, \dots, x^r)$ and A the triangular matrix $(a_{(k,i)})$ $(k, i=1, 2, \dots, r)$, we can rewrite (17) in the form $\mathbf{x}.\ a=A\mathbf{x}$. From $\mathbf{x}(ab)= =(\mathbf{x}a)b=A(\mathbf{x}b)=AB\mathbf{x}$, we see that if $a \to A$ and $b \to B$, then $ab \to AB$. We observe that $a \to (a_{(k,i)})=A$ is a matrix representation of K (of degree r).

By (9) the mapping $\alpha : a \to a_1$ is an endomorphism of K and $a_{(r,r)}=a\alpha^r$.

If $\alpha=\delta_1$ would be the zero endomorphism, then it easily follows from (8) that $\delta_2=\delta_3=\dots=\delta_r=0$. This contradicts assumption (B), so $\alpha \neq o$, i.e. α is necessarily a monomorphism of K. The inverse β of α is defined on $K\alpha$.

Now (17) says e.g.

$$x^r a=(a\alpha^r)x^r, \tag{18}$$

$$x^{r-1}a=a_{(r-1,r)}x^r+(a\alpha^{r-1})x^{r-1}. \tag{19}$$

Now we continue with the properties of the mappings δ_i. From $x^r(xa)= =x(x^r a)$ we obtain by equating coefficients of powers of x:

$$\delta_k\alpha^r=\alpha^r\delta_k. \tag{20}$$

Further we can derive

$$x(x^{r-1}a)=x^{r-1}(xa)=x^r a(=(a\alpha^r)x^r),$$

so that equating coefficients of $x^{r+k}(k=1, 2, \ldots)$ we conclude

(21) $$a_{(r-1,r)k} + a_{(r-1,r-1)(k+1)} = 0$$

and

(22) $$a_{k(r-1,r)} + a_{(k+1)(r-1,r-1)} = 0.$$

Now $k=1$ gives

(23) $$a_{(r-1,r)1} + a_{(r-1,r-1)2} = 0$$

and

(24) $$a_{1(r-1,r)} + a_{2(r-1,r-1)} = 0.$$

If we replace in (21) a by a_1 and use (24) we obtain

$$a_{(r,r)(k+1)} = a_{2(r-1,r-1)k},$$

so

(25) $$\alpha^r\delta_{k+1} = \delta_2\alpha^{r-1}\delta_k \qquad (k=1, 2, \ldots).$$

By (20) this can be rewritten as

(26) $$\delta_{k+1}\alpha^r = (\delta_2\alpha^{r-1})\delta_k.$$

Repeated application of (26) gives

(27) $$\delta_{k+1}\alpha^{kr} = (\delta_2\alpha^{r-1})^k\alpha,$$

so

(28) $$\delta_{k+1} = (\delta_2\alpha^{r-1})^k\beta^{kr-1} \qquad (k=0, 1, 2, \ldots).$$

In virtue of $\delta_{r+1} = (\delta_2\alpha^{r-1})^r\beta^{r^2-1} = 0$, we see $(\delta_2\alpha^{r-1})^r = 0$, thus by (10) we observe that $\Delta = \delta_2\alpha^{r-1}$ is a nilpotent (α^{r+1}, α^r)-derivation of index r. A further assumption (C) will appear important.

Assumption (C). $\delta_1 = \alpha$ is an automorphism of K, with inverse β.

We shall also introduce the mapping δ by

(29) $$\delta_2 = \delta\alpha, \text{ hence } a_2 = a\delta\alpha \text{ for all } a \in K.$$

