

A THEOREM ON THE FACTORIZATION OF POLYNOMIALS OF A CERTAIN TYPE*

BY

LLOYD L. DINES

Polynomials of the form

$$P \equiv y^n + P_{n-1}(x)y^{n-1} + \dots + P_1(x)y + P_0(x),$$

where the coefficients $P_j(x)$ are power series converging in a sufficiently restricted neighborhood of the origin and vanishing with x , are of importance in the theory of implicit functions. The form of analytic expression of the functions $y(x)$ defined by the equation $P = 0$ depends upon the *characteristic line* (often called the Newton polygon) of the polynomial P . The nature of this dependence strongly suggests the theorem of the present paper, and may in fact afford a proof of it. The proof here given, however, proceeds along more direct lines, and like the theorem itself is algebraic in character. An analogous theorem, relative to polynomials of which the coefficients are "Hensel series" proceeding according to powers of a prime number p , and for which the factorization is to be made in this domain was proved by Dumas.† The theorem of the present paper, like that of Dumas, admits a converse, and both converses are special cases of the "Theorem on the Product" proved by Blumberg‡ for a very general class of polynomials defined postulationally.

1. THE THEOREM

If the characteristic line§ of the polynomial P is a broken line of k segments, then P is a product of k polynomials of the same type as P , each of which has

* Presented to the Society, September 7, 1923.

† *Journal de Mathématiques*, ser. 6, vol. 2 (1906), p. 223.

‡ *These Transactions*, vol. 17 (1916), p. 530.

§ The characteristic line is to be thought of as drawn relative to a system of rectangular i, j axes, its form being determined by the "degree points" (i, j) for which the coefficients p_{ij} of the double series

$$P \equiv \sum_{i=0}^{\infty} \sum_{j=0}^n p_{ij} x^i y^j$$

are different from zero. It is not, strictly speaking, an entire "Newton polygon", but only that part which is convex toward the origin $i = 0, j = 0$. For a definition and discussion of the Newton polygon, see Chrystal's *Algebra*, Part II, pp. 386-96.

an unbroken characteristic line the same in length and direction as one segment of the characteristic line of P .

This theorem is evidently an immediate consequence, by repetition, of the following somewhat simpler proposition:

If the characteristic line of P is a broken line, and if m is the ordinate of its lowest vertex, then P is a product of two polynomials of the same type as P , of degrees m and $n - m$ respectively, and such that their characteristic lines can be obtained by translation from the two sections of the characteristic line of P of which this vertex is the upper end and the lower end respectively.

We will prove the proposition in this simpler form. To do this we must prove the existence of two polynomials

$$A \equiv y^m + A_{m-1}(x)y^{m-1} + \dots + A_1(x)y + A_0(x),$$

$$B \equiv y^\mu + B_{\mu-1}(x)y^{\mu-1} + \dots + B_1(x)y + B_0(x) \quad (\mu = n - m),$$

satisfying the identity

$$(1) \quad A \cdot B \equiv P.$$

Substituting for A , B , and P in (1) their polynomial expressions, and equating coefficients of like powers of y on the two sides of the equality, we obtain a system of equations which may be written

$$(2) \quad A_j B_0 + A_{j-1} B_1 + \dots + A_{j-r} B_r + \dots + A_0 B_j = P_j(x) \\ (j = 0, 1, \dots, n-1),$$

with the following interpretation of notations:

$$(3) \quad A_m = B_\mu = 1, \quad A_{m+q} = B_{\mu+q} = 0 \quad (q > 0).$$

The system (2) consists of n equations for the determination of the n coefficients

$$A_j \quad (j = 0, 1, \dots, m-1), \quad B_j \quad (j = 0, 1, \dots, \mu-1).$$

We will show that under the conditions of the hypothesis upon the characteristic line of P , the determination is possible, and unique.

