Theory of Non-Commutative Polynomials

Oystein Ore

The Annals of Mathematics, Second Series, Volume 34, Issue 3 (Jul., 1933), 480-508.

THEORY OF NON-COMMUTATIVE POLYNOMIALS.*

By Oystein Ore.

In the present paper I have tried to give the principal results of a general non-commutative polynomial theory. The polynomials considered have coefficients in an arbitrary commutative or non-commutative field, while the multiplication of polynomials is so restricted that the degree of a product is equal to the sum of the degrees of the factors. In this form the theory contains not only the equations studied in non-commutative algebras but also most of the linear operational equations of analysis as, for instance, linear differential and difference equations.

One could have deduced this theory using the theory of moduli studied by Noether and Schmeidler.[1] One would then have to study the residue-classes of non-commutative polynomials; these form a generalized Abelian group according to the terminology introduced by Krull,[2] and there exists a correspondence between the structure of this generalized Abelian group and the corresponding polynomial such that the properties of a polynomial can be deduced from the general theorems on generalized Abelian groups. I have preferred, however, to build up the theory directly, that is, to use only the properties of the polynomials themselves as, for instance, in the ordinary polynomial theory. This seems preferable for various reasons. It makes the theory independent of the more general theory and it brings out more clearly some of the specific properties of polynomials.

In Chapter I, section 1, one finds the principal properties of the two operations *conjugation* and *differentiation* defined by means of the rules of multiplication. In section 2 the Euclid algorithm is introduced; of particular interest is the formula for the least common multiple for non-commutative multiplication; one also finds a construction of a quotient-field for general polynomials. Sections 4 and 5 are devoted to the transformation of polynomials, a very important notion which is characteristic for the non-commutative polynomials. In section 6 the connection between left-hand and right-hand divisibility properties is considered. The general theorems about the structure of non-commutative polynomials then follow in Chapter II.

* Received December 7, 1932.

[1] E. Noether and Schmeidler, *Moduln in nichtkommutativen Bereichen, insbesondere aus Differential- und Differenzenausdrücken*, Math. Zeitschrift 8 (1920), pp. 1-35.

[2] W. Krull, *Über verallgemeinerte endliche Abelsche Gruppen*, Math. Zeitschrift 23 (1925), pp. 161-196. *Theorie und Anwendung der verallgemeinerten Abelschen Gruppen.* Sitz.-Ber. Heidelberger Akademie 1926.

Chapter I.

Fundamental Properties.

1. **Multiplication.** In the following let K denote an arbitrary commutative or non-commutative field with an arbitrary characteristic p in the sense of Steinitz. The objects of our investigations are then the formal polynomials

$$F(x) = a_0 x^n + a_1 x^{n-1} + \cdots + a_n, \tag{1}$$

where the coefficients a belong to K while x is a formal variable or symbol. Let

$$G(x) = b_0 x^m + b_1 x^{m-1} + \cdots + b_m \tag{2}$$

be a second polynomial of the same kind; the sum and difference $F(x) \pm G(x)$ is defined as the polynomial one obtains from (1) and (2) by adding or subtracting corresponding coefficients. The polynomial $c_0 F(x)$, where c_0 is an arbitrary element of K, is the polynomial one obtains from $F(x)$ by multiplying all coefficients on the left with c_0. The polynomials (1) therefore form an additive Abelian group with K as domain of multipliers. When $a_0 \neq 0$ the number n is said to be the *degree* of $F(x)$. $F(x)$ is said to be reduced when $a_0 = 1$.

We shall now define multiplication for the additive group formed by the polynomials (1) so that the *group* is made a *ring*. We assume that the multiplication of polynomials shall be associative and both-sided distributive. One can obtain a definition satisfying these conditions in an infinite number of ways, but we shall here further limit the possibilities by means of the following postulate:

The degree of a product shall be equal to the sum of the degrees of the factors.

It is clear that, due to the distributive property, it suffices to define the product of two monomials $bx^r \cdot ax^s$, or even more specifically, to define the product $x \cdot a$. According to our assumption one must have

$$x \cdot a = \bar{a} x + a', \tag{3}$$

where $\bar{a}$ and a' are elements of K. We shall call $\bar{a}$ the *conjugate* and a' the *derivative* of a. It follows from our postulate that $\bar{a} = 0$ only when $a = 0$. We can introduce higher conjugates and derivatives and we shall use the notation

$$\bar{a} = a^{[1]}, \ \bar{\bar{a}} = a^{[2]}, \cdots, a^{[n]}, \cdots, \qquad a' = a', \ a'' = a^{(2)}, \cdots, a^{(n)}, \cdots.$$

From (3) one easily obtains

$$x(a+b) = (\bar{a} + \bar{b})x + a' + b',$$

$$x(ab) = \bar{a}\bar{b}x + \bar{a}b' + a'b.$$

This leads to the following rules for the conjugates and derivatives of a sum and a product:

$$(\overline{a+b}) = \bar{a} + \bar{b}, \qquad \overline{ab} = \bar{a}\,\bar{b}, \tag{4}$$

$$(a+b)' = a' + b', \qquad (ab)' = \bar{a}\,b' + a'b. \tag{5}$$

One also obtains

$$\overline{a^{-1}} = \bar{a}^{-1}, \qquad (a^{-1})' = -\bar{a}^{-1}a'a^{-1}. \tag{6}$$

It may not be superfluous to mention a few special cases which often appear in the applications. In the common non-commutative theory of polynomials one assumes that the variable x is permutable with the coefficients with the result that

$$\bar{a} = a, \qquad a' = 0.$$

In the theory of linear differential equations K is commutative and $\bar{a} = a$ so that

$$(a+b)' = a' + b', \qquad (ab)' = ab' + a'b, \tag{7}$$

and the derivative has the ordinary properties of a derivative. It is therefore possible to introduce the notion of a derivative into non-commutative fields, letting it satisfy the relations (5). We mention finally that in the theory of difference equations one puts $a' = 0$, though $\bar{a}$ usually is different from a.

The special properties of the operations $\bar{a}$ and a' will not be discussed further; this can be done by the methods of the abstract theory of fields. Only a few almost trivial facts will be mentioned.

THEOREM 1. *Through the correspondence $a \rightarrow \bar{a}$ one obtains a homeomorphism in K, the elements $\bar{a}$ form a subfield $\bar{K}$ of K.*

This follows immediately from (4), (5), and (6). The correspondence between K and the conjugate field $\bar{K}$ is a one-to-one correspondence, for from $\bar{a} = \bar{b}$, where $a \neq b$, it follows that $x(a-b)$ has the degree zero.

One also observes, using (4), (5), and (6), that those elements for which $\bar{a} = a$ form a new field $K^{(1)}$ which we shall call the invariant field of K. When a is an element of $K^{(1)}$ one obtains

$$xa = ax + a',$$

and for the elements of the invariant field one has the same rules of operation as for the ordinary linear differential polynomials.

THEOREM 2. *Those elements a for which $a' = 0$ form a field $K^{(0)}$ which we shall call the constant field of K.*

For if $a' = b' = 0$, then

$$(a \pm b)' = 0, \quad (ab)' = 0, \quad (a^{-1})' = 0.$$

For an element of the constant field, $xa = \bar{a}x$. This leads to

THEOREM 3. *The elements permutable with x form a field which is the greatest common subfield of the invariant field and the constant field.*

If one wants to determine those elements of K which are permutable with all polynomials $F(x)$ one easily finds that they form a field which is the greatest common subfield of $K^{(0)}$, $K^{(1)}$ and the center Z of K.

When one repeats the multiplication (3) one obtains

$$x^2 a = \bar{\bar{a}} x^2 + (\bar{a}' + \overline{a'}) x + a'',$$

$$x^3 a = \bar{\bar{\bar{a}}} x^3 + (\bar{\bar{a}}' + \overline{\bar{a}'} + \overline{\overline{a'}}) x^2 + (\bar{a}'' + \overline{a'}' + \overline{a''}) x + a''',$$

and, in general,

$$x^n a = S_{n,0}(a) x^n + S_{n,1}(a) x^{n-1} + \cdots + S_{n,n}(a), \tag{8}$$

where

$$S_{n,0}(a) = a^{[n]}, \quad S_{n,n}(a) = a^{(n)}, \tag{9}$$

and

$$S_{n,i}(a) = (a^{[n-i]})^{(i)} + \cdots + (a^{(i)})^{[n-i]}. \tag{10}$$

$S_{n,i}(a)$ denotes the sum of the $\binom{n}{i}$ elements which one obtains when a is conjugated $n - i$-times and differentiated i-times in an arbitrary order.

For an arbitrary product one obtains finally

$$G(x)\, F(x) = c_0 x^{n+m} + c_1 x^{n+m-1} + \cdots + c_{n+m}, \tag{11}$$

where

$$c_0 = b_0\, a_0^{[m]} \tag{12}$$

when $G(x)$ and $F(x)$ are defined by (1) and (2), and, in general,

$$c_i = \sum_{\alpha=0}^{i} b_\alpha \sum_{\beta=0}^{i-\alpha} S_{m-\alpha,\, i-\alpha-\beta}(a_\beta). \tag{13}$$

2. The Euclid algorithm.

When $F(x) = D_1(x)\, D_2(x)$ we shall call $D_2(x)$ a *right-hand* and $D_1(x)$ a *left-hand divisor* of $F(x)$. The congruence

$$F(x) \equiv G(x) \quad (\text{mod } M(x))$$

denotes that $F(x) - G(x)$ is divisible by $M(x)$ on the right, while

$$(\text{mod } M(x)) \quad F(x) \equiv G(x)$$

indicates that this difference is divisible by $M(x)$ on the left.

