ON THE STRUCTURE OF DIFFERENTIAL POLYNOMIALS
AND ON THEIR THEORY OF IDEALS

BY
HOWARD LEVI

In the first part of this paper a special class of differential ideals\(^1\) is investigated. The results of this section are used in the following one to derive some structural properties of differential polynomials. The last part of the paper is devoted to a special differential ideal.

With the help of some conventions of notation, more precise indications of the scope of our work may be given. Let \(\mathcal{R}\) denote the ring of differential polynomials, with rational numbers for coefficients, in the unknown \(y\). The special class of differential ideals studied in Part I is composed of those generated by \(y^p\), where \(p\) is a positive integer. These ideals are among the most simple ideals encountered in the theory of differential equations. Viewed as algebraic entities, however, they are by no means trivial. We denote the \(i\)th derivative of \(y\) by \(y_i\); \(\mathcal{R}\) thus appears as a polynomial ring with infinitely many indeterminates \(y, y_1, y_2, \ldots\). Since the Hilbert basis theorem does not hold on \(\mathcal{R}\), one would expect almost any ideal in \(\mathcal{R}\) to be unruly. By introducing order relations into \(\mathcal{R}\) we have been able to proceed despite the absence of the basis theorem and to obtain fairly comprehensive results concerning these differential ideals. In particular a simple criterion for determining the membership in such an ideal of an element of \(\mathcal{R}\) is obtained which plays a fundamental role in Part II. This second part establishes the abstract counterparts of some results of J. F. Ritt concerning essential manifolds which figure in the decomposition of a manifold into irreducible ones. It has been found possible to present results which cover situations not discussed by him.

The differential ideal discussed in Part III is that generated by \(uv\), where \(u\) and \(v\) are unknowns. Among other properties, it is shown that this ideal has no representation as the intersection or product of two differential ideals, whose manifolds are respectively \(u = 0\) with \(v\) arbitrary, and \(v = 0\) with \(u\) arbitrary. This result owes its interest to the fact that the manifold of the equation \(uv = 0\) is evidently reducible into the union of the two manifolds just defined.

In a narrow sense, this paper is independent of other literature; the argu-

---

Presented to the Society February 22, 1941, under the title, On the ideal theory and structure of differential polynomials; received by the editors March 27, 1941.

ments make almost no appeal to outside sources. Less strictly, however, the writings of J. F. Ritt and H. W. Raudenbush, Jr., should be cited as furnishing both starting point and direction for this investigation. Indeed, the whole point of Part II lies in its connection with two papers of Ritt (more detailed reference is given in Part II). The congruence notation used in our work is that systematized by E. R. Kolchin\footnote{On the exponents of differential ideals, Annals of Mathematics, (2), vol. 42 (1941), p. 741.}. Brackets \([ \, \]\) and braces \({ \, \})\) mean respectively the differential ideal and the perfect differential ideal generated by the set of elements they include. The congruences

\[
a \equiv b \ [m, n, \cdots ]
\]

\[
a \equiv b \ {m, n, \cdots }
\]

mean, respectively, \(a - b\) is in \([m, n, \cdots ]\) and \(a - b\) is in \({m, n, \cdots }\). We use the term "form" exclusively as an abbreviation for the term "differential polynomial."

**Part I. The differential ideal generated by \(y^p\)**

**The form \(y^p\) and its derivatives**

1. Let \(p\) be any positive integer and let \(A = y^p\). We investigate the differential ideal \(\Sigma\) generated by \(A\). Denoting the \(i\)th derivative of \(A\) by \(A_i\), we see that \(\Sigma\) consists of all polynomials \(E_0 A + E_1 A_1 + \cdots + E_r A_r\), where the \(E_i\) are any elements of \(\mathbb{R}\). It is sometimes convenient to let \(y = y_0, A = A_0\).

We shall discuss power products in \(y\) and its derivatives, and make a few definitions for this purpose. Let \(P = y^{p_0} y_1^{q_1} \cdots y_r^{p_r}\) be such a power product, the \(p_i\) being non-negative integers. The degree of \(P\) is defined as \(\sum p_i\) and its weight is defined as \(\sum i p_i\). A power product \(P\) is different from a power product \(Q = y^{q_0} y_1^{q_1} \cdots y_r^{q_r}\) if some \(p_i - q_i\) is different from zero. We understand that if \(t > r\) then \(p_i\), the exponent of \(y_i\) in \(P\), is zero. If \(P\) is different from \(Q\) we say that \(P\) is *higher* than \(Q\), and \(Q\) is *lower* than \(P\), if the first nonzero difference \(p_i - q_i\) is positive. If \(P\) is higher than \(Q\) and \(R\) is any power product, then \(RP\) is higher than \(RQ\). If \(P\) is higher than \(Q\) and \(Q\) is higher than \(R\), then \(P\) is higher than \(R\). A power product \(P\) will be called an \(\alpha\) term, if \(p_i + p_{i+1} < p\); \(i = 0, 1, 2, \cdots\). Every factor of an \(\alpha\) term is an \(\alpha\) term. Every power product not an \(\alpha\) term will be called a \(\beta\) term. Every \(\beta\) term is divisible by an expression \(y^r y_{r+1}\) with \(r \leq p\).