In virtue of (10) δ satisfies the condition

(30) $$(ab)\delta = (a\delta)(b\alpha) + a(b\delta), \qquad a, b \in K,$$

in other words, δ is an α-derivation of K. The constant field C is defined as the subfield of δ-constants, so

$$C = \{a \in K / a\delta = 0\}.$$

Combining (20) and (29) we obtain after cancelling an application of α

(31) $$\delta\alpha^r = \alpha^r\delta.$$

With formula (28) we now conclude

$$\delta_{k+1}=(\delta_2\alpha^{r-1})^k\beta^{kr-1}=(\delta\alpha^r)^k\beta^{kr-1},$$

hence

$$(32)\qquad \delta_{k+1}=\delta^k\alpha \qquad (k=0,\,1,\,2,\,\ldots),$$

so

$$(33)\qquad a_1=a\alpha,\ a_2=a\delta\alpha,\ a_{k+1}=a\delta^k\alpha.$$

In virtue of $\delta_{r+1}=\delta^r\alpha=0$, we see $\delta^r=0$, i.e. δ is a nilpotent α-derivation of index r. Note that (31) is a consequence of the nilpotence of δ, cf. [9], Theorem 1. The results obtained may now be summed up as follows:

Theorem 1. Let R be the ring of "skew" polynomials in x, consisting of all polynomials $\sum a^{(i)}x^i$ (where $a^{(i)}$ belongs to the (skew) field K), with the usual addition and with a multiplication which is defined by

(A) $x\cdot a=a_1x+a_2x^2+\ldots+a_rx^r$, $r=2,\,3,\,\ldots$, then

(i) the mapping $\alpha : a\rightarrow a_1$ is an endomorphism of K;
(ii) the mapping $\delta_2 : a\rightarrow a_2$ is an $(\alpha^2,\,\alpha)$-derivation of K.
If the ring R satisfies the condition

(B) the mappings $\delta_i : a\rightarrow a_i$ $(i=2,\,3,\,\ldots,\,r)$ are right K-independent, then one has

(iii) the mapping $\delta_1=\alpha$ is a monomorphism of K (inverse β);
(iv) $\delta_{n+1}=(\delta_2\alpha^{r-1})^n\beta^{nr-1}$, $n=0,\,1,\,2,\,\ldots$;
(v) $\Delta=\delta_2\alpha^{r-1}$ is a nilpotent $(\alpha^{r+1},\,\alpha^r)$-derivation of index r.
Under the further condition

(C) α is an automorphism of K, with inverse β, one has

(vi) the mapping δ defined by $a_2=a\delta\alpha$ is an α-derivation of K;
(vii) $\delta_{n+1}=\delta^n\alpha$, $a_{n+1}=a\delta^n\alpha$ (higher α-derivations), $n=0,\,1,\,2,\,\ldots$;
(viii) $\delta^r=0$, $\delta^{r-1}\neq 0$, i.e. δ is a nilpotent α-derivation of index r. This implies e.g. $\alpha^r\delta=\delta\alpha^r$;
(ix) all the relations induced by the distributive and associative laws follow directly from the properties of the nilpotent α-derivation δ.

§ 3. Linear and quadratic units

Let $R=K[x]$ be the polynomial ring satisfying all the conditions of Theorem 1. In this section we give some examples of polynomial units in R. We can write e.g.

$$(x-a\alpha)(x+a)=(a\delta^{r-1}\alpha)x^r+(a\delta^{r-2}\alpha)x^{r-1}+\ldots+(a\delta^2\alpha)x^3+$$
$$+(a\delta\alpha+1)x^2-(a\alpha)a=-(a\alpha)a \text{ in the case } a\delta=-1.$$

We observe that $(x+a)$ with $a\delta=-1$ is a linear unit. The equation $a\delta=-1$ had been considered in [9]. We found $a=(b\delta^{r-1})^{-1}(b\delta^{r-2})$ for all $b\in K$ with $b\delta^{r-1}\neq 0$. To find quadratic units we write

$$(x+g)(ax^2+bx+c)=s\ (g,\ a,\ b,\ c,\ s\in K).$$

After long calculations we find

$$\{(g\delta\alpha+1)(g\alpha^2)^{-1}-x+g\alpha\}(x+g)=(g\alpha)g, \tag{34}$$

where $g\in K$ satisfies the relation

$$g\delta^2=(g\delta+1)(g\alpha)^{-1}\{g(\alpha\delta+\delta\alpha)+1\}. \tag{35}$$

The last equation has a solution $g=-(t\delta)^{-1}t$ for all $t\in K$ with $t\delta\neq 0$ and $t\delta^3=0$.