When $F(x)$ and $G(x)$ are given by (1) and (2), and $n \geqq m$, the difference

$$F_1(x) = F(x) - a_0 (b_0^{-1})^{[n-m]} x^{n-m}\, G(x) \tag{14}$$

is of degree lower than n. It follows that one can always perform a right-hand *division*

$$F(x) = Q(x)\, G(x) + R(x),$$

where $Q(x)$ is of degree $n - m$ and the degree of $R(x)$ does not exceed $m - 1$. One furthermore concludes that a Euclid Algorithm

$$
(15)\qquad
\begin{aligned}
F_1(x) &= Q_1(x) \quad F_2(x) \;\; + F_3(x),\\
F_2(x) &= Q_2(x) \quad F_3(x) \;\; + F_4(x),\\
&\cdots\cdots\cdots\cdots\\
F_{n-2}(x) &= Q_{n-2}(x)\; F_{n-1}(x) + F_n(x),\\
F_{n-1}(x) &= Q_{n-1}(x)\; F_n(x)
\end{aligned}
$$

exists for two arbitrary polynomials $F_1(x)$ and $F_2(x)$.

It follows from the Euclid Algorithm that there exists a unique, reduced, greatest common right-hand divisor or *cross-cut*

$$D(x) = F_n(x) = (F_1(x),\ F_2(x)),$$

where of course $F_n(x)$ is assumed to be reduced. A further consequence is

THEOREM 4. *When $C(x)$ is divisible by the cross-cut $D(x)$ of $F_1(x)$ and $F_2(x)$ one can determine polynomials $A_1(x)$ and $A_2(x)$ such that*

$$(16)\qquad A_2(x)\,F_1(x) + A_1(x)\,F_2(x) = C(x).$$

When $D(x) = 1$ we say that $F_1(x)$ is relatively prime to $F_2(x)$. From Theorem 4 one obtains

THEOREM 5. *When $F_1(x)$ is relatively prime to $F_2(x)$ the congruence*

$$(17)\qquad Y(x)\,F_1(x) \equiv C(x) \quad (\operatorname{mod} F_2(x))$$

always has a solution $Y(x)$ for an arbitrary $C(x)$.

We have here considered only right-hand divisibility; it should be observed in the following that when the type of divisibility is not specified, right-hand divisibility is always implied.

Left-hand division differs from right-hand division in that it cannot always be performed. If one wants to determine, as in (14), an element k of K such that

$$F_1(x) = F(x) - G(x)\cdot k x^{n-m}$$

is of degree less than n, then one obtains for the determination of k, according to (12), $a_0 = b_0 k^{[m]}$. From this condition it follows that one can determine k in general only when each element a of K is also an element $\bar{a}$ of $\bar{K}$.

THEOREM 6. *In general, left-hand division can be performed in K only if the correspondence $a \to \bar{a}$ is an automorphism of K.*

When this condition is satisfied there exist a left-hand Euclid Algorithm, a unique left-hand cross-cut, and theorems analogous to Theorems 4 and 5.

We point out that every polynomial

$$H(x) = x^n b_0 + x^{n-1} b_1 + \cdots + x\, b_{n-1} + b_n$$

with the coefficients as right-hand multipliers can be written according to the definition of multiplication as

$$H(x) = B_0 x^n + B_1 x^{n-1} + \cdots + B_{n-1} x + B_n$$

with left-hand coefficients. According to (8) one obtains the following expressions for these coefficients:

$$B_i = S_{n,i}(b_0) + S_{n-1,i-1}(b_1) + \cdots + S_{n-i,0}(b_i). \tag{18}$$

One can ask, however, whether inversely every polynomial with left-hand coefficients can always be represented as a polynomial with right-hand coefficients. One directly obtains

THEOREM 7. *The necessary and sufficient condition that every right-hand polynomial be also a left-hand polynomial and every left-hand polynomial be also a right-hand polynomial is that* $K = \bar{K}$.

If one denotes the inverse conjugate of a by $a^{[-1]}$ it follows that $\overline{a^{[-1]}} = a$,

$$a x = x a^{[-1]} - (a^{[-1]})', \tag{19}$$

and, in general,

$$a x^n = x^n T_{n,0}(a) - x^{n-1} T_{n-1,1}(a) + \cdots + (-1)^n T_{0,n}(a),$$

where the $T_{n,i}(a)$ are constructed from the inverse conjugates and derivatives in a way similar to that by which the $S_{n,i}(a)$ were obtained from the conjugates and derivatives of a.

3. **Union and quotient field.** Let $A(x)$ and $B(x)$ be arbitrary polynomials; the reduced polynomial $M(x)$ of lowest degree which is right-hand divisible by both $A(x)$ and $B(x)$ will be called the right-hand union of $A(x)$ and $B(x)$ and will be denoted by

$$M(x) = [A(x), B(x)] = A_1(x) \cdot B(x) = B_1(x) \cdot A(x). \tag{20}$$

It is obvious that if the union exists it must be uniquely determined.

We shall now show that the existence of a Euclid algorithm implies the existence of not only the cross-cut but also the union of two polynomials. We shall also show that the Euclid algorithm gives an explicit formula for the union; this result holds for arbitrary non-commutative domains with a Euclid algorithm, but the formula for the union does not seem to have been observed even in special cases. We want to prove

THEOREM 8. *When a Euclid algorithm of* $F_1(x)$ *and* $F_2(x)$ *of the form* (15) *is given, then the union of these polynomials is given by the formula*

$$\begin{aligned}[F_1(x), F_2(x)] = a F_{n-1}(x) F_n(x)^{-1} F_{n-2}(x) F_{n-1}(x)^{-1} \cdots \\ \cdots F_3(x) F_4(x)^{-1} F_2(x) F_3(x)^{-1} F_1(x).\end{aligned} \tag{21}$$

The constant a must be chosen so that the union is reduced.

In order to prove Theorem 8 we first deduce a necessary property of the union.

THEOREM 9. *When*

$$A(x) \equiv B(x) \pmod{C(x)} \tag{22}$$

then

$$[A(x),\ C(x)] = k[B(x),\ C(x)]\cdot B(x)^{-1}\cdot A(x). \tag{23}$$

If the $H(x)$ of $A(x)$ and $C(x)$ exists, then one must have $H(x) = C_1(x)A(x)$, and $C_1(x)$ is the polynomial of lowest degree such that this product is also divisible by $C(x)$. According to (22) one has

$$A(x) = B(x) + Q(x)\,C(x)$$

and one obtains

$$H(x) = C_1(x)\,A(x) = C_1(x)\,B(x) + C_1(x)\,Q(x)\,C(x),$$

and the product $C_1(x)\,B(x)$ must be divisible by $C(x)$.

The lowest degree of $C_1(x)$ is therefore obtained when

$$C_1(x)\,B(x) = [B(x),\ C(x)],$$

and this relation obviously gives (22).

When the formula (23) is repeatedly applied to the union $[F_1(x),\ F_2(x)]$, it follows according to (15) that if the union exists it must have the form (21). It therefore remains only to show that the right-hand side of (21) is a polynomial which is divisible by both $F_1(x)$ and $F_2(x)$.

We prove this by induction; let

$$\varphi_i(x) = F_{n-1}(x)\,F_n(x)^{-1}\,F_{n-2}(x)\cdots F_i(x).$$

It follows from the last equation (15) that $\varphi_{n-1}(x) = F_{n-1}(x)$ is integral and divisible by both $F_{n-1}(x)$ and $F_n(x)$. Let us therefore assume it has already been shown that $\varphi_{i+1}(x)$ is integral and divisible by $F_{i+1}(x)$ and $F_{i+2}(x)$. It is then obvious that

$$\varphi_i(x) = \varphi_{i+1}(x)\,F_{i+2}(x)^{-1}\,F_i(x)$$

is integral and divisible by $F_i(x)$. In order to prove the divisibility by $F_{i+1}(x)$ we observe that

$$F_i(x) = Q_i(x)\,F_{i+1}(x) + F_{i+2}(x)$$

and consequently that

$$\varphi_i(x) = \varphi_{i+1}(x)\,F_{i+2}(x)^{-1}\,Q_i(x)\,F_{i+1}(x) + \varphi_{i+1}(x).$$

This shows the divisibility by $F_{i+1}(x)$.

From (21) it follows furthermore that when the degrees of $F_1(x)$ and $F_2(x)$ are n_1 and n_2 respectively, the cross-cut being of degree d, the union must be of degree $n_1 + n_2 - d$.

In the following let $\delta_{A,B}$ denote the degree of the cross-cut $(A(x),\ B(x))$ and let $\mu_{A,B}$ denote the degree of the union; we have shown then that

$$\alpha + \beta = \mu_{A,B} + \delta_{A,B}, \tag{24}$$

where α and β are the degrees of $A(x)$ and $B(x)$.

For an arbitrary number of polynomials one defines the union

$$[A(x), B(x), C(x), \cdots]$$

in a corresponding manner. Among the more important properties we observe that

$$(25) \qquad [A(x), [B(x), C(x)]] = [[A(x), B(x)], C(x)]$$

and that

$$(26) \qquad [A(x), B(x), D(x)] = [A(x), B(x)],$$

where $D(x)$ is a divisor of $[A(x), B(x)]$. Further,

$$(27) \qquad d_0^{[\mu_{A,B}]}[A(x)D(x), B(x)D(x)] = [A(x), B(x)]D(x),$$

where d_0 is the highest coefficient in $D(x)$ and $\mu_{A,B}$ is the degree of $[A(x), B(x)]$.