2. The polynomial \(A_i\), the \(i\)th derivative of \(A\), is homogeneous of degree \(p\) and isobaric of weight \(i\). \(A_i\) is a sum over \(j\) of terms \(h_{ij} P_j\) where the \(h_{ij}\) are positive integers, and the \(P_j\) are power products of degree \(p\) and weight \(i\). Each power product \(P_j\) of this weight and degree is present in \(A_i\) with a coefficient \(h_{ij}\) different from zero. In particular if \(i = rp + s\) \((r\) and \(s\) non-negative integers and \(s < p\)) the term \(L_i = y^{rp+s} y_{r+1}\) has the proper weight and degree.
and is present effectively in $A_i$. It will be called the leader of $A_i$. We show that it is lower than any other term of $A_i$. It is certainly lower than any term involving a $y_k^{p_k}$ with $k < r$, $p_k > 0$. Any term of $A_i$ lower than $L_i$ would thus be of the form $y_k^{p_k-1} \cdots y_i^{p_i}$, with $q \leq p - s$. This would imply that $q_1 + q_2 + \cdots + q_i \geq s$. Thus the weight of such a term would be greater than $q(r + q_1 + \cdots + q_i) + q$ unless $q_2, \cdots, q_i$ were all zero. This last expression exceeds $i$ if $q < p - s$. It follows that any term of $A_i$ distinct from $L_i$ must be higher than $L_i$.

**Reduction of power products**

3. We prove the following lemma.

**Lemma 1.1.** For every $\beta$ term $F$ of $\mathcal{R}$ there is a congruence

$$F = \sum_i h_i P_i [\Sigma]$$

where the $P_i$ are $\alpha$ terms of the same weight and degree as $F$ and the $h_i$ are rational numbers (they may of course be zero).

$F$ is divisible by the leader $L_i$ of some $A_i$. Let $c_i$ stand for the coefficient of $L_i$ in $A_i$. Then $c_iL_i = A_i + (c_iL_i - A_i)$, where the terms in the parenthesis are higher than $L_i$, or are zero (if $i$ is zero or unity). If $F = c_iL_iF'$, then

$$F = F'A_i + F'((c_iL_i - A_i)$$

(1.1)

$$\equiv F'((c_iL_i - A_i) [\Sigma].$$

All the terms of the right member of this congruence are higher than $F$ and are of the same weight and degree as $F$. There may be some $\beta$ terms among them. Each such term is likewise congruent to a sum of higher terms of the same weight and degree. In particular the lowest $\beta$ term effectively present in (1.1) is congruent to such a sum. This term may be replaced in (1.1) by the appropriate combination of higher terms, yielding a new congruence for $F$ free of this $\beta$ term and all lower ones. Since there is only a finite number of power products of given weight and degree, this process eventually terminates; what remains in the right member is a linear combination of $\alpha$ terms with rational coefficients.

**Canonical representations**

4. The above lemma will be complemented by the fact established later that no linear combination of $\alpha$ terms with rational coefficients is in $\Sigma$ unless the coefficients are all zero. In addition, a canonical representation for the elements of $\Sigma$ will be obtained, in the following sense. Every element of $\Sigma$ has a representation $E_0A + \cdots + E_sA_s$ but the same element may have different representations. A simple example of this is given by the polynomial $2y^2y_1$ which is in the ideal generated by $y^2$ and may be written $2y_1A$ or $yA_1$. Our canonical representation will be obtained by choosing the coefficients $E_i$
from a restricted set of polynomials, with the result that these coefficients are uniquely determined, while still furnishing representations for every element of $\Sigma$.

5. In securing the canonical representation for the elements of $\Sigma$ we shall use forms $H$ of the types

$$H = EA_{i_0}A_{i_1} \cdots A_{i_s}$$

where $E$ is any power product in the $y_i$ and the other factors of $H$ constitute an arbitrary power product in $A$ and its derivatives. It is convenient to write this latter power product as above, without using exponents, in such a way that $i_0 \leq i_1 \leq \cdots \leq i_s$. $H$ is homogeneous and isobaric. Its degree is the degree of $E$ plus $(s+1)p$. Its weight is the weight of $E$ plus $i_0+i_1+\cdots+i_s$. Evidently $H$ is in $\Sigma$ and conversely every element of $\Sigma$ is a linear combination of such forms with constant coefficients. We order these forms in the following way: $H$ is higher than $H' = EA_{j_0}A_{j_1} \cdots A_{j_s}$ if either

(a) $A_{i_0}A_{i_1} \cdots A_{i_s}$ is higher than $A_{j_0}A_{j_1} \cdots A_{j_s}$ when both expressions are considered as power products in the $A_i$ and are compared by the method used for power products in the $y_i$, or

(b) $r = s$, $i_k = j_k$ ($k = 0, 1, \cdots, s$) and $E$ is higher than $E'$ in the sense previously explained.