It is surprising that the quadratic unit in (34) can be decomposed into a linear unit and a constant term. After complicated calculations we obtain the decomposition

$$\begin{aligned}\{(g\delta\alpha+1)(g\alpha^2)^{-1}x^2-x+g\alpha\}&=\\ &=\{-(g\delta\alpha+1)(g\alpha^2)^{-1}x+(g\alpha)(1+g\delta)(g\alpha)^{-1}\}.\,(g\alpha)(1+g\delta)^{-1}.\end{aligned}$$

Of course one might be inclined to believe that all quadratic units can be written as the product of a linear unit and a constant term. However quadratic units with quadratic complementary units cannot be decomposed in general into linear units and constant terms. In § 5 we derive a general decomposition theorem on units. In the following section we make some preparations for this theorem.

§ 4. The units x^nay^n and y^nax^n

We begin with the introduction of the symbol $y=x^{-1}$ and the rather remarkable result:

Theorem 2. Let $R=K[x]$ be the polynomial ring defined by all the assumptions of Theorem 1, then the indeterminate $y=x^{-1}$ satisfies the commutation formula

$$ya=(a\beta)y-a\beta\delta \qquad \text{(for all } a\in K). \tag{36}$$

Proof. Relation (23) can be rewritten as

$$a_{(r-1,r)}\alpha+a\alpha^{r-1}\delta\alpha=0,$$

hence

$$a_{(r-1,r)}=-a\alpha^{r-1}\delta \qquad \text{(all } a\in K). \tag{37}$$

Combining (19) and (37) we obtain

$$x^{r-1}a=(a\alpha^{r-1})x^{r-1}-(a\alpha^{r-1}\delta)x^r, \tag{38}$$

also we remember

$$x^r b = (b\alpha^r)x^r, \tag{39}$$

hence

$$ya = x^{-1}a = x^{r-1}(x^{-r}a) = x^{r-1}\{(a\beta^r)x^{-r}\} = \{x^{r-1}(a\beta^r)\}x^{-r} =$$
$$= \{(a\beta)x^{r-1} - (a\beta\delta)x^r\}x^{-r} = (a\beta)y - (a\beta\delta).$$

Corollary 2.1. In order to get the general case of (3) and subjected to the conditions of Theorem 1, take any skew polynomial ring $K[y;\beta, -\beta\delta]$ with $\delta^r = 0$, adjoin $y^{-1} = x$ and take subring of polynomials in x.

Conversely given this, we have

$$ya = a\beta y - a\beta\delta = a\beta(y - \delta),$$

thus with $y = x^{-1}$

$$ax = xa\beta(1 - \delta x),$$

hence

$$xa = ax(1 - \delta x)^{-1}\alpha = ax(1 + \delta x + \delta^2 x^2 + \ldots + \delta^{r-1}x^{r-1})\alpha =$$
$$= (a\alpha)x + (a\delta\alpha)x^2 + (a\delta^2\alpha)x^3 + \ldots (a\delta^{r-1}\alpha)x^r.$$

Further we have symbolically (using again the rules $\delta x = x\delta$, $\alpha x = x\alpha$)

$$xay = a_1 + a_2 x + \ldots + a_r x^{r-1} = (a\alpha) + (a\delta\alpha)x + \ldots + (a\delta^{r-1}\alpha)x^{r-1} =$$
$$= a\{1 + \delta x + (\delta x)^2 + \ldots + (\delta x)^{r-1}\}\alpha = a(1 - \delta x)^{-1}\alpha,$$

and

$$xaby = xay \cdot xby.$$

This expresses the fact, that the mapping $a \to xay$ is an isomorphism of the field K; in fact it is the automorphism generated by the derivation δ(=infinitesimal automorphism) by the Taylor formula.

Unless stated otherwise we always assume that the ring R satisfies all the assumptions of Theorem 1 (especially Assumption (C)).