When left-hand division can be performed, that is, when $K = \bar{K}$ in accordance with Theorem 6, it follows in the same way that also a left-hand union exists and has analogous properties.

It should be mentioned that it is possible to enlarge the ring of polynomials considered here to a non-commutative field through the introduction of formal quotients. This quotient field then corresponds in the commutative case to the field of all rational functions with coefficients in the field K.

The existence of the quotient-field is a consequence of the existence of a union.[3] If we define the quotient as

$$Q(x) = \frac{B(x)}{A(x)} = A(x)^{-1}B(x),$$

we can define the sum of two quotients as

$$\frac{B(x)}{A(x)} + \frac{D(x)}{C(x)} = \frac{C_1(x)B(x) + A_1(x)D(x)}{M(x)},$$

where

$$M(x) = [A(x), C(x)] = A_1(x)C(x) = C_1(x)A(x)$$

is the right-hand union of $A(x)$ and $C(x)$. We define the product by

$$\frac{B(x)}{A(x)} \cdot \frac{D(x)}{C(x)} = \frac{B_1(x)D(x)}{C_1(x)A(x)},$$

where

$$[B(x), C(x)] = B_1(x)C(x) = C_1(x)B(x).$$

Finally,

$$\left(\frac{B(x)}{A(x)}\right)^{-1} = \frac{A(x)}{B(x)}.$$

One easily shows that addition is commutative and multiplication is associative and distributive on both sides.

[3] O. Ore, *Linear equations in non-commutative fields*, Annals of Mathematics 32 (1931), pp. 463–477.

4. **Transformation.** Having established the existence of a union we can proceed to introduce the fundamental notion of *transformation* of a polynomial.

The polynomial

$$A_1(x) = a_0 b_0^{[\alpha-\delta_{A,B}]} \, [A(x),\; B(x)] \cdot B(x)^{-1} \tag{28}$$

is called the transform of $A(x)$ by $B(x)$ and will be denoted by

$$A_1(x) = BA(x)B^{-1}. \tag{29}$$

Here α denotes the degree of $A(x)$. The transform $A_1(x)$ is of degree $\alpha - \delta_{A,B}$; consequently, when $A(x)$ and $B(x)$ are relatively prime, $A(x)$ and $A_1(x)$ are of the same degree. The multiplicative constant in (28) is chosen so that the highest coefficient in $A_1(x)$ is equal to a_0. It might have been possible to choose this constant in a different way or even to omit it, but after several trials I have found it most satisfactory to use the form given here.

When $A(x)$ is relatively prime to $B(x)$, and $A_1(x)$ is consequently of the same degree as $A(x)$, we shall call the transformation (29) a special transformation and $A_1(x)$ is said to be of *the same kind* or *similar* to $A(x)$. When, however, $A(x)$ and $B(x)$ have a common factor, we shall call the transformation *general*.[4] In both cases one obtains, according to (28) and (29),

$$BA(x)B^{-1}B(x) = a_0 b_0^{[\alpha-\delta_{A,B}]} \, [A(x),\; B(x)]. \tag{30}$$

We shall now deduce various properties of the transform, and we commence with the following theorem:

THEOREM 10. *When*

$$A(x) \equiv B(x) \pmod{C(x)}$$

then

$$AC(x)A^{-1} = BC(x)B^{-1}.$$

In order to prove this theorem we have to go back to the identity (23) in Theorem 9. Since here both sides are reduced, one easily obtains for the factor k the value

$$k = (a_0^{-1})^{[\gamma-\delta_{B,C}]} \, b_0^{[\gamma-\delta_{B,C}]}$$

and (23) takes the form

$$a_0^{[\gamma-\delta_{B,C}]} \, [A(x), C(x)] \cdot A(x)^{-1} = b_0^{[\gamma-\delta_{B,C}]} \, [B(x), C(x)] \; B(x)^{-1}.$$

Since $\delta_{B,C} = \delta_{A,C}$, according to assumption, our theorem follows.

The next theorem is particularly important in its applications.

[4] The general transformation can always be reduced to the special. When $D(x)$ is the cross-cut of $A(x) = A_1(x)D(x)$ and $B(x) = B_1(x)D(x)$, then it follows from the definition of the transformation that

$$AB(x)A^{-1} = A_1B_1(x)A_1^{-1}.$$

THEOREM 11. *When a product $A(x)\,B(x)$ is divisible by $C(x)$, then $A(x)$ is divisible by $B\,C(x)\,B^{-1}$.*

Since $A(x)\,B(x)$ is divisible both by $B(x)$ and $C(x)$, the product must also be divisible by the union of these polynomials, and it is therefore possible to find a polynomial $K(x)$ such that

$$A(x)\,B(x) = K(x)\,c_0\,b_0^{[\gamma-\delta_{C,B}]}\,[B(x), C(x)].$$

Division by $B(x)$ gives

$$A(x) = K(x)\,B\,C(x)\,B^{-1}.$$

We shall usually apply Theorem 11, as in the commutative theory of polynomials, to the case where $C(x)$ is relatively prime to $B(x)$. The more general formulation 11 seems, however, to command some interest.

THEOREM 12. *When the transformer is a product, one can transform by the factors in order from right to left, that is,*

$$(31) \qquad (CB)\,A(x)\,(CB)^{-1} = C(B\,A(x)\,B^{-1})\,C^{-1}.$$

From the identity

$$[A(x), C(x)\,B(x)] = [A(x), B(x), C(x)\,B(x)]$$

one obtains, applying (27),

$$(CB)\,A(x)\,(CB)^{-1} = a_0\,c_0^{[\alpha-\delta_{A,CB}]}\,[B\,A(x)\,B^{-1}, C(x)]\cdot C(x)^{-1},$$

and since

$$\delta_{A,CB} = \delta_{A,B} + \delta_{BAB^{-1},C},$$

one has obtained the theorem. The associative law for the transformation can also easily be shown to hold.

The symmetry of the notion of similarity follows from

THEOREM 13. *When $A_1(x)$ is similar to $A(x)$, then $A(x)$ is similar to $A_1(x)$, that is, when*

$$A_1(x) = B\,A(x)\,B^{-1},$$

where $B(x)$ is relatively prime to $A(x)$, then one can determine $B_1(x)$ so that

$$A(x) = B_1\,A_1(x)\,B_1^{-1}.$$

According to Theorems 10 and 12 it is sufficient to choose $B_1(x)$ such that

$$B_1(x)\,B(x) \equiv 1 \pmod{A(x)},$$

and this always possible according to Theorem 5.

We also see that Theorem 5 can be stated in the following more complete form:

THEOREM 14. *When the congruence*

$$X(x)\,B(x) \equiv C(x) \pmod{A(x)},$$

has a solution $X_0(x)$, the most general solution has the form

$$X(x) = X_0(x) + K(x) B A(x) B^{-1},$$

where $K(x)$ is an arbitrary polynomial.

It follows that

$$(X(x) - X_0(x)) B(x) \equiv 0 \pmod{A(x)},$$

and the theorem follows from Theorem 11.

THEOREM 15. *The transform of a union is equal to the union of the transforms of the components, that is,*

$$C[A(x), B(x)] \cdot C^{-1} = [CA(x) C^{-1}, CB(x) C^{-1}]. \tag{32}$$

The theorem follows immediately when one multiplies both sides by $C(x)$.

Finally it should be mentioned that the corresponding theorem does not hold for the cross-cut. One sees, however, that in any case

$$(CA(x) C^{-1}, CB(x) C^{-1}) \equiv 0 \pmod{C(A(x), B(x)) C^{-1}}, \tag{33}$$

and one concludes that, when $C(x)$ is relatively prime to the union $[A(x), B(x)]$, then

$$(CA(x) C^{-1}, CB(x) C^{-1}) = C(A(x), B(x)) C^{-1}. \tag{34}$$

5. **Transformation of a product.** We shall finally examine the transformation af a product; the results are in this case not quite as simple as for the preceding theorems on transformation. The complications are due chiefly to the multiplicative constant in the definition (28) of the transform. It may be possible to choose this constant in a different way, making the transformation of a product more symmetric, but after various attempts I have found that the former theorems then become correspondingly more complicated.

Since one has for an arbitrary product

$$B(x) A(x) = [B(x) A(x),\ A(x)]$$

one obtains according to Theorem 5

$$C(B(x) A(x)) C^{-1} \equiv 0 \pmod{CA(x) C^{-1}}. \tag{35}$$

It should be observed that the congruence (33) is a consequence of (35). According to (35) one can always write

$$C(B(x) A(x)) C^{-1} = K(x) CA(x) C^{-1}, \tag{36}$$

and the problem is to determine $K(x)$. We shall simplify the formulas by assuming that the transformation is a special one in that $C(x)$ is relatively prime to $B(x) A(x)$. The same method is, however, applicable even in the most general case.