It should be emphasized that what we order are the symbols used to denote the forms rather than the forms themselves. For instance, for $A = y^2$, criterion (a) implies that $H = A \cdot A$ is higher than $H' = y^3A$ even though $H$ and $H'$ both denote the same form $y^4$. Thus an expression $H$ is to be considered as different from $H'$ for purposes of ordering, if the set $(i_0, i_1, \cdots, i_s)$ is different from the set $(j_0, j_1, \cdots, j_r)$ or if $E$ is different from $E'$. We do not insist that the represented forms be different. On the other hand, equations connecting $H$, $H'$, $\cdots$, are to refer in the usual way to the forms denoted by the symbols.

Evidently, of two different expressions $H$, $H'$, one must be higher than the other. It is clear that our ordering is transitive. Furthermore, if $H$ is higher than $H'$ and $G$ is any power product in the $y_i$, then $GH$ is higher than $GH'$.

6. We now introduce the notion of a $\gamma$ term. An expression $III = EA_{i_0}A_{i_1} \cdots A_{i_s}$ will be called a $\gamma$ term if both (a) and (b) below hold.

(a) $i_s \geq sp$,

(a1) if $i_s > sp$ then $E$ is an $\alpha$ term in the letters $y_{r+1}$, $y_{r+2}$, $\cdots$,

(a2) if $i_s = sp$ then $E$ is any power product in the letters $y_r$, $y_{r+1}$, $\cdots$.

(b) If $s > 0$ then $i_k > kp$, $k = 0, 1, \cdots, s-1$.

7. The role of these $\gamma$ terms is revealed by the following lemma.

**Lemma 1.2.** Every expression $H$ is equal to a sum $\sum r_iR_i$, where the $R_i$ are $\gamma$ terms of the same weight and degree as $H$. The $r_i$ are rational numbers.
This implies that every element of \( \Sigma \) is a linear combination of \( \gamma \) terms with rational coefficients. We shall see later that such a sum is zero only if all the coefficients are zero.

Our proof will consist mainly in showing that if \( H \) itself is not a \( \gamma \) term it is a linear combination with constant coefficients of \( \gamma \) terms and of expressions \( H', H'', \ldots, H^{(i)} \); the expressions \( H^{(i)} \) all being higher than \( H \) and of the same weight and degree as \( H \). Once this is accomplished the proof can quickly be completed. By replacing the lowest \( H^{(i)} \) by its linear combination of \( \gamma \) terms and expressions \( H_f^{(i)} \) we obtain for \( H \) a new linear combination of \( \gamma \) terms and expressions \( H_f^{(0)} \) which is free of that lowest \( H^{(i)} \) and all lower ones. A finite number of repetitions of this procedure yields a linear combination of \( \gamma \) terms for \( H \).

8. To devise methods for obtaining this sum of higher terms we consider the obvious equality \( yA_1 = py_1A \) and those obtained by differentiating both members of this equation \( r \) times, \( r = 1, 2, \ldots \). We obtain

\[
yA_{r+1} + \sum_{i=1}^{r+1} C_{r,i} y_i A_{r+1-i} = p \sum_{i=1}^{r+1} C_{r,i-1} y_i A_{r+1-i}.
\]

The symbols \( C_{r,i} \) in equation (1.2) are binomial coefficients. This equation and the original equality together express \( yA_s \) with \( c > 0 \) as a sum \( \sum c_i y_i A_{s-i} \), where \( i \) runs through all positive integers not greater than \( s \). An analogous expression may be obtained for \( y_kA_s \), where \( k \) and \( s \) are positive integers and where \( s > kp \). Let \( r \) and \( k \) be positive integers with the property \( r+1-k > kp \). By subtracting the coefficient of \( y_kA_{r+1-k} \) in the right member of (1.2) from its coefficient in the left member we obtain

\[
\frac{r!}{(k-1)!(r-k)!} \left( \frac{1}{k} - \frac{p}{r+1-k} \right).
\]

This number is not zero, since \( r+1-k > kp \). Thus under these circumstances \( y_kA_{r+1-k} \) is effectively present in (1.2). Equations (1.2) show, then, that \( y_kA_s \) with \( s > kp \) may be written in the form

\[
\sum_{i=1}^{s} d_{ki} y_{k+i} A_{s-i} + \sum_{i=1}^{k} e_{ki} y_{k-i} A_{s+i}
\]

where the \( d_{ki} \) and \( e_{ki} \) are rational numbers which depend on \( s \) as well as on \( i \) and \( k \). Observe that in the first of these sums the subscript of each \( A_{s-i} \) is less than \( s \) and in the second the subscript of each \( y_{k-i} \) is less than \( k \).

