

# IDEAL THEORY AND ALGEBRAIC DIFFERENCE EQUATIONS\*

BY

J. F. RITT AND H. W. RAUDENBUSH, JR.

Recent work of J. L. Doob, F. Herzog, W. C. Strodt and J. F. Ritt furnishes a theory of manifolds for systems of algebraic difference equations.† We present here a basis theorem for infinite systems of difference polynomials, and a restricted theory of ideals; there is obtained thus a counterpart, for difference equations, of Raudenbush's work on differential equations.‡

In the theory of algebraic polynomials, one derives an infinite system from a basis by forming linear combinations. A system of differential polynomials is obtained from a basis by differentiations, linear combinations and the extraction of roots. For difference polynomials, one performs "shufflings" in succession, each shuffling consisting in taking *transforms*, performing linear combinations and factoring forms into products of transforms.

We leave open the question as to how many shufflings are necessary in order to produce a system from its basis. Conceivably, an infinite number may be necessary in certain cases; we give an example, in §15, for which two shufflings are required.

We shall, in this paper, work with difference polynomials whose coefficients lie in an abstract field. If we have decided to relinquish the meromorphic coefficients used in R. D., it is because, with the present undeveloped state of the analytic theory of nonlinear difference equations, it appears tactical to proceed in an algebraic direction, hoping for analytical developments to follow. For example, the establishment of existence theorems which will permit the translation into analytic terms of the *Nullstellensatz* presented in §13, is a problem with a distinct challenge.

## DIFFERENCE RINGS

1. Let  $\mathcal{R}$  denote a commutative ring§ possessing a unit element. Let us

---

\* Presented to the Society, September 8, 1939; received by the editors May 24, 1939.

† For references, see *Semicentennial Addresses of the American Mathematical Society*, New York, 1938, pp. 54, 55. The present paper attaches particularly to Ritt and Doob, *Systems of algebraic difference equations*, American Journal of Mathematics, vol 55 (1933). That paper will be designated below by R. D.

‡ Actually, our considerations hold for equations involving a substitution of any type; for instance, for  $q$ -difference equations.

§ Defined as in van der Waerden, *Moderne Algebra*, chap. 3. Small italics will, until §9, usually represent elements of  $\mathcal{R}$ .

suppose that, for every element  $a$  of  $\mathcal{R}$ ,  $\mathcal{R}$  contains a unique element  $a_1$ , called the *transform* of  $a$ , the correspondence between elements and their transforms being such that:

( $\alpha$ ) the transform of unity is unity;

( $\beta$ ) for every  $a$  and  $b$  in  $\mathcal{R}$ ,  $(a+b)_1 = a_1 + b_1$  and  $(ab)_1 = a_1 b_1$ .

We shall, under these circumstances, call  $\mathcal{R}$  a *difference ring*. If  $\mathcal{R}$ , in addition to being a ring, is a field, we shall call  $\mathcal{R}$  a *difference field*.\*

In everything which follows, we deal with a fixed difference ring  $\mathcal{R}$ .

2. We denote  $(a_1)_1$  by  $a_2$  and, by induction, define  $a_n$  as  $(a_{n-1})_1$  for every  $n > 1$ . We shall call  $a_n$  the *transform of  $a$  of order  $n$*  or the  *$n$ th transform of  $a$* . The element  $a$  will be described as its own transform of order 0 and will be denoted, at times, by  $a_0$ . We shall refer to the  $a_n$ ,  $n=0, 1, \dots$ , as *transforms of  $a$* ; the transform of  $a$  will continue to mean, as above,  $a_1$ .