2. A TRANSFORMATION OF THE SYSTEM (2)

We denote by the symbol $[j]$, the smallest integer i for which the point (i, j) does not lie to the left of the characteristic line of P . As a consequence of this definition, each coefficient $P_j(x)$ contains $x^{[j]}$ as a factor. We introduce the quantities \bar{P}_j , \bar{A}_j , and \bar{B}_j as follows:

$$(4) \quad P_j \equiv x^{[j]} \bar{P}_j, \quad A_j \equiv x^{[j]-[m]} \bar{A}_j, \quad \bar{B}_j \equiv x^{[j+m]} \bar{B}_j \quad (j = 0, 1, \dots, n-1).$$

Substituting these values of P_j , A_j , B_j into the system of equations (2), and dividing the j th equation by $x^{[j]}$ for each j , we obtain a system which may be written

$$(5) \quad \bar{A}_j \bar{B}_0 + x^{\{j,1\}} \bar{A}_{j-1} \bar{B}_1 + \dots + x^{\{j,r\}} \bar{A}_{j-r} \bar{B}_r + \dots + x^{\{j,j\}} \bar{A}_0 \bar{B}_j = P_j(x) \\ (j = 0, 1, \dots, n-1),$$

where the symbol $\{j, r\}$ has the meaning

$$\{j, r\} \equiv [j-r] - [m] + [m+r] - [j].$$

We shall see that the exponent $\{j, r\}$ is never negative in a non-vanishing term of (5), but that in a certain well defined class of terms it is indeed positive.

3. THE EXPONENT $\{j, r\}$ AND THE ARTIFICIAL ORDER

Let us denote by k_α and m_α , respectively, the projections upon the i and j axes of that segment of the characteristic line of P which meets and lies above the line $j = \alpha$. For the sake of brevity, we will denote the projections of the lowest segment merely by k and m . That is, when $\alpha < m$, $k_\alpha = k$, and $m_\alpha = m$. And from the defining properties of the characteristic line it follows that if $\alpha \geq m$, then $k_\alpha/m_\alpha < k/m$.

From the definition of the symbol $[j]$ and the properties of the characteristic line follow certain inequalities which we shall need, and which may be stated as follows:

$$(6) \quad [m-q] - [m] \geq q \frac{k}{m}, \quad 0 \leq q \leq m,$$

$$(7) \quad [m] - [m+r] < r \frac{k}{m}, \quad 0 < r \leq \mu,$$

$$(8) \quad [j] - [j+s] < s \frac{k_j}{m_j} + 1, \quad 0 \leq s \leq m + \mu - j.$$

We are now in a position to study the exponent $\{j, r\}$. Since by (3) and (4) $\bar{A}_{j-r} = 0$ when $j-r > m$, this exponent can occur in non-vanishing terms of (5) only if $j-r \leq m$, that is if $j \leq m+r$.

Taking $q = m - (j-r)$ in (6), and $s = (m+r) - j$ in (8) we obtain the two relations

$$(6') \quad [j-r] - [m] \geq (m-j+r) \frac{k}{m},$$

$$(8') \quad [j] - [m+r] < (m+r-j) \frac{k_j}{m_j} + 1.$$

Subtracting the second of these relations from the first, we find that

$$\{j, r\} > (m-j+r) \left(\frac{k}{m} - \frac{k_j}{m_j} \right) - 1,$$

and since the right hand side of this inequality cannot be less than -1 , we have shown that *the exponent $\{j, r\}$ cannot be negative in a non-vanishing term of (5).*

Certain other important statements to be made relative to the exponent $\{j, r\}$ depend upon what we shall call the "artificial order" for the subscripts $j = 0, 1, 2, \dots, n-1$. This artificial order is suggested by the rôle which the lowest segment of the characteristic line plays in our theorem. If the coördinates of the upper and lower ends of this last segment be denoted by (l, m) and $(k+l, 0)$ respectively, the equation of the last segment is

$$mi + kj = (k+l)m.$$

By the artificial order of the subscripts j , we shall mean the order in which the corresponding points $([j], j)$ occur on the sequence of parallel lines

$$mi + kj = \alpha, \quad \alpha = 1, 2, 3, \dots,$$

the order on each individual line being downward.

More precisely, we define the *level* of a subscript j to be the integer given by the formula

$$\text{level of } j = m[j] + kj.$$

We then order the subscripts $j = 0, 1, \dots, n-1$, according to increasing level, with the additional convention that of two subscripts having the same level, the greater subscript shall precede.