Because of $y = x^{-1}$ we have

$$(x^n a y^n)(x^n a^{-1} y^n) = 1 \text{ and } (y^n a x^n)(y^n a^{-1} x^n) = 1,$$

and it is interesting to consider the units

$$U_n(a) = x^n a y^n \qquad (n = 0, 1, 2, \ldots)$$

and

$$E_n(a) = y^n a x^n \qquad (n = 0, 1, 2, \ldots).$$

To begin with E_n we have for example (cf. (36))

$$E_1(a) = yax = a\beta - a\beta\delta x,$$

so

$$E_1(a) = -a\beta\delta\{x - (a\beta\delta)^{-1}(a\beta)\}; \tag{40}$$

$$E_2(a) = y^2ax^2 = a(\beta\delta)^2x^2 - a\beta(\beta\delta + \delta\beta)x + a\beta^2. \tag{41}$$

Further we have

$$\left\{\begin{aligned} E_{r-1}(a) &= y^{r-1}ax^{r-1} = xa\beta^r y = \\ &= a\beta^r\delta^{r-1}\alpha x^{r-1} + a\beta^r\delta^{r-2}\alpha x^{r-2} + \ldots + a\beta^r\delta\alpha x + a\beta^{r-1} = \\ &= xa\beta^r\delta + a\beta^{r-1} = \{x + a\beta^{r-1}(a\beta^r\delta)^{-1}\}\, a\beta^r\delta, \end{aligned}\right. \tag{42}$$

$$E_r(a) = y^rax^r = a\beta^r, \tag{43}$$

$$E_{r+k}(a) = y^{r+k}ax^{r+k} = y^ka\beta^rx^k = E_k(a\beta^r). \tag{44}$$

In the same way we find

$$\left\{\begin{aligned} U_1(a) &= xay = a_rx^{r-1} + a_{r-1}x^{r-2} + \ldots + a_2x + a_1 = \\ &= (a\delta^{r-1}\alpha)\, x^{r-1} + (a\delta^{r-2}\alpha)\, x^{r-2} + \ldots + (a\delta\alpha)\, x + (a\alpha) = \\ &= xa\delta + a\alpha = \{x + (a\alpha)\, (a\delta)^{-1}\}\, (a\delta), \end{aligned}\right. \tag{45}$$

$$U_2(a) = x^2ay^2 = a_{(2,r)}x^{r-2} + a_{(2,r-1)}x^{r-3} + \ldots + a_{(2,3)}x + a_{(2,2)}, \tag{46}$$

$$U_{r-1}(a) = x^{r-1}ay^{r-1} = y(a\alpha^r)x = E_1(a\alpha^r), \tag{47}$$

$$U_r(a) = x^ray^r = a\alpha^r, \tag{48}$$

$$U_{r+k}(a) = x^{r+k}ay^{r+k} = x^ka\alpha^ry^k = U_k(a\alpha^r), \tag{49}$$

$$E_n(a) = y^nax^n = x^{kr-n}(y^{kr}ax^{kr})y^{kr-n},$$

so

$$E_n(a) = U_{kr-n}(a\beta^{kr}), \tag{50}$$

for all positive integers k such that $kr - n \geqq 0$.

Also it is easy to verify the relations

$$E_1(a^{-1}\alpha)\cdot a\delta = -a^{-1}\delta\cdot U_1(a) \tag{51}$$

or

$$E_1(b)\cdot b^{-1}\beta\delta = -b\beta\delta\cdot U_1(b^{-1}\beta). \tag{52}$$

It is important to notice that all units of the form x^nay^n (where n is a fixed integer) constitute a (skew) field K_n isomorphic with the coefficient field K. Of course $K_r = K$, so that the polynomial ring R contains $(r-1)$ isomorphic copies $K_1, K_2, \ldots, K_{r-1}$ of K.

Finally we remark that the quadratic unit $U_{r-2}(a)$ cannot be written in general as the product of a linear unit and a constant term.