### IDEALS

3. An ideal  $\pi$  contained in  $\mathcal{R}$  will be called a *difference ideal* if, given any element  $a$  in  $\mathcal{R}$ , the presence in  $\pi$  of either of  $a$  and  $a_1$ , where  $a_1$  is the transform of  $a$ , implies the presence in  $\pi$  of the other. Thus, if  $a$  is in  $\pi$ ,  $\pi$  contains every transform of  $a$  and also contains every element in  $\mathcal{R}$  of which  $a$  is a transform of some order.†

4. A difference ideal  $\pi$  will be called *perfect* if, whenever  $a$  is such that some product of positive integral powers of transforms of  $a$  is contained in  $\pi$ ,  $a$  is also contained in  $\pi$ . That is, if

$$a_p a_q^i \cdots a_r^k,$$

where  $p, q, \dots, r$  are distinct nonnegative integers and  $i, j, \dots, k$  are positive integers, is in  $\pi$ ,  $a$  is in  $\pi$ .

A difference ideal  $\pi$  will be called *prime* if, whenever  $ab$  is in  $\pi$ , at least one of  $a$  and  $b$  is in  $\pi$ ;  $\pi$  will thus be a prime ideal in the sense in which that term is regularly used in algebra. *Every prime difference ideal is perfect.*

Henceforth, unless other indications are given, *ideal* will mean difference ideal.

\* Let  $b \neq 0$ . By ( $\alpha$ ),  $b_1(1/b)_1 = 1$ . Hence  $b_1 \neq 0$  and  $(a/b)_1 = a_1/b_1$  for every  $a$ .

† Our insistence that  $\pi$  contain, together with  $a$ , all transforms of  $a$  "of negative orders" which may exist in  $\mathcal{R}$ , is explained by the material of §4. Our definition of ideal appears to have sufficient generality for the purposes of the concrete applications; one will notice, for instance, that a difference form with meromorphic coefficients has the same manifold as its transform. (Cf. R. D.) The fact that many investigations on difference equations deal with half of the complex plane seems to make it undesirable to assume that, for every  $a$  in  $\mathcal{R}$ , there is an element in  $\mathcal{R}$  of which  $a$  is a transform.

## PERFECT IDEALS GENERATED BY A SET OF ELEMENTS

5. Let  $\sigma$  be any set of elements in  $\mathcal{R}$ .<sup>\*</sup> There exist perfect ideals in  $\mathcal{R}$ , for instance  $\mathcal{R}$  itself, which contain  $\sigma$ . The intersection of all such perfect ideals is a perfect ideal which contains  $\sigma$ . We denote this intersection by  $\{\sigma\}$  and call it the *perfect ideal generated by  $\sigma$* .

We shall study the relationship of  $\{\sigma\}$  to  $\sigma$ .

Let  $\tau$  be any set of elements of  $\mathcal{R}$ . Consider all elements of  $\mathcal{R}$  which are of the form

$$au + bv + \cdots + cw,$$

where  $u, v, \cdots, w$  are transforms of any orders of elements of  $\tau$  and  $a, b, \cdots, c$  are in  $\mathcal{R}$ . The totality of such elements will be denoted by  $[\tau]$ ;  $[\tau]$  may not be a difference ideal, but it will be an ideal in the sense of algebra and it will be closed with respect to "transforming."

Again, let  $a$  be any element in  $\mathcal{R}$  which is such that some product of positive powers of transforms of  $a$  is in  $\tau$ . The totality of such elements  $a$  will be denoted by  $\tau'$ .<sup>†</sup>

Returning to  $\sigma$  above, let  $\sigma_1 = [\sigma]'$  and, continuing inductively, let  $\sigma_n = [\sigma_{n-1}]'$  for every  $n > 1$ . The logical sum, or what is the same, the *limit*, of the sets  $\sigma_n$  is easily seen to be  $\{\sigma\}$ .

Because  $[\sigma]$  and the  $[\sigma_n]$  are closed with respect to "transforming," every  $\sigma_n$  is also so closed. Thus, for  $n \geq 1$ , each element of  $[\sigma_n]$  is a linear combination of elements of  $\sigma_n$ .

In what follows, a plus sign between two sets will indicate that the logical sum of the sets is to be taken.