From (36) it follows by multiplication with $C(x)$ that

$$b_0\, a_0^{[\beta]}\, c_0^{[\alpha+\beta]}\, [B(x)\, A(x),\, C(x)] = K(x)\, a_0\, c_0^{[\alpha]}\, [A(x),\, C(x)].$$

Division by $A(x)$ gives

$$b_0\, a_0^{[\beta]}\, c_0^{[\alpha+\beta]}\, (a_0^{-1})^{[\beta+\gamma]}\, [B(x),\, A\, C(x)\, A^{-1}] = K(x)\, a_0\, c_0^{[\alpha]}\, (a_0^{-1})^{[\gamma]}\, c_0^{-1}\, A\, C(x)\, A^{-1}.$$

Another division by $C_1(x) = A\, C(x)\, A^{-1}$ gives

$$(37)\; K(x) = b_0\, a_0^{[\beta]}\, c_0^{[\alpha+\beta]}\, (a_0^{-1})^{[\beta+\gamma]}\, (c_0^{-1})^{[\beta]}\, b_0^{-1}\, C_1 B(x) C_1^{-1}\, c_0\, a_0^{[\gamma]}\, (c_0^{-1})^{[\alpha]}\, a_0^{-1}.$$

One should now observe that, when an arbitrary polynomial $E(x)$ of degree ϵ with highest coefficient e_0 is transformed by an element k_0 in K, one obtains the polynomial

$$e_0\, k_0^{[\epsilon]}\, e_0^{-1}\, E(x)\, k_0^{-1}.$$

It is then seen that (37) can be written in the simpler form

$$(38) \qquad K(x) = C_2\, B(x)\, C_2^{-1},$$

where

$$(39) \qquad C_2(x) = a_0\, c_0^{[\alpha]}\, (a_0^{-1})^{[\gamma]}\, c_0^{-1}\, A\, C(x)\, A^{-1}.$$

This proves

THEOREM 16. *When $C(x)$ is relatively prime to $B(x)\, A(x)$ then*

$$(40) \qquad C(B(x)\, A(x))\, C^{-1} = C_2\, B(x)\, C_2^{-1} \cdot C\, A(x)\, C^{-1},$$

where $C_2(x)$ is determined by (39).

The theorem can be extended to an arbitrary number of factors and then gives

$$(41)\; C(A_n(x) \cdots A_1(x))\, C^{-1} = C_n\, A_n(x)\, C_n^{-1} \cdots C_2\, A_2(x)\, C_2^{-1}\, C\, A_1(x)\, C^{-1},$$

where the transformation, as before, is a special one. *The factors of the transform are similar to the factors of the original product.*

If one assumes that $A(x)$ and $B(x)$ are reduced, or in general, that all $A(x)$ in (41) are reduced, then similar formulas hold also for a general transformation.

6. **Left-hand transformation.** Having thus deduced the main properties of the right-hand transformation, we shall briefly discuss the left-hand properties. It has already been mentioned in section 3 that if $\bar{K}$ is identical with K, there exists a left-hand Euclid algorithm and consequently a left-hand union

$$H_l(x) = [A(x),\, B(x)]_l$$

for two arbitrary polynomials $A(x)$ and $B(x)$. The left-hand union is of course defined as the reduced polynomial of lowest degree which is left-hand divisible by both $A(x)$ and $B(x)$. The degree of $H_l(x)$ is $\alpha + \beta - \delta_{A,B}$, where $\delta_{A,B}$ is the degree of the left-hand cross-cut $(A(x), B(x))_l$ of $A(x)$ and $B(x)$.

The left-hand transformation can now be defined by putting

$$B^{-1}A(x)_l B = B(x)^{-1}\cdot[A(x), B(x)]_l\cdot b_0^{[-\alpha-\beta+\delta_{A,B}]}\ a_0^{[-\alpha-\delta_{A,B}]},$$

where the constants are chosen so that the transform of $A(x)$ has the highest coefficient a_0. One then proves without great difficulty the corresponding theorems for left-hand transformation as before for right-hand transformation.

In many cases however, the left-hand notions can be replaced by the corresponding right-hand notions. We first prove

THEOREM 17. *When $A_1(x)$ and $B_1(x)$ are left-hand relatively prime, then the left-hand union of these polynomials can also be considered as a right-hand union*

$$[A_1(x), B_1(x)]_l = [A(x), B(x)] \tag{42}$$

and this can be done in such a way that $A_1(x)$ is right-hand similar to $A(x)$ and $B_1(x)$ right-hand similar to $B(x)$.

According to the definition of the left-hand union

$$[A_1(x), B_1(x)]_l = A_1(x)B_2(x) = B_1(x)A_2(x), \tag{43}$$

and here $A_2(x)$ is right-hand relatively prime to $B_2(x)$, for if they had a common factor, one could find a polynomial of lower degree that would be left-hand divisible by both $A_1(x)$ and $B_1(x)$. From Theorem 11 and (43) it follows that

$$A_1(x) = K(x)\cdot B_2A_2(x)B_2^{-1}, \qquad B_1(x) = L(x)\cdot A_2B_2(x)A_2^{-1}. \tag{44}$$

If α is the degree of $A_1(x)$ and β the degree of $B_1(x)$, then according to (43) $A_2(x)$ is of degree α and $B_2(x)$ is of degree β; one finds furthermore for the highest coefficients in $A_2(x)$ and $B_2(x)$

$$a_2 = (b_1^{-1})^{[-\beta]}, \qquad b_2 = (a_1^{-1})^{[-\alpha]},$$

where a_1 and b_1 are the highest coefficients of $A_1(x)$ and $B_1(x)$. If one substitutes the expressions (44) into (43), one can divide on the right-hand side by $[A_2(x), B_2(x)]$, getting

$$K(x)\,(b_1^{-1})^{[-\beta]}\,a_1^{-1} = L(x)\,(a_1^{-1})^{[-\alpha]}\,b_1^{-1}.$$

Since (44) shows that $K(x)$ and $L(x)$ are left-hand relatively prime, this identity can hold only when $K(x)$ and $L(x)$ are both of degree zero. It follows then from (44) that

$$K(x) = a_1 b_1^{[-\beta]}, \qquad L(x) = b_1 a_1^{[-\alpha]}.$$

The relation (44) also gives

$$(45)\qquad \begin{aligned} A_1(x) &= a_1 b_1^{[-\beta]} B_2 A_2(x) B_2^{-1} = B_2 A(x) B_2^{-1}, \\ B_1(x) &= b_1 a_1^{[-\alpha]} A_2 B_2(x) A_2^{-1} = A_2 B(x) A_2^{-1}, \end{aligned}$$

where one has put

$$(46)\qquad A(x) = a_1 b_1^{[-\beta]} A_2(x), \qquad B(x) = b_1 a_1^{[-\alpha]} B_2(x).$$

Then one has, however, according to (43)

$$[A_1(x), B_1(x)]_l = [A_2(x), B_2(x)] = [A(x), B(x)],$$

and the theorem is demonstrated.

From the definition of left-hand transform one deduces the notion of *left-hand similarity*: Two polynomials $A_1(x)$ and $A(x)$ are said to be left-hand similar if $A_1(x)$ is obtained from $A(x)$ by transformation by some polynomial $B(x)$ left-hand relatively prime to $A(x)$. Left-hand similarity offers, however, little new, as the following theorem shows:

THEOREM 18. *When two polynomials are left-hand similar they are also right-hand similar.*

One obtains, according to (43) for the left-hand transform of an arbitrary polynomial $A(x)$ by another $B_1(x)$ left-hand relatively prime to $A_1(x)$,

$$(47)\qquad \begin{aligned} H(x) = B_1^{-1} A_1(x)_l B_1 &= A_2(x) b_1^{[-\alpha-\beta]} a_1^{[-\alpha]} \\ &= (b_1^{-1})^{[-\beta]} a_1^{-1} A(x) b_1^{[-\alpha-\beta]} a_1^{[-\alpha]}, \end{aligned}$$

where the notation is as above. The theorem to be proved then asserts that $H(x)$ can also be obtained from $A_1(x)$ through right-hand transformation. According to (45) $A_1(x)$ is right-hand similar to $A(x)$ and one obtains $H(x)$ from $A(x)$ by right-hand transformation with the element

$$(a_1^{-1})^{[-\alpha]} (b_1^{-1})^{[-\alpha-\beta]}.$$

CHAPTER II.

The Theorems of Decomposition.

1. **First theorem.** A reduced polynomial $P(x)$ is called a *prime* polynomial when $P(x)$ has no reduced factors aside from constants and $P(x)$ itself.

Among the properties of prime polynomials we observe that

Every polynomial which is similar to a prime polynomial is a prime polynomial.

This follows from Theorems 13 and 16. One deduces furthermore from Theorem 11 the following:

When a product $A(x)\,B(x)$ is right-hand divisible by a prime polynomial $P(x)$, and $B(x)$ is not divisible by $P(x)$, then $A(x)$ is divisible by the prime polynomial $B\,P(x)\,B^{-1}$.

Let us now consider the decomposition of an arbitrary polynomial into prime factors. To describe the following results more completely we shall need the notion of *interchangeability:*

A polynomial $A(x)$ is said to be interchangeable with a second polynomial $B(x)$ if one can determine a polynomial $A_1(x)$ similar to $A(x)$ such that

$$A(x) = B A_1(x)\, B^{-1}. \tag{1}$$

It is hardly necessary to mention that the notion of interchangeability is not symmetric.