9. We use these equations to derive a useful fact about expressions \( y_kA_{i_1}A_{i_2} \cdots A_{i_s} \) when \( s \) and \( k \) are non-negative integers with \( k \leq s \) and \( i_m > mp, \; m = 0, \ldots, s \). We show that such an expression is equal to a sum of certain products \( F_i G_j \) where the \( F_i \) are forms, and the \( G_j \) are power products in the \( A_i \). It will be seen that the degree of each \( G_j \) in the \( A_i \) does not
exceed \( s+1 \), and that each \( G_j \) as a power product in the \( A_i \) is higher than \( A_{i_0} \cdots A_{i_s} \). For \( s=0, i_0>0 \) we already know that \( y A_{i_0} \) is a sum of such products, namely \( c y A_{i_0-j} \). We establish the result for \( s>0 \) by induction. It follows from equations (1.2) that \( y A_{i_0} A_1 \cdots A_{i_s} \) is equal to

\[
\left( \sum_i \bar{d}_{ki} y_{k+i} A_{i_0} \right) A_{i_0} \cdots A_{i_{s-1}} + \left( \sum_i \bar{e}_{ki} y_{k-i} A_{i_{s+i}} \right) A_{i_0} \cdots A_{i_{s-1}}.
\]

The terms in the first group meet our requirements since each contains a power product in the \( A_i \) which is higher than \( A_{i_0} A_1 \cdots A_{i_s} \). Each term of the second group contains a factor \( y_{k-i} A_{i_0} \cdots A_{i_{s-1}} \). Assume the result true for all integers less than \( s \). Such a factor is then equal to a sum of products \( F_j G_j \) where the \( G_j \) are power products in the \( A_i \) which are higher than \( A_{i_0} \cdots A_{i_{s-1}} \) and whose degrees in the \( A_i \) do not exceed \( s \). Consequently the \( G_j \) are all higher than \( A_{i_0} A_1 \cdots A_{i_{s-1}} \). By letting \( F_j = e_{ki} A_{i_{s+i}} F_j \) it is seen that the terms in the second group are likewise equal to a sum of the required type.

10. We are now in a position to carry out the proof. An expression \( H \) which is not a \( \gamma \) term must fail to satisfy at least one of the conditions (a) and (b). We enumerate the various possibilities and show how for each one the required sum of higher terms may be obtained. It is both permissible and convenient first to discuss those terms \( H = E A_{i_0} \cdots A_{i_s} \) which satisfy (b) but not (a), and then to give a complete discussion for those expressions which do not satisfy (b). We follow this plan.

Suppose \( H \) satisfies (b), \( i_s > sp \) and (a\( \bar{1} \)) is not satisfied. If \( E \) involves only the letters \( y_{s+1}, y_{s+2}, \ldots, \) it must be a \( \beta \) term. It must then be divisible by the leader \( L_j \) of some \( A_j \). As in the proof of Lemma 1.1, we have \( E = c E' A_j + F \) where \( c \) is a constant and \( F \) is a form every one of whose terms is higher than \( E \). Consequently

\[
H = c E' A_{i_0} \cdots A_{i_s} A_j + FA_{i_0} \cdots A_{i_s}.
\]

The first term in the right member of this equation is the product with \( c \) of an expression \( H' \) which by criterion (a) above is higher than \( H \). The rest of the right member of this equation consists of a linear combination with rational coefficients of terms \( H^{(i)} \) all higher than \( H \) by criterion (b). We therefore have the required sum of higher terms.

If \( i_s = sp \), if (a\( \bar{1} \)) is not satisfied and if \( E \) contains effectively some letter \( y_k \) with \( k \leq s \), we first write \( E = y_k E' \). Because \( H \) is supposed to satisfy (b), we know that \( H' = y_k A_{i_0} \cdots A_{i_s} \) is a sum of products \( F_j G_j \) as described earlier. It follows that \( H \) is a sum of products \( (E' F_j) G_j \). Since the \( G_j \) are power products in the \( A_i \) whose degrees do not exceed \( s+1 \) and which are higher than \( A_{i_0} \cdots A_{i_s} \), it follows from criterion (a) that all the expressions \( H^{(i)} \) in each product \( (E' F_j) G_j \) are higher than \( H \). This disposes of expressions \( H \) which satisfy (b) but not (a\( \bar{1} \)). If an \( H \) satisfying (b) does not satisfy (a\( \bar{2} \)) and is such
that \( i_s = s \), then its coefficient \( E \) must contain effectively some letter \( y_k \) with \( k < s \). Let \( E = y_k E' \) and consider \( y_k A_{i_0} \cdots A_{i_{s-1}} \). It equals a sum of products \( F_i G_j \) where the \( G_j \) are power products in the \( A_i \) of degree not more than \( s \) and which are higher than \( A_{i_0} \cdots A_{i_{s-1}} \). What is important for us, is that the \( G_j \) are consequently also higher than \( A_{i_0} A_{i_1} \cdots A_{i_{s-1}} A_{i_s} \). It follows that \( H = (A_{i_0} E') y_k A_{i_0} \cdots A_{i_{s-1}} \) is a sum of higher expressions of the required sort.