6. We prove the following:

LEMMA I. *Let  $\sigma$  be any set of elements of  $\mathcal{R}$  and  $a$  and  $b$  any two elements of  $\mathcal{R}$ . If  $d$  is contained in  $(\sigma + a)_n$  and  $e$  in  $(\sigma + b)_n$ ,  $n \geq 1$ , then  $de$  is contained in  $(\sigma + ab)_{n+1}$ .*<sup>‡</sup>

First, let  $n = 1$ . There exist a product  $\bar{d}$  of positive powers of transforms of  $d$ , and an  $\bar{e}$  similarly related to  $e$ , which have expressions

$$\bar{d} = gu + \cdots + hv + ka_i + \cdots + la_j,$$

$$\bar{e} = g'u' + \cdots + h'v' + k'b_p + \cdots + l'b_q,$$

<sup>\*</sup> For the purposes of §§11, 12, it is desirable to allow a given element of  $\mathcal{R}$  to occur more than once in  $\sigma$ . Thus, the elements in  $\sigma$  are supposed to be provided with marks and a single element of  $\mathcal{R}$  may appear many times in  $\sigma$ , each time in association with a different mark. When we have to do with ideals, however, a given element will be assumed to appear only once.

<sup>†</sup> In  $[\tau]$  and in  $\tau'$  a given element will be understood to occur only once. The notation in the present paragraphs, as regards accents and subscripts, is of an episodic character.

<sup>‡</sup> The parentheses are ordinary symbols of aggregation. Thus,  $(\sigma + a)_1 = [\sigma + a]'$ .

where  $u, \dots, v$  and  $u', \dots, v'$  are transforms of elements of  $\sigma$  and the subscripted  $a$  and  $b$  are transforms of  $a$  and  $b$ . Thus  $\bar{d}\bar{e}$  has an expression in which some terms are in  $[\sigma]$  and in which the others are of the type  $fa_r b_s$ . For any  $r$  and  $s$ , the product of  $a_r b_s$  by a suitable transform of itself is a multiple\* of a transform of  $ab$ . Thus every  $a_r b_s$  is in  $[ab]'$  and, a fortiori, in  $(\sigma+ab)_1$ . Then  $\bar{d}\bar{e}$  is in  $[(\sigma+ab)_1]$ . Some product of powers of transforms of  $de$  is a multiple of  $\bar{d}\bar{e}$ . Thus  $de$  is in  $(\sigma+ab)_2$ .

Now, let  $n=2$ . Let  $\bar{d}$ , described as above, be in  $[(\sigma+a)_1]$ . By §5,  $\bar{d}$  is a linear combination of elements of  $(\sigma+a)_1$ . We use an  $\bar{e}$ , described as above, which is linear in elements of  $(\sigma+b)_1$ . Then  $\bar{d}\bar{e}$  has an expression in which each term is of the type  $guv$  with  $u$  in  $(\sigma+a)_1$  and  $v$  in  $(\sigma+b)_1$ . Now  $uv$ , by the case of  $n=1$ , is in  $(\sigma+ab)_2$ . Hence  $\bar{d}\bar{e}$  is in  $[(\sigma+ab)_2]$ . This puts  $de$  in  $(\sigma+ab)_3$ .

The proof continues by induction.

**LEMMA II.** *Let  $\sigma$  be any set of elements of  $\mathcal{R}$  and  $a$  and  $b$  any two elements of  $\mathcal{R}$ . Then  $\{\sigma+ab\}$  is the intersection of  $\{\sigma+a\}$  and  $\{\sigma+b\}$ .*

We have only to show that,  $c$  being any element in the intersection,  $c$  is contained in  $\{\sigma+ab\}$ . Let  $n$  be such that  $c$  is contained in  $(\sigma+a)_n$  and in  $(\sigma+b)_n$ . Then  $c^2$  is in  $(\sigma+ab)_{n+1}$ . Thus  $c$  is also in  $(\sigma+ab)_{n+1}$ .