§ 5. The general decomposition theorem on units

Let $R=K[x]$ be the ring again satisfying all the assumptions of Theorem 1. Assume

$$[a^{(n)}x^n+a^{(n-1)}x^{n-1}+\ldots+a^{(1)}x+a^{(0)}][b^{(m)}x^m+b^{(m-1)}x^{m-1}+\ldots+ \\ +b^{(1)}x+b^{(0)}]=1,$$

in which $a^{(l)}$ does not denote a derivative, but plainly the l-th coefficient. Hence

$$[a^{(0)}y^n+a^{(1)}y^{n-1}+\ldots+a^{(n-1)}y+a^{(n)}]x^n[b^{(0)}y^m+b^{(1)}y^{m-1}+\ldots+ \\ +b^{(m-1)}y+b^{(m)}]x^m=1.$$

Let k be a positive integer such that $n\leqq kr<n+r$, then the previous relation can be rewritten as

$$(53)\qquad \begin{cases} [a^{(0)}y^n+a^{(1)}y^{n-1}+\ldots+a^{(n-1)}y+a^{(n)}]\,y^{kr-n}[b^{(0)}\alpha^{kr}y^m+ \\ \quad +b^{(1)}\alpha^{kr}y^{m-1}+\ldots+b^{(m-1)}\alpha^{kr}y+b^{(m)}\alpha^{kr}]=y^{m+kr}, \end{cases}$$

where we have $(m+kr)$ prime factors y on the right-hand side. Now it follows from a theorem of Oystein Ore ([8], Theorem 1, p. 494, or cf. Jacobson [7], Theorem 5, p. 34) that the left-hand side of (53) can also be decomposed into the product of $(m+kr)$ linear prime factors similar to y. (For the concept of similarity see the references just quoted).

Thus it follows that the polynomial

$$a^{(0)}y^n+a^{(1)}y^{n-1}+\ldots+a^{(n-1)}y+a^{(n)}$$

can be represented as a product of n linear prime factors similar to y. It is easy to prove that every polynomial similar to y can be written as $eyf(e, f\in K)$, cf. [7], p. 36.

Consequently

$$(54)\qquad a^{(0)}y^n+a^{(1)}y^{n-1}+\ldots+a^{(n-1)}y+a^{(n)}=\varrho^{(0)}y\varrho^{(1)}y\varrho^{(2)}y\ \ldots\ y\varrho^{(n-1)}y\varrho^{(n)},$$

where the $\varrho^{(i)}$ are certain elements of K. After multiplication on the right with x^n relation (54) can be rewritten in the form

$$(55)\qquad \begin{cases} a^{(n)}x^n+a^{(n-1)}x^{n-1}+\ldots+a^{(1)}x+a^{(0)}= \\ =\varrho^{(0)}(y\varrho^{(1)}x)\,(y^2\varrho^{(2)}x^2)\ \ldots\ (y^{n-1}\varrho^{(n-1)}x^{n-1})\,(y^n\varrho^{(n)}x^n)= \\ =E_0(\varrho^{(0)})\cdot E_1(\varrho^{(1)})\cdot E_2(\varrho^{(2)})\cdot\ \ldots\ \cdot E_{n-1}(\varrho^{(n-1)})\cdot E_n(\varrho^{(n)}). \end{cases}$$

By (40) and (43) we observe that in the case $r=2$ all the $E_i(\varrho^{(i)})$ are elements of K or linear units. For $r\geqq 3$ we found in the preceding section that $E_0, E_1, \ldots, E_{r-2}$ were units of degree $0, 1, 2, \ldots, r-2$ and E_{r-1} was a unit of degree $r-1$ that could be written as the product of a linear unit and an element of K (cf. (42)). The results of § 4 and § 5 may be stated as

THEOREM 3. Let R be the polynomial ring satisfying all the assumptions of Theorem 1. Let $K_n \subset R$ be the (skew) field of all units $x^n a y^n$ ($a \in K$; $n=0, 1, 2, \ldots, r-1$; $K_0 = K$). Then K_n is an isomorphic copy of K and further

(i) every unit of R can be written as a product of units from the fields K_n;
(ii) for $r>2$ every unit of R can be written as a product of units of degree $\leqq r-2$, for $r=2$ as a product of linear units;
(iii) the fields K and $K_1 = xKx^{-1}$ generate the whole ring R;
(iv) if $r=2$, then the ring R is the free product over the constant field C of the fields K and K_1;
(v) if $r>2$, then the ring R is a proper homomorphic image of the free product of the fields K and K_1 over the constant field C.