There remains the case of an \( H \) which does not satisfy (b). Let \( r \) be the smallest integer for which \( i_r \leq r \) so that if \( r > 0 \) then \( i_m > m p, m = 0, \cdots, r-1 \). Our procedure depends on whether \( i_r = r p \) or \( i_r < r p \). If \( i_r = r p \) let \( E A_{i_{r+1}} \cdots A_{i_s} \) be expanded into a form \( F = \sum h_i E_i \), the \( h_i \) being constants and the \( E_i \) power products. We have \( H = (\sum h_i E_i) A_{i_0} \cdots A_{i_s} \). We consider the expressions \( H_i = E_i A_{i_0} \cdots A_{i_s} \), noting that they all satisfy (b). Certain of the \( E_i \) may be free of the letters \( y, y_1, \cdots, y_{r-1} \). For these \( E_i \) the corresponding \( H_i \) are \( \gamma \) terms and require no further discussion. If \( r = 0 \), all the \( E_i \) have this property and all the \( H_i \) are \( \gamma \) terms. On the other hand an \( E_i \) which contains effectively some \( y_k \) with \( k < r \) leads to an \( H_i \) which is not a \( \gamma \) term. Such an \( H_i \) satisfies (b) but not (a2). As we have seen, such an \( H_i \) is a sum of terms \( q_j H_j^{(0)} \) where the \( q_j \) are constants and the expressions \( H_j^{(0)} \) are all higher than \( H_i \). This of course does not itself imply that the \( H_j^{(0)} \) are higher than our original \( H \). But by recalling that the \( H_j^{(0)} \) must each contain a power product in the \( A_i \) which is higher than \( A_{i_0} A_{i_1} \cdots A_{i_s} \), and whose degree does not exceed \( r + 1 \), we see that the \( H_j^{(0)} \) are actually higher than \( H \) by criterion (a). Our procedure for an \( H \) which does not satisfy (b) and for which \( i_r < r p \) is the following. We note that \( r \) must be greater than zero, since \( i_s \) is non-negative. Let the form \( E A_{i_{r+1}} \cdots A_{i_s} \) be expanded as above into the form \( F = \sum h_i E_i \). The fact that \( i_r < r p \) implies that every term of \( A_{i_{r+1}} \), and consequently every term of \( F \), contains effectively some \( y_k \) with \( k < r \). Then \( H = (\sum h_i E_i) A_{i_0} \cdots A_{i_{s-1}} \) is a linear combination of expressions \( H_i = E_i A_{i_0} \cdots A_{i_{s-1}} \) which satisfy (b) but do not satisfy the requirement of (a1) which asks that \( E \) be free of \( y, y_1, \cdots, y_r \). It is easy to see how the methods of the previous case apply here, and we omit the details of showing that \( H \) must be a sum of terms \( h_j H_j^{(0)} \) where \( H_j^{(0)} \) is higher than \( H \).

11. We have now carried out our program of showing that each \( H \) not a \( \gamma \) term is a linear combination of expressions \( H_i \), those \( H_i \) which are not \( \gamma \) terms being higher than \( H \). The remarks made at the outset of the proof suffice to establish the lemma.

THE FUNDAMENTAL LEMMA

12. We prove the following lemma.

Lemmas 1.3. Let \( d \) and \( w \) be positive integers. The number \( n_\gamma \) of \( \gamma \) terms of degree \( d \) and weight \( w \) does not exceed the number of \( n_\beta \) of \( \beta \) terms which have this weight and degree.
It will be shown later that we actually have $n_\gamma = n_\beta$. We remind the reader that in computing $n_\gamma$ one counts the number of distinct symbols which stand for $\gamma$ terms without considering whether or not the symbols stand for distinct forms.

The proof will consist in associating a unique $\beta$ term of degree $d$ and weight $w$ with each $\gamma$ term of this degree and weight. The association will be such that to different $\gamma$ terms there will correspond different $\beta$ terms.

13. We require a few more definitions. Let $R_t$ denote the ring of polynomials with rational coefficients in the letters $y_i, y_{i+1}, \cdots, i=1, 2, 3, \cdots$, and let $R_0$ denote our original ring $R$. Let $t$ be any non-negative integer. A form $EA_t$ where $i \geq tp$ and $E$ is a power product in $y$ and its derivatives will be called an expression $K_t$ if the appropriate one of the following three conditions is satisfied by $E$.

(i) If $i = tp$, $E$ is any power product of $R_t$.
(ii) If $tp < i \leq (t+1)p$, $E$ is any power product of $R_{t+1}$. Let $E = y_{t+1}^{a_1} \cdots y_{t+k}^{a_k}$.
(iii) If $(t+1)p < i$, $E$ is a special power product of $R_{t+1}$. Let $E = y_{t+1}^{a_1} \cdots y_{t+k}^{a_k}$.

We ask that there exist an integer $k$ for which $E' = y_{t+1}^{a_1} \cdots y_{t+k}^{a_k}$ is an $\alpha$ term, and in addition we require for this $k$ that $i \leq (t+k+1)p - (a_1 + \cdots + a_k)$.