## BASES

7. Let  $\sigma$  be a system of elements in  $\mathcal{R}$ . A finite subset  $\phi$  of  $\sigma$  will be called a *basis* of  $\sigma$  if  $\{\phi\}$  contains  $\sigma$ .

A finite system of elements is a basis for itself. If every infinite system of elements in  $\mathcal{R}$  has a basis, we shall call  $\mathcal{R}$  a *difference ring with a basis theorem*.

## DECOMPOSITION OF PERFECT DIFFERENCE IDEALS

8. Let  $\mathcal{R}$  have a basis theorem. We prove the theorem:

**THEOREM.** *Every perfect ideal in  $\mathcal{R}$  is the intersection of a finite set of prime ideals.*

Let  $\pi$  be a perfect ideal for which our statement is false. Then  $\pi$  is not prime. Let  $ab$  be in  $\pi$  while neither  $a$  nor  $b$  is. Then  $\pi$  is the intersection of  $\{\pi+a\}$  and  $\{\pi+b\}$  (Lemma II). At least one of the two latter ideals does not have the property of being the intersection of a finite set of prime ideals. Of the two ideals, let  $\pi_1$  designate one which lacks the property. We give  $\pi_1$  the treatment accorded to  $\pi$  and continue, forming a sequence of perfect ideals

$$(1) \quad \pi, \pi_1, \dots, \pi_n, \dots,$$

---

\* The meaning is obvious.

each a proper part of its successor. Let  $\sigma$  be the logical sum of the ideals in (1) and let  $\phi$  be a basis of  $\sigma$ . There is some  $\pi_n$  which contains  $\phi$ . That  $\pi_n$  will contain  $\sigma$ . This contradiction proves the theorem.

It is easy now to see that every perfect ideal in  $\mathcal{R}$  has a *unique* representation as the intersection of a finite number of prime ideals *none of which contains any other*.

#### IDEALS OF DIFFERENCE POLYNOMIALS

9. Let  $n$  be any positive integer. We consider  $n$  symbols  $y_1(x), \dots, y_n(x)$ . If  $j$  is any nonnegative integer, we shall call the symbol  $y_i(x+j)$  the *jth transform* of  $y_i(x)$ .

Let  $\mathcal{F}$  be a given difference field. By a *difference polynomial* we shall mean a polynomial in a certain (eo ipso finite) number of the  $y_i(x+j)$ , with coefficients in  $\mathcal{F}$ . As a rule, we shall substitute the briefer term *form* for "difference polynomial." By the *transform* of a form  $A$ , we mean the form obtained when  $x$  is replaced by  $(x+1)$  in the  $y_i(x+j)$  appearing in  $A$ , and when the coefficients in  $A$  are replaced by their transforms. Transforms of higher order are defined similarly. Because the transform of unity is unity, these definitions are consistent with the definition given above of the *jth transform* of  $y_i(x)$ .

The totality of forms with coefficients in  $\mathcal{F}$  is a difference ring, which we shall call the *ring of forms in the unknowns*  $y_1, \dots, y_n$ . Any form of this ring will be called a *form in*  $y_1, \dots, y_n$ .

10. We prove the theorem:

**THEOREM.** *For any difference field  $\mathcal{F}$ , the ring of difference polynomials in the unknowns  $y_1, \dots, y_n$  is a difference ring with a basis theorem.*

We assume the theorem to be false and work towards a contradiction. We shall use methods and results of R. D. The items of that paper which will be employed here acquire validity, with no essential change, for an abstract field  $\mathcal{F}$ .