Proof. In the beginning of this section we proved (i) and (ii). Statement (iii) follows from the fact that there are polynomial units of every degree, so that an arbitrary polynomial of R can be written as a sum of units, hence as a sum of products of elements from the K_i ($i=0, 1, \ldots, r-1$). However also the elements of $K_2, K_3, \ldots, K_{r-1}$ can be written as sums of products of elements of K and K_1. To show this let $g \in K$ satisfy $g\delta = 1$ (such an element exists, cf. § 3), then we have

$$g^{(1)} = xgy = (g\alpha) + x, \text{ so } x = g^{(1)} - (g\alpha), \tag{56}$$

where $g^{(1)} \in K_1$ and $g\alpha \in K$. For an arbitrary element $a^{(m)} \in K_m$ we now have

$$a^{(m)} = x^m a y^m = \sum_{k=0}^{r-m} a_{(m,m+k)} x^k = \sum_{k=0}^{r-m} a_{(m,m+k)} \{g^{(1)} - (g\alpha)\}^k \tag{57}$$

with certain coefficients $a_{(m,m+k)} \in K$. Consequently R is generated by the fields K and K_1 and the polynomial ring R is a homomorphic image of the free product of K and K_1 over the common (constant) field C. This image may be proper or not proper. To prove statement (iv) we remark in the first place that K possesses a nilpotent α-derivation δ of index 2, hence by Theorem 3 of [9] every element $a \in K$ can be written uniquely in the forms

$$a = \pi + \varrho z = \sigma + z\tau \qquad (\pi, \varrho, \sigma, \tau \in C, z\delta = 1),$$

i.e. K/C is a left and right quadratic extension of the constant field C with basis 1 and z ($z\delta = 1$). Now statement (iv) can be proved by the remark that if R would be a proper homomorphic image of the free product of K and K_1, then R would be a field (COHN [3], Theorem 8) which is of course a contradiction, so R is the free product itself. However statement (v) can be proved by an inverse method. If R would be the free product itself of K and K_1, then every unit in R would be a monomial unit, i.e. every unit would be a product of elements of K and K_1 (COHN [2], Theorem 2.6). By (55) this is a contradiction, thus R is not the free product itself but a proper homomorphic image ($r \geqq 3$). This remark completes the proof of Theorem 3.

Corollary 3.1. In the case $r=2$ or 3 every unit can be written as a product of linear units. In the case $r=4$ every unit can be written as a product of linear and quadratic units etc.

§ 6. The prime factorization

The quadratic commutation formula ($r=2$) has been met first by P. M. Cohn ([3], p. 548) in a free product of two quadratic extensions. He proved that every left ideal was principal. It is not difficult to generalize his proof for an arbitrary r. In the proof we do not need the fact that the monomorphism α is also an automorphism of K.

Lemma 1. Let the polynomial ring R be defined by the assumptions (A) and (B) of Theorem 1, then R is a left principal ideal domain (left PID for short).

Proof. We note first that results (vi) up to (ix) of Theorem 1 cannot be proved any longer. In fact we only need formula (17). Now let $I \neq 0$ be a left ideal of R, and let

$$f = x^n + a x^{n-1} + b x^{n-2} + \dots$$

be a monic polynomial of least degree in I. Of course f is unique. Now we have $Rf \subseteq I$ and

$$(58) \qquad \begin{cases} xf = a_r x^{n+r-1} + (a_{r-1} + b_r)\, x^{n+r-2} + \dots \\ x^2 f = a_{(2,r)} x^{n+r-1} + (a_{(2,r-1)} + b_{(2,r)})\, x^{n+r-2} + \dots \\ x^3 f = a_{(3,r)} x^{n+r-1} + (a_{(3,r-1)} + b_{(3,r)})\, x^{n+r-2} + \dots \\ \quad \vdots \\ x^{r-1} f = (a_{(r-1,r)} + 1)\, x^{n+r-1} + (a_{(r-1,r-1)} + b_{(r-1,r)})\, x^{n+r-2} + \dots \end{cases}$$

and all these polynomials belong to I. We denote by R_k the set of all polynomials in R of degree $\leqq k$. With obvious abbreviations (58) can be rewritten as

$$(59) \qquad \begin{cases} xf = \gamma^{11} x^{n+r-1} + \gamma^{12} x^{n+r-2} + \dots + \gamma^{(1,r-1)} x^{n+1}, \\ x^2 f = \gamma^{21} x^{n+r-1} + \gamma^{22} x^{n+r-2} + \dots + \gamma^{(2,r-1)} x^{n+1}, \\ ---------------------------- \pmod{R_n} \\ x^{r-1} f = \gamma^{(r-1,1)} x^{n+r-1} + \gamma^{(r-1,2)} x^{n+r-2} + \dots + \gamma^{(r-1,r-1)} x^{n+1}. \end{cases}$$