Under condition (iii) any $\alpha$ term of $R_{t+1}$ is acceptable as a coefficient $E$, for $(t+k+1)p$ increases with $k$, whereas for large $k$ the exponent $a_k$ is zero, so that $a_1 + \cdots + a_k$ remains unchanged. On the other hand under condition (iii) an admissible $E$ need not be an $\alpha$ term. Once a suitable $k$ is found, no restriction whatever is made on the letters $y_{t+k+1}, y_{t+k+2}, \cdots$.

14. We now describe a process by which each such expression $K_t$ determines a $\beta$ term $F$ of $R_t$. Let $K = EA_t$ be a definite expression $K_t$. It comes under one of (i), (ii), (iii).

If $K$ comes under (i) we have $i = tp$. Let

$$F = y_i^p E.$$  

Clearly $F$ is a $\beta$ term of $R_t$ having the same weight and degree as $K$.

If $K$ comes under (ii) then $tp < i \leq (t+1)p$. Let $b = (t+1)p - i$. Then $b$ is a non-negative integer and $b < p$. Let

$$F = y_i^{b} y_{t+1}^{p-b} E.$$  

$F$ is a $\beta$ term of $R_t$ obtained by replacing $A_t$ in $K$ by the term $y_i^{b} y_{t+1}^{p-b}$. The degree of this term is $p$ and its weight is $(t+1)p - b = i$. Thus $F$ has the same weight and degree as $K$. For this case $E$ does not contain the letter $y_i$ so that the exponent of $y_i$ in $F$ is $b$ which is less than $p$. This distinguishes the $F$ obtained from a $K$ which comes under (ii) from that obtained from a $K$ which comes under (i). However, for both cases the sum of the exponents of $y_i$ and $y_{t+1}$ in $F$ is at least $p$.  

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
If $K$ comes under (iii) then $(t+1)p < i$. We define

$$s_0 = (t + 1)p, \quad s_f = (t + f + 1)p - (a_1 + \cdots + a_f), \quad f = 1, \cdots, k.$$ 

We have $s_f - s_{f-1} = p - a_f$, $f = 1, \cdots, k$. By hypothesis $\gamma_{t+1}^i \cdots \gamma_{t+k}^i$ is an $\alpha$ term and in particular each $a_f$ is less than $p$. Thus each $p - a_f$ is positive, so that $s_0 < s_1 < \cdots < s_k$. Since by hypothesis $i \leq s_k$ there is an integer $m$ which is such that $1 \leq m \leq k$ and for which $s_{m-1} < i \leq s_m$. Let $b = s_m - i$. Then $b$ is a non-negative integer. Since $b < s_m - s_{m-1}$ and $s_m - s_{m-1} = p - a_m$ we have $b + a_m < p$. Let

$$(1.3) \quad F = y_t^{a_1} \cdots y_{t+m-1}^{a_m} y_{t+m}^b (y_{t+m}^{p-b}) G$$

where $G = y_{t+m+1}^{a_{m+1}} \cdots y_{t+m+t}^{a_t}$. $F$ is a $\beta$ term of $R_t$. In the transition from $K$ to $F$ the expression $A_i$ is suppressed, $\gamma_{t+j}^i$ is replaced by $\gamma_{t+j-1}^i$, $j = 1, \cdots, m$, and the term $y_{t+m}^{p-b}$ is introduced. $G$ is carried over unchanged. The first operation lowers the weight by $i$ and the degree by $p$. The second lowers the weight by $a_1 + \cdots + a_m$ and does not change the degree. The introduction of $y_{t+m}^{p-b}$ augments the degree by $p$ and the weight by $(t + m + 1)p - b$. Since

$$(1.4) \quad (t + m + 1)p - b = s_m + (a_1 + \cdots + a_m) - b = i + (a_1 + \cdots + a_m)$$

we see that the net effect of these alterations is to leave the weight and degree unchanged. Note that $F$ contains the factor $y_t^{a_1} \cdots y_{t+m-1}^{a_m} y_{t+m}^b$ and that its other letters all have subscripts which exceed $t + m$. This factor is an $\alpha$ term, since $b + a_m < p$ and $y_t^{a_1} \cdots y_{t+m-1}^{a_m}$ is an $\alpha$ term by hypothesis. Since $m$ is positive, it follows that the sum of the exponents of $y_t$ and $y_{t+1}$ in $F$ is less than $p$. This is a characteristic property of a term $F$ obtained from an expression $K$ which comes under (iii).