When the identity (1) holds, where $A(x)$ and $B(x)$ are reduced, then the product $A(x)\,B(x)$ is a union

$$A(x)\,B(x) = [A_1(x), B(x)] = A_1\, B(x)\, A_1^{-1}\cdot A_1(x). \tag{2}$$

One has therefore in this case two different decompositions of the product $A(x)\,B(x)$ wherein the factors are similar but occur in inverse order. We shall say that the second decomposition (2) is obtained from the first through *interchange of factors.*

The first decomposition theorem then takes the following form:

THEOREM 1. *Every reduced polynomial has a representation as the product of prime factors. Two different decompositions of the same polynomial have the same number of prime factors and the factors are similar in pairs. One decomposition can be obtained from the other through interchanges of factors.*

It is obvious that every polynomial possesses at least one decomposition into prime factors. Let

$$F(x) = P_r(x)\cdots P_2(x)\,P_1(x) = Q_s(x)\cdots Q_2(x)\,Q_1(x) \tag{3}$$

be two different decompositions. The prime polynomial $Q_1(x)$ then divides the left-hand side of (3). If $P_1(x) = Q_1(x)$, this factor can be cancelled. If $P_1(x) \neq Q_1(x)$, let k be the first number such that the product $P_k(x)\cdots P_1(x)$ is divisible by $Q_1(x)$. The product $P_{k-1}(x)\cdots P_1(x)$ is then relatively prime to $Q_1(x)$ and according to Theorem 11, Chapter 1, $P_k(x)$ is then divisible by the prime polynomial $Q_1'(x)$ obtained from $Q_1(x)$ by transformation with $P_{k-1}(x)\cdots P_1(x)$. Since $P_k(x)$ is a prime polynomial one obtains

$$P_k(x) = (P_{k-1}\cdots P_1)\,Q_1(x)\,(P_{k-1}\cdots P_1)^{-1}. \tag{4}$$

Then, however, according to the definition (1), $P_k(x)$ is interchangeable with $P_{k-1}(x)\cdots P_1(x)$ and one obtains according to (2) and (4)

$$P_k(x)\,P_{k-1}(x)\cdots P_1(x) = Q_1\,(P_{k-1}(x)\cdots P_1(x))\;Q_1^{-1}\cdot Q_1(x). \tag{5}$$

We have shown by interchange of factors that also the first decomposition (3) contains $Q_1(x)$ as a right-hand factor, so that this factor may be cancelled. $P_k(x)$ and $Q_1(x)$ must be similar according to (4), and when one performs the transformation

$$Q_1(P_{k-1}(x)\cdots P_1(x))\,Q_1^{-1}$$

one obtains, according to Theorem 16, Chapter 1, a product of prime factors which are similar to the $P_{k-1}(x), \cdots, P_1(x)$. After division by $Q_1(x)$ the remaining products in (3) can be treated in the same way and the theorem follows.

The number of prime factors which occur in an arbitrary prime factor decomposition of a polynomial $F(x)$ will sometimes be called the *length* of $F(x)$.

We point out as a consequence of Theorem 16, Chapter 1, that similar polynomials must have decompositions of the same length and with similar factors.

It may not be superfluous to mention that the number of factors of a polynomial is in general not limited by the degree of the polynomial; it may even happen, as for instance in the theory of linear differential polynomials, that the number of divisors is infinite.

2. **Completely reducible polynomials.** A polynomial is said to be *completely reducible* when it is representable as the union of a finite or infinite number of prime polynomials.

THEOREM 2. *A completely reducible polynomial $F(x)$ can always be represented in the form*

$$F(x) = [P_1(x),\ P_2(x),\ \cdots,\ P_r(x)], \tag{6}$$

where the $P_i(x)$ are prime polynomials such that none of them divides the union of the others.

Let $P_1(x)$ be an arbitrary prime divisor of $F(x)$; if then $F(x) \neq P_1(x)$, let $P_2(x)$ be a second prime divisor. The union $[P_1(x), P_2(x)]$ is then a divisor of $F(x)$, and if it is not equal to $F(x)$, let $P_3(x)$ be a prime divisor of $F(x)$ not dividing $[P_1(x), P_2(x)]$. Then $[P_1(x), P_2(x), P_3(x)]$ is a divisor of $F(x)$. This procedure can be continued, and one must finally obtain a representation (6) since the degree of $F(x)$ is finite.

When a completely reducible polynomial (6) is transformed by an arbitrary polynomial $H(x)$, one obtains according to Theorem 15, Chapter 1,

$$HF(x)\,H^{-1} = [HP_1(x)\,H^{-1},\ \cdots,\ HP_r(x)\,H^{-1}]. \tag{7}$$

The transform of a completely reducible polynomial is again completely reducible.

For two similar completely reducible polynomials one can choose the basis representations so that the prime polynomials of the bases are similar in pairs.

One obtains the prime factor representation of a completely reducible polynomial (6) by writing

$$F(x) = [P_1 P_2(x) P_1^{-1}, \cdots, P_1 P_r(x) P_1^{-1}] P_1(x)$$

and repeatedly applying the same process to the perfectly reducible factor

$$[P_1 P_2(x) P_1^{-1}, \cdots, P_1 P_r(x) P_1^{-1}].$$

It is obvious that the length of the decomposition is r and that the factors are similar to the $P_i(x)$.

One can also characterize the completely reducible polynomials by:

THEOREM 3. *The necessary and sufficient condition that a polynomial be completely reducible is that two arbitrary prime factors in an arbitrary decomposition be interchangeable.*

A completely reducible polynomial may therefore also be termed a completely interchangeable polynomial.

We prove first that a completely interchangeable polynomial is also completely reducible. This is obvious when the length of the decomposition does not exceed 2, for when

$$F(x) = P_1(x) P_2(x)$$

then, according to assumption, $P_1(x) = P_2 \bar{P}_1(x) P_2^{-1}$, where $\bar{P}_1(x)$ also denotes a prime polynomial. Hence

$$F(x) = [\bar{P}_1(x), P_2(x)].$$

The general proof follows by induction. We write

$$F(x) = \mathrm{P}_r(x) \cdots P_2(x) P_1(x) = F_1(x) P_1(x), \tag{8}$$

where $F_1(x)$ contains a smaller number of prime factors and, since it is completely interchangeable, it is also completely reducible according to assumption. Let

$$F_1(x) = [\bar{P}_r(x), \cdots, \bar{P}_2(x)]. \tag{9}$$

Since all prime polynomials $\bar{P}_i(x)$ $(i = 2, \cdots, r)$ may appear as right-hand factors of $\mathrm{F}_1(x)$, they are all interchangeable with $P_1(x)$, so that

$$\bar{P}_i(x) = P_1 Q_i(x) P_1^{-1}, \qquad (i = 2, \cdots, r),$$

where $Q_i(x)$ denote prime polynomials. From (7) and (9) one obtains

$$F_1(x) = P_1[Q_r(x), \cdots, Q_2(x)]\,P_1^{-1}, \tag{10}$$

and from (8),

$$F(x) = [Q_r(x), \cdots, Q_2(x), P_1(x)]. \tag{11}$$

The converse part of the theorem also follows by induction. When $F(x) = P_1(x)\,P_2(x)$ is completely reducible, then $F(x)$ must be divisible by a second prime function $Q_1(x)$ other than $P_1(x)$, and one deduces from Theorem 11, Chapter 1, that $P_1(x) = P_2\,Q_1(x)\,P_2^{-1}$, so that $P_1(x)$ is interchangeable with $P_2(x)$. Now let (8) be a representation of a completely reducible polynomial as a product of prime polynomials. Then $F_1(x)$ has the form (10), it is completely reducible, and therefore, by assumption, it is completely interchangeable. It remains only to show that $P_2(x)$ is always interchangeable with $P_1(x)$. This follows easily from the fact that every prime divisor of $F_1(x)$ must be of the form $P_1\,Q(x)\,P_1^{-1}$.

From Theorem 3 one derives:

THEOREM 4. *Every divisor of a completely reducible polynomial is completely reducible, and the basis of a divisor can be completed to a basis for $F(x)$.*

The second part of the theorem follows by the method of construction of a basis. One can also show somewhat more generally that every factor occurring anywhere in a product decomposition of a completely reducible polynomial must also be completely reducible.

Let us say that a completely reducible polynomial $F(x)$ is *uniform* when it is the union of similar prime polynomials so that it may be expressed as

$$F(x) = [A_1\,P(x)\,A_1^{-1}, \cdots, A_r\,P(x)\,A_r^{-1}].$$

Every prime divisor of $F(x)$ is then similar to $P(x)$. The union of all similar prime polynomials which divide a given completely reducible polynomial (6) shall be called a *maximal uniform* divisor. One can then prove

THEOREM 5. *Every completely reducible polynomial is uniquely representable as the union of its maximal uniform divisors.*

It should be mentioned finally that the theory of left-hand completely reducible polynomials is quite analogous. A polynomial is said to be *left-hand completely reducible* when it is the *left union* of prime polynomials. One easily proves

THEOREM 6. *Every left-hand completely reducible polynomial is also right-hand completely reducible and conversely.*

One also finds that the left-hand and the right-hand basis representations contain the same number of prime polynomials and that these are similar in pairs.

3. Second decomposition theorem.[5] Let $F(x)$ be an arbitrary polynomial; the union $H_1(x)$ of all prime functions $P(x)$ dividing $F(x)$ from the right shall be called the *maximal completely reducible* divisor of $F(x)$. This name is justified by the fact that every completely reducible divisor of $F(x)$ must also be a divisor of $H_1(x)$. $H_1(x)$ is uniquely determined and one can write $F(x) = F_1(x) H_1(x)$. $F_1(x)$ has then also a unique maximal completely reducible divisor $H_1(x)$, etc.