11. We prove the following lemma:

**LEMMA III.** *Let  $\Sigma$  be a system of forms in  $y_1, \dots, y_n$  which has no basis. Let  $F_1, \dots, F_s$  be such that, by multiplying each form in  $\Sigma$  by some product of nonnegative powers of transforms of  $F_1, \dots, F_s$ , a system  $\Lambda$  is obtained which has a basis. Then at least one of the systems  $\Sigma + F_i$ ,  $i = 1, \dots, s$ , has no basis.*

Suppose that every  $\Sigma + F_i$  has a basis. Then, for each  $i$ , there is a finite subset  $\Phi_i$  of  $\Sigma + F_i$  such that  $\{\Phi_i\}$  contains  $\Sigma + F_i$ . As  $\Phi_i$  may evidently be replaced by any finite subset of  $\Sigma + F_i$  which contains  $\Phi_i$ , we may (and shall) suppose  $\Phi_i$ , for every  $i$ , to be of the type

$$(2) \quad F_i, A_1, \dots, A_q,$$

with the set

$$(3) \quad A_1, \dots, A_q$$

independent of  $i$ . Enlarging (3) if necessary, we assume that the subset of  $\Lambda$  obtained from (3) by the above described multiplications is a basis of  $\Lambda$ . Thus, the perfect ideal generated by the set (3) contains  $\Lambda$ .

Let  $L$  be any form in  $\Sigma$ . Then  $L$  is contained, for every  $i$ , in the perfect ideal generated by the set (2). Certainly,  $L$  is contained in the perfect ideal generated by

$$L, A_1, \dots, A_q.$$

By Lemma II of §6,  $L$  is contained in the perfect ideal generated by

$$(4) \quad F_1 F_2 \dots F_s L, A_1, \dots, A_q.$$

Some  $KL$ , with  $K$  a product of powers of transforms of the  $F_i$ , belongs to  $\Lambda$ . Now an appropriate product of powers of transforms of  $F_1 F_2 \dots F_s L$  is a multiple of  $KL$ . Since  $\Lambda$  is contained in the perfect ideal generated by (3),  $F_1 \dots F_s L$  is also so contained. Inspecting (4), we see that  $L$  is contained in the perfect ideal generated by (3). Then (3) is a basis of  $\Sigma$ . This contradiction proves the lemma.

12. From among all systems of forms in  $y_1, \dots, y_n$  which have no basis, we select one,  $\Sigma$ , whose basic sets are not higher than those of any other system which has no basis. Let

$$(5) \quad A_1, \dots, A_r$$

be a basic set of  $\Sigma$ . Then  $A_1$  must be of class greater than 0, else  $\Sigma$  would be contained in  $\{A_1\}$ . Let  $I_i$  be the initial of  $A_i$ ,  $i=1, \dots, r$ .

For every form of  $\Sigma$  not in (5), let a remainder with respect to (5) be found. Let  $\Lambda$  be the system composed of the forms in (5) and of the products of the forms of  $\Sigma$  not in (5) by the power products of transforms of the  $I_i$  used in forming the remainders. Let  $\Omega$  be the system composed of (5) and of the remainders of the forms of  $\Sigma$  not in (5).

By the considerations of R. D.,  $\Omega$  has a basis. Such a basis is a basis of  $\{\Omega\}$ . Now  $[\Lambda]$  is identical with  $[\Omega]$ , so that  $\{\Lambda\}$  is identical with  $\{\Omega\}$ . Thus  $\{\Lambda\}$  has a basis. It is easy now to see that  $\{\Lambda\}$  has a basis composed of forms of  $\Lambda$ . Such a basis is a basis of  $\Lambda$ .

The lemma of §11 informs us now that some  $\Sigma + I_i$  has no basis. As in R. D., this is impossible. The theorem is proved.