First we observe that $xf, x^2 f, \dots, x^{r-1} f$ are left K-independent modulo R_n, for if

$$\varrho^{(1)} xf + \varrho^{(2)} x^2 f + \dots + \varrho^{(r-1)} x^{r-1} f = 0 \pmod{R_n},\ \varrho^{(i)} \in K,$$

then since

$$(\varrho^{(1)} x + \varrho^{(2)} x^2 + \dots + \varrho^{(r-1)} x^{r-1}) f \in Rf \subseteq I,$$

we must have

$$(\varrho^{(1)}x+\varrho^{(2)}x^2+\ldots+\varrho^{(r-1)}x^{r-1})f=\varrho^{(r)}f,$$

where $\varrho^{(r)}\in K$. But R has no zero-divisors, hence $\varrho^{(1)}=\varrho^{(2)}=\ldots=\varrho^{(r)}=0$. If there would exist a linear relation between the right-hand sides of (59), then also between the left-hand sides of (59), which is impossible because they were left K-independent. Hence the system (59) has one and only one solution (mod R_n) for x^{n+r-k}, $k=1, 2, \ldots, r-1$. E.g.

$$x^{n+r-k}=s_kf-q_n,\ s_k\in R_{r-1},\ q_n\in R_n,$$

hence

$$s_kf=x^{n+r-k}+q_n.$$

We can even obtain a relation of the form

$$(60)\qquad \bar{s}_kf=x^{n+r-k}+\bar{q}_{n-1},\ \bar{s}_k\in R_{r-1},\ \bar{q}_{n-1}\in R_{n-1},\ k=1, 2, \ldots, r-1.$$

So finally it follows that Rf contains monic polynomials of degree n, $n+1, n+2, \ldots, n+r-1$. After multiplication of those polynomials on the left by x^r we conclude that Rf contains a monic polynomial of degree N for every $N\geqq n$ and an arbitrary element h of R can be written as

$$(61)\qquad h=qf+t,\ q\in R,\ t\in R_{n-1}.$$

In particular if $h\in I$, than $t\in I$, so $t=0$ and $h=qf\in Rf$ which yields $I\subseteq Rf$. Because of $Rf\subseteq I$ we conclude $I=Rf$, which proves the lemma.

From (61) it is obvious how to develop a right Euclidean algorithm in R. If we want to divide $h_l\in R$ of degree l by $g_m\in R$ of degree m, we look in the ideal Rg_m for the monic polynomial of least degree, say f_n, so $Rg_m=Rf_n$ with $n\leqq m$. Clearly we have $f_n=ug_m$, $g_m=vf_n$, where u and v are units, so f_n and g_m are left associated. By (61) we can write h_l in the form

$$h_l=qf_n+t,\ t\in R_{n-1},$$

thus

$$(62)\qquad h_l=(qu)g_m+t,\ t\in R_{n-1}\subseteq R_{m-1}.$$

We observe that $\deg t<n\leqq m=\deg g_m$. Notice that it is possible that $l<m$.

Lemma 2. Let the polynomial ring R be defined by the assumptions (A) and (B) of Theorem 1, then R is a unique factorization domain (UFD for short).

Proof. First we show that the descending chain condition holds for left ideals having an intersection $\neq 0$ (restricted descending chain condition). Let

$$(63)\qquad R\supset Ra\supset Rb\supset Rc\supset\ldots\supset Rz$$

be a strictly descending chain of principal left ideals, which all contain a fixed element $z \neq 0$. If we rewrite the ideals by means of the corresponding minimal polynomials we obtain the sequence

$$R \supset Ra_{min} \supset Rb_{min} \supset Rc_{min} \supset \ldots \supset Rz_{min}$$

and we see that the chain (63) is finite. The ascending chain condition for left ideals follows as usual for left principal domains, thus both chain conditions hold for left ideals. The proof of [7, Theorem 5, p. 34] remains valid for the principal left ideal domain R. Hence R is a UFD.