THEOREM 7. *Every polynomial has a unique representation as the product of maximal completely reducible factors*

$$F(x) = a_0 H_r(x) \cdots H_2(x) H_1(x). \tag{12}$$

There exists a certain relation between the decomposition of $F(x)$ and any divisor of $F(x)$ which we shall express in the following theorem:

THEOREM 8. *Let $F_1(x)$ be a divisor of $F(x)$ and let*

$$F_1(x) = b_0 G_s(x) \cdots G_2(x) G_1(x) \tag{13}$$

be any decomposition of $F_1(x)$ into completely reducible factors, while the decomposition of $F(x)$ into maximal completely reducible factors is given by (12). *Every product $G_i(x) \cdots G_1(x)$ $(i = 1, \cdots, s)$ is then a divisor of $H_i(x) \cdots H_1(x)$.*

According to the definition of a maximal completely reducible factor, $H_1(x)$ is always divisible by $G_1(x)$. The theorem will be proved by induction. Let us suppose that it has been shown that

$$H_{i-1}(x) \cdots H_1(x) = K_{i-1}(x) G_{i-1}(x) \cdots G_1(x). \tag{14}$$

Since $F_1(x)$ is a divisor of $F(x)$ one can write $F(x) = L(x) F_1(x)$; from (14) one also obtains, when multiplying on the left by $a_0 H_r(x) \cdots H_i(x)$,

$$F(x) = a_0 H_r(x) \cdots H_i(x) K_{i-1}(x) G_{i-1}(x) \cdots G_1(x).$$

Consequently, when one divides on the right by $G_{i-1}(x) \cdots G_1(x)$,

$$L(x) b_0 G_s(x) \cdots G_i(x) = a_0 H_r(x) \cdots H_i(x) K_{i-1}(x). \tag{15}$$

According to Theorem 11 of Chapter 1, and (15), the product $a_0 H_r(x) \cdots H_i(x)$ is divisible by $K_{i-1} G_i(x) K_{i-1}^{-1}$. Since $G_i(x)$ is completely reducible, while $H_i(x)$ is the maximal completely reducible factor of the product $a_0 H_r(x) \cdots H_i(x)$, one concludes that $H_i(x)$ is divisible by $K_{i-1} G_i(x) K_{i-1}^{-1}$. Consequently

$$H_i(x) = T(x) K_{i-1} G_i(x) K_{i-1}^{-1}.$$

We multiply this relation right-hand by $K_{i-1}(x)$ and, since all the polynomials are reduced, one obtains

[5] Compare Krull, Heidelberger Akademie 1926.

$$H_i(x)\,K_{i-1}(x) = T(x)\cdot[G_i(x), K_{i-1}(x)] = K_i(x)\cdot G_i(x). \tag{16}$$

When finally both sides of (16) are multiplied right-hand by $G_{i-1}(x)\cdots G_1(x)$, one obtains according to (14)

$$H_i(x)\cdots H_1(x) = K_i(x)\,G_i(x)\cdots G_1(x).$$

Q. E. D.

As an application of Theorem 8 one can prove

THEOREM 9. *An arbitrary representation of a polynomial as a product of completely reducible factors never contains a smaller number of factors than the decomposition in maximal completely reducible factors.*

The proof follows when one applies Theorem 8 to the special case $F_1(x) = F(x)$. If one had $s > r$ in (12) and (13), then one would obtain

$$H_s(x)\cdots H_1(x) = K_s(x)\,G_s(x)\cdots G_1(x) = K(x)\,F(x)$$

and a divisor of $F(x)$ would be divisible by $F(x)$ itself.

Every decomposition of a polynomial in completely reducible factors which contains the same number of factors as the decomposition in maximal completely reducible factors may therefore be called a *shortest completely reducible decomposition.*

We have, up till the present time, considered only right-hand completely reducible representations; it follows immediately, however, that corresponding theorems hold also for left-hand maximal completely reducible decompositions. It follows also from Theorem 6 that every left-hand completely reducible decomposition is also a right-hand completely reducible decomposition and it is therefore natural to investigate the connection between the two types.

Let us suppose that $F(x)$ is reduced; the left-hand decomposition of $F(x)$ in maximal completely reducible factors is also a right-hand (not maximal) completely reducible decomposition and, according to Theorem 9, it does not contain fewer factors than the right-hand maximal decomposition. The analogous reasoning holds for the right-hand maximal decomposition, and it follows that the right-hand and left-hand maximal decompositions must contain the same number of factors. From Theorem 8 there follows

THEOREM 10. *The decompositions*

$$F(x) = H_r(x)\cdots H_2(x)\,H_1(x), \tag{17}$$

$$F(x) = L_r(x)\cdots L_2(x)\,L_1(x) \tag{18}$$

of a reduced polynomial $F(x)$ into right-hand and left-hand maximal completely reducible factors are both shortest completely reducible decompositions, and, for every $i = 1, \cdots, r$, $H_i(x)\cdots H_1(x)$ is right-hand divisible by $L_i(x)\cdots L_1(x)$ and $L_r(x)\cdots L_i(x)$ is left-hand divisible by $H\ (x)\cdots H_i(x)$.

We shall finally prove a general property of the shortest completel; reducible decompositions which gives a clearer picture of the particula properties of the special shortest decompositions (17) and (18).

THEOREM 11. *Let*

$$F(x) = S_r(x) \cdots S_2(x) S_1(x)$$

be an arbitrary shortest completely reducible decomposition of $F(x)$ whil the two maximal completely reducible decompositions are given in (17 *and* (18). *Then, for every $i = 1, \cdots, r$,*

(19) $$H_i(x) \cdots H_1(x) \equiv 0 \quad (\mathrm{mod}\ S_i(x) \cdots S_1(x)),$$

(20) $$S_i(x) \cdots S_1(x) \equiv 0 \quad (\mathrm{mod}\ L_i(x) \cdots L_1(x)),$$

and similarly for left-hand divisibility,

(21) $$(\mathrm{mod}\ S_r(x) \cdots S_i(x)) \quad 0 \equiv L_r(x) \cdots L_i(x),$$

(22) $$(\mathrm{mod}\ H_r(x) \cdots H_i(x)) \quad 0 \equiv S_r(x) \cdots S_i(x).$$

The congruence (19) is a consequence of Theorem 8. The congruence (20) is correct for $i = r$ and we prove it by induction for all smaller i. Let us suppose, therefore, that it has been shown that

(23) $$S_{i+1}(x) \cdots S_1(x) = K_{i+1}(x) L_{i+1}(x) \cdots L_1(x).$$

Since the same theorems will hold for left-hand divisibility as for right-hand divisibility (assuming $K = \bar{K}$), one can conclude that the product $L_{i+1}(x) \cdots L_1(x)$ is left-hand divisible by the left-hand transform of $S_{i+1}(x)$ by $K_{i+1}(x)$. The polynomial $K_{i+1}^{-1} S_{i+1}(x)_l K_{i+1}$ is completely reducible and, since it divides $L_{i+1}(x) \cdots L_1(x)$, it is also a left-hand divisor of $L_{i+1}(x)$. Consequently

$$L_{i+1}(x) = K_{i+1}^{-1} S_{i+1}(\mathrm{x})_l K_{i+1} T(x)$$

and

$$K_{i+1}(x) L_{i+1}(x) = [S_{i+1}(x), K_{i+1}(x)]_l T(x) = S_{i+1}(x) K_i(x).$$

When this expression is substituted in (23) one can divide on the left by $S_{i+1}(x)$ and one obtains the congruence (20). The congruences (21) and (22) can be obtained in a similar way.

Theorem 11 shows that right-hand maximal completely reducible decomposition is uniquely characterized among the shortest completely reducible decompositions by the fact that $H_i(x) \cdots H_1(x)$ has a maximal degree for each i, while for the left-hand maximal completely reducible decomposition the products $L_i(x) \cdots L_1(x)$ have a minimal degree.

4. The third decomposition theorem. A reduced polynomial is said to be *decomposible* when it can be represented as the union of reduced polynomials

(24) $$F(x) = [A(x), B(x)],$$

where $A(x)$ and $B(x)$ are relatively prime; when no such representation exists, $F(x)$ is said to be *indecomposible.* When $A(x)$ or $B(x)$ in (24) can be decomposed further, we may continue the process and obtain the result that *every polynomial is representable as the union*

$$F(x) = [A_1(x), \cdots, A_r(x)]$$

of mutually prime indecomposible polynomials. The term *mutually prime* applied to a system of polynomials $A_i(x)$ will indicate that every polynomial is relatively prime to the union of the remaining polynomials.

In the proof of the principal theorem about such decompositions we shall need an auxiliary theorem which we now deduce.

THEOREM 12. *When $F(x)$ is decomposible and has the decomposition* (24), *then every divisor $F_1(x)$ of $F(x)$ which is divisible by $A(x)$ is also decomposible and has the representation*

$$F_1(x) = [A(x), (B(x), F_1(x))]. \tag{25}$$

It is obvious that the right-hand side of (25) is a divisor of the left-hand side; to prove the converse, let

$$F_1(x) = Q_1(x)A(x), \qquad F(x) = Q_2(x)F_1(x).$$

Dividing (24) right-hand by $A(x)$ one obtains

$$AB(x)A^{-1} = Q_2(x)Q_1(x),$$

and $Q_1(x)$ is consequently of the form $AD(x)A^{-1}$. Hence $F_1(x) = [A(x), D(x)]$, where $D(x)$ is a divisor of $B(x)$.