## HILBERT-NETTO THEOREM

13. Following the procedure of Raudenbush for differential equations, one can prove that if  $\Phi$  is a finite system of difference polynomials with coefficients in a difference field  $\mathcal{F}$  and if  $G$  is a form which is not in  $\{\Phi\}$ , then there exists an extension  $\mathcal{F}'$ , of  $\mathcal{F}$ , containing a solution of  $\Phi$  which is not a solution of  $G$ .\*

## DERIVATION OF AN INFINITE SYSTEM FROM A BASIS

14. Let  $\Sigma$  be an infinite system of forms in  $y_1, \dots, y_n$  and  $\Phi$  a basis of  $\Sigma$ . Given any form  $A$  in  $\Sigma$ , there is some  $\Phi_i$  (notation as in §5) which contains  $A$ . It is natural to ask whether there is some  $\Phi_i$  which contains  $\Sigma$ , and, indeed, whether  $\Phi_1$  contains  $\Sigma$ .

We shall present a system  $\Sigma$  which has no basis  $\Phi$  for which  $\Phi_1$  holds  $\Sigma$ . Whether there is a system with no basis for which  $\Phi_2$  contains  $\Sigma$ , we do not know.

15. Our example will deal with forms in a single unknown  $y$ . We use the field of constants, each constant being its own transform. The  $j$ th transform of  $y$  will be denoted by  $y_j$ . The system  $\Sigma$  will consist of the sequence of forms  $A_0, A_1, \dots$ , where

$$A_j = (y_0 + y_1 + \dots + y_{2^j-1})(y_{2^j} + \dots + y_{2^{j+1}-1}).$$

Let  $m$  be any nonnegative integer. Let  $\Phi$  consist of  $A_0, \dots, A_m$ . We shall show that  $A_{m+1}$  is not contained in  $\Phi_1$ .

Consider any  $A_j$  with  $j \leq m$ . Each term in  $A_j$  is of the type  $y_a y_b$  with

$$(6) \quad 0 < b - a \leq 2^{m+1} - 1.$$

Hence, if  $G$  is any form in  $[\Phi]$ , each term in  $G$  has, among its letters, two letters  $y_a$  and  $y_b$  where  $a$  and  $b$  satisfy (6). We shall show that if  $K$  is a product of powers of transforms of  $A_{m+1}$ ,  $K$  contains a term each of whose letters  $y_j$  has its subscript  $j$  divisible by  $2^{m+1}$ . This will prove that  $A_{m+1}$  is not contained in  $\Phi_1$ .

Let  $B_i$  represent the  $i$ th transform of  $A_{m+1}$ ,  $i=0, 1, \dots$ . Let  $B_j^r$ , with  $r>0$ , be one of the powers of which  $K$  is a product. Let

$$j = p2^{m+1} + q,$$

with  $p$  and  $q$  nonnegative integers and  $q < 2^{m+1}$ . We have

$$(7) \quad B_j = (y_j + \dots + y_{2^{m+1}+j-1})(y_{2^{m+1}+j} + \dots + y_{2^{m+2}+j-1}).$$

\* These Transactions, vol. 36 (1934), p. 368. The term *extension* is self-explanatory.

If  $q=0$ , the first parenthesis in (7) contains  $y_{p2^{m+1}}$  and the second contains  $y_{(p+1)2^{m+1}}$ . If  $q>0$ , the first parenthesis contains  $y_{(p+1)2^{m+1}}$  and the second  $y_{(p+2)2^{m+1}}$ . In any case,  $B_j$  contains one and only one term  $y_a y_b$  in which  $a$  and  $b$  are both divisible by  $2^{m+1}$ . Furthermore,  $B_j^r$  contains a term which is a power of  $y_a y_b$ , and that term of  $B_j^r$  is the only one in which every subscript is divisible by  $2^{m+1}$ . Our statement with respect to  $K$  follows.

We observe that  $A_0$  is a basis of  $\Sigma$  and that, if  $\Psi$  represents  $A_0$ ,  $\Psi_1$  contains  $y_0$  so that  $[\Psi_1]$  and  $\Psi_2$  contain  $\Sigma$ .

COLUMBIA UNIVERSITY,  
NEW YORK, N. Y.,  
QUEENS COLLEGE,  
FLUSHING, N. Y.