In the statements (a) and (b) which follow, the term "order" is to be interpreted as the artificial order just defined.

(a) If the order of j is lower than the order of $j-r$, then $\{j, r\} > 0$.

(b) If $j < m+r$, and the order of j is lower than the order of $m+r$, then $\{j, r\} > 0$.

To prove (a), we have by hypothesis

$$m[j] + kj \leq m[j-r] + k(j-r)$$

and hence

$$[j-r] - [j] \geq r \frac{k}{m}.$$

This inequality, together with (7), yields the conclusion of (a), namely $\{j, r\} > 0$.

To prove (b), we have by hypothesis

$$m[j] + kj < m[m+r] + k(m+r)$$

or

$$[m+r] - [j] > (j-m-r) \frac{k}{m}.$$

From this inequality, and the inequality (6') follows the conclusion of (b).

4. THE SOLUTION OF THE SYSTEM OF EQUATIONS

Turning now to the system (5), we note that since $\bar{A}_j = 0$ when $j > m$, the equations fall under two slightly different types, namely

$$(9) \quad \begin{aligned} \bar{A}_j \bar{B}_0 + \sum_{r=1}^j x^{\{j,r\}} \bar{A}_{j-r} \bar{B}_r &= \bar{P}_j(x) & (j < m), \\ \bar{B}_{j-m} + \sum_{r=j-m+1}^j x^{\{j,r\}} \bar{A}_{j-r} \bar{B}_r &= \bar{P}_j(x) & (j \geq m). \end{aligned}$$

We will hereafter think of the equations of this system as arranged in the order indicated by the artificial order of the subscripts j occurring on the right hand sides. Furthermore the n unknowns \bar{A}_j ($j = 0, 1, \dots, m-1$)

and \bar{B}_r ($r = 0, 1, \dots, \mu - 1$) are to be ordered according to the following conventions: \bar{A}_j is to have the same place in the sequence of unknowns that its subscript j has in the artificial order of subscripts; \bar{B}_r is to have the same place in the sequence of unknowns that $m + r$ has in the artificial order of subscripts.

By virtue of this ordering of unknowns, it follows from (a) and (b) that the equations (9) can be written in the forms

$$(10) \quad \begin{aligned} G_j + \bar{A}_j \bar{B}_0 + x H_j &= \bar{P}_j(x) & (j < m), \\ G_j + \bar{B}_{j-m} + x H_j &= \bar{P}_j(x) & (j \geq m), \end{aligned}$$

where G_j and H_j are binary forms in the unknowns \bar{A} and \bar{B} , and furthermore each G_j has the important property that it contains no \bar{A} or \bar{B} of order as high as the order of the subscript j .

Consider now the system of equations obtained from (10) by putting $x = 0$:

$$(11) \quad \begin{aligned} G_j + \bar{A}_j \bar{B}_0 &= \bar{P}_j(0) & (j < m), \\ G_j + \bar{B}_{j-m} &= \bar{P}_j(0) & (j \geq m). \end{aligned}$$

Since the first subscript j in the artificial order is $j = m$,* the first equation of this system is

$$\bar{B}_0 = \bar{P}_m(0).$$

The value of \bar{B}_0 so determined is different from zero, since $\bar{P}_m(0) = p_m \neq 0$.

Since, as we have seen, each form G_j contains only \bar{A} 's and \bar{B} 's of order lower than the order of the corresponding subscript j , and since $\bar{B}_0 \neq 0$, it follows that the equations (11) can be solved successively in their artificial order for \bar{A}_j ($j = 0, 1, \dots, m - 1$) and \bar{B}_{j-m} ($j = m, m + 1, \dots, n - 1$). We thus obtain a solution of the system (10) corresponding to the value $x = 0$, which we may denote by

$$(12) \quad x = 0, \quad \bar{A}_j = a_j, \quad \bar{B}_{j-m} = b_{j-m}.$$

The fundamental existence theorem for implicit functions insures the existence of power series $\bar{A}_j(x)$, $\bar{B}_{j-m}(x)$ converging in a neighborhood of

* The point (l, m) or $([m], m)$ being the upper end of the lowest segment of the characteristic line.

the origin, and identically satisfying the system of equations (10), provided the functional determinant of the system with respect to the n variables \bar{A}_j , \bar{B}_{j-m} is different from zero for the set of values (12). We therefore consider now this functional determinant.