We can now prove the *third* decomposition theorem.[6]

THEOREM 13. *Every polynomial has a representation as the union*

$$F(x) = [A_1(x), \cdots, A_r(x)] \tag{26}$$

of mutually prime indecomposible polynomials. If a second such representation

$$F(x) = [B_1(x), \cdots, B_s(x)] \tag{27}$$

exists, then the number of components in both must be the same and they are similar in pairs. An arbitrary component $A_i(x)$ in (26) *can always be replaced by a suitably chosen component $B_j(x)$ in* (27), *and every $B_j(x)$ can be used for some such replacement.*

[6] By means of the method used by Krull, Math. Zeitschr. 8, one could have deduced the fact that the components are similar. The present proof seems somewhat simpler since the "Zurückleitungsgruppen" of Krull are avoided. This proof also yields the fact that the components are replaceable.

It has already been shown that there exists such a decomposition. We shall prove the general theorem by induction and assume it to be true for all polynomials of degree lower than that of $F(x)$.

We show first: *If two decompositions* (26) *and* (27) *exist, then an* $A_i(x)$ (*or a* $B_i(x)$) *can replace a* $B_j(x)$ (*or an* $A_j(x)$). Let $A_1(x)$ be the component of highest degree in the two representations. If then $A_1(x)$ is relatively prime to the union $[B_2(x), \cdots, B_s(x)]$, one must have

$$F(x) = [A_1(x),\ B_2(x), \cdots, B_s(x)],$$

and the conjecture is proven. Let $P(x)$ be a common prime factor of $A_1(x)$ and $[B_2(x), \cdots, B_s(x)]$. Then

$$B(x) = [B_2(x), \cdots, B_s(x)] = \bar{B}(x)\,P(x) \tag{28}$$

and $A_1(x) = \bar{A}_1(x)\,P(x)$. We divide both sides of (26) and (27) by $P(x)$ and obtain

$$[\bar{A}_1(x),\ PA_2(x)\,P^{-1}, \cdots, PA_r(x)\,P^{-1}] = [PB_1(x)\,P^{-1},\ \bar{B}(x)]. \tag{29}$$

According to our assumption the theorem holds for the polynomial (29). The indecomposible component $PB_1(x)\,P^{-1}$ can therefore replace a component on the left-hand side, that is, $PB_1(x)\,P^{-1}$ can replace a $PA_i(x)\,P^{-1}$ $(i = 2, \cdots, r)$ or a component of $\bar{A}_1(x)$.

In the first case let $PB_1(x)\,P^{-1}$ replace $PA_2(x)\,P^{-1}$, giving

$$F(x)\,P(x)^{-1} = [\bar{A}_1(x), PB_1(x)\,P^{-1},\ PA_3(x)\,P^{-1}, \cdots, PA_r(x)\,P^{-1}].$$

Multiplication with $P(x)$ then gives

$$F(x) = [A_1(x),\ B_1(x),\ A_3(x), \cdots, A_r(x)].$$

In the second case one obtains

$$F(x)\cdot P(x)^{-1} = [PB_1(x)\,P^{-1},\ \bar{\bar{A}}_1(x),\ PA_2(x)\,P^{-1}, \cdots, PA_r(x)\,P^{-1}],$$

and consequently

$$F(x) = [B_1(x),\ \bar{\bar{A}}_1(x)\,P(x),\ A_2(x), \cdots, A_r(x)]. \tag{30}$$

This is, however, impossible; if one puts

$$A(x) = [A_2(x), \cdots, A_r(x)], \tag{31}$$

one concludes from (26) and (30), when dividing by $A(x)$, that

$$A\,A_1(x)\,A^{-1} = [A\,B_1(x)\,A^{-1},\ A\,(\bar{\bar{A}}(x)\,P(x))\,A^{-1}],$$

and it would follow that $A_1(x)$ is decomposible, contrary to assumption.

We have consequently shown that, with a suitable notation, one always has a identity of the form

$$(32) \qquad [A_1(x), \cdots, A_r(x)] = [B_1(x), A_2(x), \cdots, A_r(x)].$$

It then follows easily that the two decompositions (26) and (27) have the same number of components similar in pairs. If one divides (32) by $A(x)$ one obtains

$$A A_1(x) A^{-1} = A B_1(x) A^{-1},$$

and $A_1(x)$ and $B_1(x)$ are similar. From the identity

$$[B_1(x), A_2(x), \cdots, A_r(x)] = [B_1(x), B_2(x), \cdots, B_s(x)]$$

there follows, after division by $B_1(x)$,

$$(33) \quad [B_1 A_2(x) B_1^{-1}, \cdots, B_1 A_r(x) B_1^{-1}] = [B_1 B_2(x) B_1^{-1}, \cdots, B_1 B_s(x) B_1^{-1}],$$

and, since the theorem holds for this polynomial, one must have $r = s$ and the $A_i(x)$ are similar to the $B_i(x)$ in pairs.

It now remains only to show that all components in (27) can replace an $A_i(x)$ and that all $A_i(x)$ can be thus replaced, a similar relation holding for the $B_i(x)$. Since this is true for (33), one may replace any $B_1 B_i(x) B_1^{-1}$ by a $B_1 A_i(x) B_1^{-1}$, and when one multiplies afterward by $B_1(x)$ one obtains

$$(34) \qquad F(x) = [B_1(x), B_2(x), \cdots, A_i(x), \cdots, B_r(x)] \qquad (i = 2, \cdots, r),$$

and all $B_i(x)$, with the possible exception of $B_1(x)$, can be replaced. From (34) one concludes, as from (32), that every $A_j(x)$, with the possible exception of $A_i(x)$, may be replaced by a $B_j(x)$ and every $B_j(x)$, except possibly $B_i(x)$, may be used for such a replacement. When $r \geqq 3$, one might have used a different index $j \neq i$ in (34), and the exceptions are thereby avoided. It follows then in the same way that all $B_j(x)$ may be replaced by $A_j(x)$ and that all $A_j(x)$ may be used for replacements.

This leaves only the special case

$$(35) \qquad F(x) = [A_1(x), A_2(x)] = [B_1(x), A_2(x)] = [B_1(x), B_2(x)]$$

unaccounted for. We shall have to show that $A_2(x)$ may also be replaced by $B_1(x)$ or $B_2(x)$, that is, that at least one of the relations

$$(36) \qquad F(x) = [A_1(x), B_1(x)], \qquad F(x) = [A_1(x), B_2(x)]$$

holds. Let us suppose, for instance, that $F_1(x) = [A_1(x), B_2(x)]$ is only a divisor of $F(x)$; then one has, according to Theorem 12,

$$(37) \qquad F_1(x) = [A_1(x), (A_2(x), F_1(x))] = [(F_1(x), B_1(x)), B_2(x)].$$

The theorem holds for $F_1(x)$; from (35) it follows that $A_2(x)$ and $B_2(x)$ are similar and consequently of the same degree, while the degree of $(A_2(x),\ F_1(x))$ must be less than the degree of $B_2(x)$. In (37), therefore, $A_1(x)$ can replace only $B_2(x)$, giving

$$F_1(x) = [(F_1(x),\ B_1(x)),\ A_1(x)],$$

so that

$$F(x) = [B_1(x),\ F_1(x)] = [B_1(x),\ (F_1(x),\ B_1(x)),\ A_1(x)] = [A_1(x),\ B_1(x)].$$

In the same way one shows all the other replacements.

5. **Connection between representations of the second and third kind.** One may ask for the connection between the three main types of representations which we have discussed up till now. It is fairly obvious how one can deduce the prime function decomposition of a polynomial when its representation as a product of maximal completely reducible factors or as a union of indecomposible components is given. We shall study here the connection between the two last types of representations.

Let

$$(38) \qquad F(x) = [A(x), B(x)], \qquad (A(x), B(x)) = 1,$$

and let

$$(39)\ A(x) = A_r(x) \cdots A_2(x)\, A_1(x), \qquad B(x) = B_s(x) \cdots B_2(x)\, B_1(x)$$

be the maximal completely reducible representations of $A(x)$ and $B(x)$. It is easily seen that every prime divisor of $F(x)$ also divides $F_1(x) = [A_1(x), B_1(x)]$ and that $F_1(x)$ is the maximal completely reducible factor of $F(x)$. One also finds that $F(x)$ can be represented in the form

$$F(x) = [\bar{A}(x), \bar{B}(x)]\, F_1(x),$$

where $\bar{A}(x)$ is similar to $A_r(x) \cdots A_2(x)$ and $\bar{B}(x)$ similar to $B_s(x) \cdots B_2(x)$. When the same conclusions are drawn for $[\bar{A}(x), \bar{B}(x)]$, one finds that the maximal completely reducible factor of this polynomial is $F_2(x) = [\bar{A}_2(x), \bar{B}_2(x)]$, where $\bar{A}_2(x)$ is similar to $A_2(x)$ and $\bar{B}_2(x)$ to $B_2(x)$ and so on.