15. We have described a procedure for obtaining from any expression $K_1$ a definite $\beta$ term $F$ of $R_t$. We shall investigate this procedure further in order to obtain two useful facts. The first is that by this process different $\beta$ terms $F$ are assigned to different expressions $K$. The second is that when $t > 0$, then for any integer $h$ such that $(t - 1)p < h \leq i$, the expression $FA_h$ is an expression $K_{t-1}$. In other words if $F$ is obtained from any $K_1$ in the manner set forth above, then, if $t > 0$ and $h$ is as above, $FA_h$ admits one of the three characterizations, (i), (ii), (iii), where the discussion is referred to the integer $t-1$ instead of $t$. In deriving the first property of the term $F$ we need only show that a $\beta$ term $F$ cannot be obtained from two different expressions $K_1$ which both come under the same condition of the three listed. This simplification is due to the fact that in describing the procedure it was pointed out how one could infer from a given $F$ which of the three conditions governed the $K$ which determined it. We now list the three possibilities for $F$ and verify the two statements for each one.
16. Let $F$ be determined by an expression $K = EA_i$ which comes under (i). Then $F = y_t^m E$, so that given $F$ one can find $E$. Since for this case $i = tp$, there is only one $EA_i$ which could lead to $F$. This proves the first statement. To prove the second, let $h$ be any integer such that $(t-1)p < h \leq i$. We can easily verify that $FA_h$ is an expression $K_{t-1}$, coming under (ii). In the case at hand $i = tp$, so that we have $(t-1)p < h \leq tp$. In addition $F$ is a power product of $R_t$. These are precisely the requirements of condition (ii).

17. Now let $F$ be determined by $K = EA_i$, $K$ coming under (ii). We have $F = y_t^m y_{t+1}^m E$ with $0 \leq b < p$ and with $E$ free of $y, y_1, \ldots, y_t$. Again it is obvious that $F$ determines $E$ uniquely and that the subscript $i$ of $A_t$ can also be uniquely determined from the equation $(t+1)p - b = i$. Thus only one expression $EA_i$ can yield $F$ by our procedure. Suppose that $t > 0$ and that $h$ is some integer for which $(t-1)p < h \leq i$. If $h \leq tp$ then $FA$ is an expression $K_{t-1}$ coming under (ii), since $F$ is in $R_t$. If $h > tp$ we show that $FA_h$ is also an expression $K_{t-1}$ but that it then comes under (iii). The inequality $h \leq i = (t+1)p - b$ enables us to draw this conclusion. The $\alpha$ term required by (iii) is simply $y_t^m$; the integer $k$ is unity.

18. The case in which $K = EA_i$ comes under (iii) remains. The $F$ which it determines is displayed in (1.3). We noted above that $F$ contains as a factor the $\alpha$ term $y_t^m \cdots y_{t+m-1}^m y_{t+m}$. In an obvious sense this factor is the "largest" $\alpha$ term which can be split off from $F$. More precisely, given an $F = y_t^m y_{t+1}^m \cdots y_{t+m-1}^m y_{t+m}$ determined by an expression $K_t$ which comes under (iii), if one chooses the largest $g$ such that $y_t^m y_{t+1}^m \cdots y_{t+g}^m$ is an $\alpha$ term, this last power product will be identical with $y_t^m \cdots y_{t+m-1}^m y_{t+m}$. We recall that in passing from $K = EA_i$ to $F$ we divided the letters of $E$ into two classes; the letters of one class were replaced by others, and the letters in the other were carried over unaltered. What we have just shown is that given an $F$ determined by a $K$ which comes under (iii) it is possible to determine exactly which letters were in each of the classes. The weight of the $A_i$ involved in $K$ may be computed from

\[ i = (t + m + 1)p - (a_1 + \cdots + a_m) - b. \]

Thus, given such an $F$ it is possible to reconstruct unequivocally the expression $K$ from which it was obtained. This establishes the first property for a $K$ coming under (iii). Assuming now that $t > 0$, we proceed to establish the second. Let $h$ be an integer such that $(t-1)p < h \leq i$. We show that $FA_h$ is an expression $K_{t-1}$. If $h \leq tp$, $FA_h$ is clearly an expression $K_{t-1}$ coming under (ii), since $F$ is a power product of $R_t$. If $h > tp$ we show that $FA_h$ is an expression $K_{t-1}$ coming under (iii). To do this we must produce an $\alpha$ term and an integer $k$ as described in (iii). Let $F = y_t^m \cdots y_{t+k}^m$. Since our calculations are now based on the integer $t-1$, the integer $k$ is required to have the property $(t+k)p - (b_0 + \cdots + b_{k-1}) \geq h$. Since $h \leq i$, we see from (1.4) that the $\alpha$ term $y_t^m \cdots y_{t+m-1}^m y_{t+m}$ and the integer $m+1$ have the required properties.
19. The proof of the lemma may now be completed. Let \( H = \sum_{i=0}^{t} A_i \) be any \( \gamma \) term of degree \( d \) and weight \( w \). \( H \) satisfies conditions (a) and (b) defining a \( \gamma \) term. We now show how these conditions make it possible to use the work immediately preceding to carry out our program of assigning a \( \beta \) term to every \( \gamma \) term.