If the order of the rows of this determinant be made to agree with the artificial order of the subscripts j , and the order of the columns to agree with the artificial order of the variables \bar{A} and \bar{B} , it is apparent that the elements of the principal diagonal will for $x = 0$ be b_0 ($= \bar{B}_0$) or 1, according as the corresponding subscript j is less than m or not. Furthermore every element to the right of the principal diagonal is zero for $x = 0$. Hence the functional determinant has for the set of values (12) the value

$$\Delta = b_0^m = p_{lm}^m \neq 0.$$

The system of equations (10) have therefore by the existence theorem, unique convergent solutions

$$(13) \quad \bar{A}_j = a_j + \mathfrak{A}_j(x) \quad (j = 0, 1, \dots, m-1),$$

$$(14) \quad \bar{B}_{j-m} = b_{j-m} + \mathfrak{B}_{j-m}(x) \quad (j = m, m+1, \dots, n-1).$$

In view of the substitutions (4), the system (2) has therefore the unique solution

$$(15) \quad A_j = x^{[j]-1} (a_j + \mathfrak{A}_j(x)) \quad (j = 0, 1, \dots, m-1),$$

$$(16) \quad B_j = x^{[j+m]} (b_j + \mathfrak{B}_j(x)) \quad (j = 0, 1, \dots, \mu-1).$$

5. THE CHARACTERISTIC LINES OF A AND B

We have shown the existence of factors

$$A \equiv y^m + A_{m-1}(x) y^{m-1} + \dots + A_0(x),$$

$$B \equiv y^\mu + B_{\mu-1}(x) y^{\mu-1} + \dots + B_0(x)$$

satisfying the identity (1). It remains to show that the characteristic lines of A and B can be obtained from parts of the characteristic line of P by suitable translation.

We first recall that the characteristic line of P passes through or to the left of all points $([j], j)$ for $j = 0, 1, \dots, n-1$.

If the lowest segment of this line be translated l units to the left, it will pass through or to the left of all points $([j]-l, j)$ for $j \leq m$. From the analytic form of the coefficients (15) it follows therefore that the characteristic line of A will certainly not lie to the left of this translated segment. But it is easy to see that the ends (l, m) and $(k+l, 0)$ of the segment subjected to translation go into points of the characteristic line of A . The former point goes into $(0, m)$ which is obviously the upper end of the characteristic line. The latter goes into $(k, 0)$ which will be the lower end of this characteristic line provided $a_0 \neq 0$. This is true since a_0 is determined by the equation $a_0 b_0 = \bar{P}_0(0) = p_{k+l,0}$ and $p_{k+l,0} \neq 0$. The desired property relative to the characteristic line of A is thus established.

Consider now the part of the characteristic line of P above the point (l, m) . If it be translated m units downward it will pass through or to the left of all points $([j+m], j)$ for $j \leq \mu$. Hence from (16), no part of the characteristic line of B will lie to the left of its new position. Furthermore the characteristic line of B must coincide with it in its new position provided in (16) $b_j \neq 0$ for every j for which $([j], j)$ is a vertex of the characteristic line of P .

Suppose then that $([j], j)$ is a vertex for P . In this case, the inequality (8') may be altered to the form

$$[j] - [m+r] \leq (m+r-j) \frac{k_j}{m_j},$$

and as a result $\{j, r\} > 0$, at least for $j < m+r$. Hence when $([j], j)$ is a vertex, the form G_j in (11) contains no terms, and the solution of the corresponding equation is $b_{j-m} = \bar{P}_j(0) = p_{[j],j} \neq 0$. Hence the b_{j-m} of (14), that is the b_j of (16), is different from zero when $([j], j)$ is a vertex for P . This completes the proof of the theorem.

UNIVERSITY OF SASKATCHEWAN,
SASKATOON, CANADA.