THEOREM 14. *Let* $[A(x), B(x)]$ *be a decomposible polynomial and let* (39) *be the maximal completely reducible representations of* $A(x)$ *and* $B(x)$. *The maximal completely reducible representation of* $F(x)$ *is then*

$$F(x) = F_r(x) \cdots F_1(x), \quad F_i(x) = [\bar{A}_i(x), \bar{B}_i(x)] \quad (i = 1, \cdots, r),$$

where $\bar{A}_i(x)$ *is similar to* $A_i(x)$ *and* $\bar{B}_i(x)$ *similar to* $B_i(x)$.

When $r > s$ one obviously must put $\bar{B}_i(x) = 1$ for $i > s$. For an arbitrary representation

$$F(x) = [F_1(x), \cdots, F_r(x)],$$

where

$$F_i(x) = \Phi_{s_i}^{(i)}(x) \cdots \Phi_1^{(i)}(x)$$

is the representation of each component as product of maximal completely reducible factors, one finds for the corresponding representation of $F(x)$,

$$F(x) = A_s(x) \cdots A_1(x),$$

where

$$A_j(x) = [\bar{\Phi}_j^{(1)}(x), \cdots, \bar{\Phi}_j^{(r)}(x)]$$

and $\bar{\Phi}_j^{(i)}(x)$ is similar to $\Phi_j^{(i)}(x)$. It is also obvious that s is equal to the greatest of the numbers s_i.

6. The fourth decomposition theorem. In the preceding discussion we have deduced some of the more important structure theorems for noncommutative polynomials. They depend on various notions of indecomposibility, each notion leading to its own type of decomposition theorems. There are various other possible definitions of indecomposibility which lead to a number of decomposition theorems. We shall discuss only one of these possibilities.

We shall say that a polynomial is *distributive* if there exists a representation

$$F(x) = [A(x), B(x)]$$

in which $A(x)$ and $B(x)$ are *proper* divisors of $F(x)$; the polynomials $A(x)$ and $B(x)$ do not have to be relatively prime as in the case of decomposible polynomials. Every decomposible polynomial is distributive, but not conversely.

It is easily seen that every reduced polynomial has a representation as the union of non-distributive polynomials

$$F(x) = [A_1(x), \cdots, A_r(x)], \tag{40}$$

where all $A_i(x)$ are supposed to be reduced. In the following we shall always assume that in a representation (40) we have omitted the superfluous components, that is, the polynomials which are divisors of the union of the remaining polynomials. We shall then say that the representation is *minimal.*

We shall first deduce a characteristic property of the non-distributive polynomials. A non-distributive polynomial cannot be divisible on the left by two different prime polynomials, for otherwise one would obtain a representation

$$A(x) = [P_1(x), P_2(x)]_l K(x) \tag{41}$$

and, according to Theorem 17, Chapter 1, the left-hand union could also be considered as a right-hand union

$$[P_1(x),\ P_2(x)]_l = [Q_1(x),\ Q_2(x)],$$

where the prime functions $Q_1(x)$ and $Q_2(x)$ are similar to $P_1(x)$ and $P_2(x)$. From (41) it would then follow that $A(x)$ must be distributive, that is,

$$A(x) = [Q_1(x)\,K(x),\ Q_2(x)\,K(x)].$$

We can also show the converse, namely, that when $A(x)$ is divisible by only a single prime function on the left, that is, when the maximal left-hand completely reducible factor is a prime polynomial, then $A(x)$ is non-distributive. If one had in fact

$$A(x) = [A_1(x),\ A_2(x)],$$

then $A_1(x)$ and $A_2(x)$ would have some right-hand cross-cut $D(x)$ which would have to be a proper divisor of both polynomials. Let us put

$$A_1(x) = \bar{A}_1(x)\,D(x), \qquad A_2(x) = \bar{A}_2(x)\,D(x), \qquad A(x) = \bar{A}(x)\,D(x).$$

Consequently

$$\bar{A}(x) = [\bar{A}_1(x),\ \bar{A}_2(x)],$$

where $\bar{A}_1(x)$ is relatively prime to $\bar{A}_2(x)$. From Theorem 17, Chapter 1, one concludes, however, that $\bar{A}(x)$ must then also be the left-hand union of two relatively prime factors, and $A(x)$ would be left-hand divisible by at least two different prime functions.

THEOREM 15. *The necessary and sufficient condition that a polynomial be non-distributive is that it be left-hand divisible by only a single prime polynomial.*

To every non-distributive polynomial $A(x)$ there exists a unique prime polynomial $P(x)$ which divides it on the left. We shall say that $A(x)$ *belongs to* $P(x)$. An immediate consequence of Theorem 15 is

THEOREM 16. *Every left-hand divisor of a non-distributive polynomial is again non-distributive.*

It is clear that a left-hand divisor of $A(x)$ can be only left-hand divisible by $P(x)$. We shall need also the following result:

THEOREM 17. *The transform of a non-distributive polynomial is non-distributive and belongs to a similar prime polynomial.*

Let us put $A_1(x) = CA(x)C^{-1}$, where we first suppose that the transformation is special, that is, that $(C(x),\ A(x)) = 1$. If then

$$A_1(x) = [L(x),\ M(x)],$$

one obtains

$$A(x) = [C_1\,L(x)\,C_1^{-1},\ C_1\,M(x)\,C_1^{-1}],$$

where $C_1(x)$ is determined such that $C_1(x)\,C(x) \equiv 1$ (mod $A(x)$). $A(x)$ would then be distributive. Now let $A(x)$ and $C(x)$ have a greatest common factor $D(x)$, and let

$$A_1(x) = \bar{A}(x)\,D(x), \qquad C(x) = \bar{C}(x)\,D(x).$$

Here $\bar{A}(x)$ is non-distributive, and, since

$$A_1(x) = CA(x)\,C^{-1} = \bar{C}\,\bar{A}(x)\bar{C}^{-1},$$

this case has been reduced to the first. The second part of the theorem follows from the rule of transformation of a product.

We are now able to prove the following fourth decomposition theorem:

THEOREM 18. *Let $F(x)$ be an arbitrary polynomial and let*

$$F_1(x) = [P_1(x), \cdots, P_r(x)]$$

be the maximal left-hand completely reducible factor of $F(x)$. Every minimal representation of $F(x)$ as a union of non-distributive polynomials has then the form

$$F(x) = [A_1(x), \cdots, A_r(x)], \tag{42}$$

where every non-distributive polynomial $A_i(x)$ belongs to a prime polynomial which is similar to a $P_i(x)$.

The theorem is obviously correct for every completely reducible polynomial. We prove it by induction, assuming it to be true for every left-hand divisor of $F(x)$ which is left-hand divisible by $F_1(x)$. We can then assume that $F(x)$ is not completely reducible, so that

$$F(x) = F_1(x) \cdot Q(x).$$

Let $P(x)$ be a right-hand prime divisor of $F(x)$ and $Q(x)$ such that $F(x)\,P(x)^{-1}$ is left-hand divisible by $F_1(x)$.

Let us first suppose that $F(x)$ has a decomposition of the form (42) with $s \neq r$ components, and let us further suppose that $P(x)$ does not divide any $A_i(x)$. One then obtains

$$F(x) \cdot P(x)^{-1} = [PA_1(x)\,P^{-1}, \cdots, PA_s(x)\,P^{-1}], \tag{43}$$

where every component $PA_i(x)\,P^{-1}$ is non-distributive according to Theorem 17. We wish to show that this representation (43) is minimal. If namely $PA_1(x)\,P^{-1}$ were a divisor of the union of the other $PA_i(x)\,P^{-1}$, then one would obtain, when multiplying by $P(x)$,

$$F(x) = [P(x), A_2(x), \cdots, A_s(x)].$$

But this is not possible, since $F(x)$ would be decomposible into

$$F(x) = [P(x), H(x)]; \qquad H(x) = [A_2(x), \cdots, A_s(x)],$$

and the results of section 5, Chapter 2, would show that the maximal left-hand completely reducible factor of $F(x)\,P(x)^{-1}$ would be of lower degree than $F_1(x)$. The representation (43) is then minimal, giving $r = s$, and the components $PA_i(x)\,P^{-1}$, and therefore also the $A_i(x)$, belong to prime polynomials similar to the $P_i(x)$.

Let us assume in the second place that the first t polynomials $A_i(x)$ in (42) are divisible by $P(x)$,

$$A_1(x) = \bar{A}_1(x)\,P(x), \cdots, A_t(x) = A_t(x)\,P(x).$$

It follows from the choice of $P(x)$ as above that the possibility $A_i(x) = P(x)$ is excluded; the polynomials $\bar{A}_i(x)$ are consequently non-distributive according to Theorem 17 and belong to the same prime polynomial as $A_i(x)$.

When (42) is right-hand divided by $P(x)$, one obtains

$$(44) \quad F(x)\,P(x)^{-1} = [\bar{A}_1(x), \cdots, \bar{A}_t(x), PA_{t+1}(x)\,P^{-1}, \cdots, PA_s(x)\,P^{-1}].$$

This non-distributive representation is also minimal. If, for instance, $PA_s(x)\,P^{-1}$ were a divisor of the union of the remaining polynomials, then it would follow from (44) that the original representation was not minimal. If, for instance, $\bar{A}_1(x)$ were a divisor of the union of the remaining, one would obtain for $t > 1$, that the original representation was not reduced and for $t = 1$

$$F(x) = [P(x), A_2(x), \cdots, A_s(x)].$$

This is shown to be impossible as before.

Since our theorem holds for (44), it follows that $s = r$ and that $A_i(x)$ belongs to a prime polynomial similar to $P_i(x)$.

YALE UNIVERSITY.