Lemma 3. Let R be the polynomial ring satisfying all the conditions of Theorem 1, then also every right ideal is principal, i.e. R is a non-commutative principal ideal domain.

Proof. This merely depends on the fact that if the monomorphism α is an automorphism of K (with inverse β) every polynomial in R with left-hand coefficients can always be represented as a polynomial with right-hand coefficients.

From $y \cdot a = (a\beta)y - a\beta\delta$ (cf. (36)) and $y^r a = (a\beta^r)y^r$ (cf. (39)) we conclude

$$(\beta\delta)^r = 0. \tag{64}$$

The last formula follows also directly from the nilpotence of the α-derivation δ, cf. [9], Theorem 1, (iii).

Using (36) we obtain successively

$$\begin{aligned} ax = x(ya)x = x(a\beta) - x(\underbrace{a\beta\delta)x}_{}= \\ = x(a\beta) - x^2(a\beta\delta\beta) + x^2(a\beta\delta\beta\delta)x = \ldots = \\ = xa\beta - x^2 a\beta\delta\beta + x^3 a(\beta\delta)^2\beta + \ldots + x^r a(-\beta\delta)^{r-1}\beta + x^r a(-\beta\delta)^r x. \end{aligned}$$

The last term becomes zero and we have the right-hand commutation formula

$$ax = xa\beta - x^2 a(\beta\delta)\beta + x^3 a(\beta\delta)^2\beta + \ldots + x^r a(-\beta\delta)^{r-1}\beta, \tag{65}$$

and instead of the nilpotent α-derivation δ we have the derivation $\bar{\delta} = -\beta\delta$ again nilpotent of index r, satisfying $(ab)\bar{\delta} = (a\bar{\delta})b + (a\beta)(b\bar{\delta})$ for all $a, b \in K$, in other words, $\bar{\delta}$ is a $(1, \beta)$-derivation of the field K. If we apply the methods of Lemma 1 to the right-hand polynomials, then Lemma 3 follows.

Summing up we have

Theorem 4. Let $R = K[x]$ be the ring of skew polynomials in an indeterminate x over a (skew) field K with an endomorphism $\alpha : a \to a_1$, multiplication in R being defined by

$$x \cdot a = a_1 x + a_2 x^2 + \ldots + a_r x^r, \; a_i \in K.$$

If we assume that the mappings $\delta_i : a \to a_i$ $(i = 2, 3, \ldots, r)$ are right K-independent we can prove

(i) R is a principal left ideal domain with a right Euclidean algorithm;

(ii) R is a unique factorization domain;

(iii) if α is an automorphism of K, then R is a non-commutative principal ideal domain.

Department of Mathematics
Technological University Delft
Delft, The Netherlands

REFERENCES

1. Cohn, P. M., On the free product of associative rings. Math. Z. **71**, 380–398 (1959).
2. ————, On the free product of associative rings II. The case of (skew) fields. Math. Z. 73, 433–456 (1960).
3. ————, Quadratic extensions of skew fields. Proc. London Math. Soc. (3) **11**, 531–556 (1961).
4. ————, Noncommutative unique factorization domains. Trans. Am. Math. Soc. **109**, 313–331 (1963).
5. ————, Rings with a weak algorithm. Trans. Am. Math. Soc. **109**, 332–356 (1963).
6. ————, On a class of binomial extensions. Illinois J. of Math. (3) **10**, 418–424 (1966).
7. Jacobson, N., Theory of rings. Amer. Math. Soc., Providence, R.I., 1943.
8. Ore, O., Theory of non-commutative polynomials. Ann. of Math. **34**, 480–508 (1933).
9. Smits, T. H. M., Nilpotent S-derivations, Proc. Kon. Ned. Akad. Wetensch., **A 70**, 72–86 (1968).