Consider the form \( K^{(s)} = \sum_{i=s}^{t} A_i \). If \( i_s = s \) then \( K^{(s)} \) is an expression \( K_s \), coming under (i). If \( i_s > s \) it is readily seen that \( K^{(s)} \) is likewise an expression \( K_s \), only in this case it comes under (iii). In fact condition (a) requires \( E \) to be an \( \alpha \) term of \( \mathcal{R}_{s+1} \) and it was pointed out above that an integer \( k \) of the type required by (iii) can always be found under these circumstances. Thus by splitting off \( \sum_{i=s}^{t} A_i \) from a \( \gamma \) term \( H \) we always obtain an expression \( K_s \). Let the weight of \( E \) be \( w_s \) and its degree be \( d_s \). \( K^{(s)} \) determines a \( \beta \) term of \( \mathcal{R}_s \) by the procedure described above. Let it be denoted by \( E^{(s)} \). Its weight is \( w_s + i_s \) and its degree is \( d_s + p \). If \( s=0 \) we associate this \( \beta \) term with \( H \). It has the same weight and degree as \( H \) because it has the same weight and degree as \( K^{(s)} \) and for this case \( K^{(s)} = H \).

If \( s>0 \) consider the expression \( K^{(s-1)} = \sum_{i=s-1}^{t} A_i \). It follows from the definition of \( \gamma \) term that \( (s-1)p < i_{s-1} \leq i_s \). This inequality permits us to conclude that \( K^{(s-1)} \) is actually an expression \( K_{s-1} \). \( K^{(s-1)} \) determines a \( \beta \) term \( E^{(s-1)} \) of \( \mathcal{R}_{s-1} \) having the same weight \( w_{s-1} + i_{s-1} \) and the same degree \( d_{s-1} + 2p \) as \( K^{(s-1)} \). If \( s=1 \) we associate this \( \beta \) term with \( H \). It clearly has the same weight and degree as \( H \).

If \( s>1 \) we continue in this way. We obtain a sequence \( K^{(s)}, K^{(s-1)}, \ldots, K^{(0)} \) and a sequence \( E^{(s)}, E^{(s-1)}, \ldots, E^{(0)} \). The sequences are obtained from \( H \) by successive applications of the procedure described for obtaining \( \beta \) terms from expressions \( K_i \). Each \( K^{(f)} = \sum_{i=f+1}^{t} A_i \). Each \( E^{(f)} \) is the \( \beta \) term of \( \mathcal{R}_f \) determined by \( K^{(f)} \). The weight of both \( K^{(f)} \) and \( E^{(f)} \) is \( w_f + i_f \). The degree of both \( K^{(f)} \) and \( E^{(f)} \) is \( d_f + p \). The sequences are to be continued until \( K^{(0)} \) and \( E^{(0)} \) are reached. \( E^{(0)} \) is a \( \beta \) term of \( \mathcal{R} = \mathcal{R}_0 \) having the same weight and degree as \( H \). We associate \( E^{(0)} \) with \( H \).

20. We now prove that if \( H_1 = \sum_{i=0}^{r} A_i A_j \) is a \( \gamma \) term different from \( H \), then the \( \beta \) term \( E^{(0)}_1 \) assigned to it in this way must be different from \( E^{(0)} \). \( H_1 \) determines the two sequences \( K_1^{(r)}, K_1^{(r-1)}, \ldots, K_1^{(0)} \) and \( E_1^{(r)}, E_1^{(r-1)}, \ldots, E_1^{(0)} \). We know that each \( E^{(f)} \) is determined by at most one \( K^{(f)} = \sum_{i=f+1}^{t} A_i \). We conclude that if \( E^{(0)} = E_1^{(0)} \), then for every \( f \) for which the symbols are defined, \( E^{(f)} = E_1^{(f)} \) and \( K^{(f)} = K_1^{(f)} \). If \( s=r \) we have immediately that \( H = H_1 \). Suppose \( s \neq r \) and, say, \( s<r \). We show that it is impossible to have \( E^{(0)} = E_1^{(0)} \) under this assumption. This last equality implies that \( K^{(0)} = K_1^{(0)} \), that is, that \( \sum_{i=0}^{r} A_i = \sum_{i=s+1}^{r} A_i \). \( E^{(s+1)} \) is determined by \( K_1^{(s+1)} \) and is consequently a \( \beta \) term. From the definition of \( \gamma \) term \( E \) may only be a \( \beta \) term if \( i_s = s \). Thus \( j_s = s \). This is impossible since \( H_1 \) is a \( \gamma \) term and for such terms we have \( j_f = fp \) only if \( f = r \), whereas here we have \( j_s = s \) and \( s<r \).

We have shown that every \( \gamma \) term determines a \( \beta \) term of \( \mathcal{R} \) having the
same weight and degree, and that distinct $\gamma$ terms determine distinct $\beta$ terms. This proves the lemma.

**The structure of the ideal of $\gamma^p$**

21. We can now prove the following lemma.

**Lemma 1.4.** Let $d$ and $w$ be positive integers. Let $n_\alpha$ denote the number of $\alpha$ terms $P_i$ of degree $d$ and weight $w$, let $n_\beta$ denote the number of $\beta$ terms $Q_j$ of this degree and weight, and let $n_\gamma$ denote the corresponding number of $\gamma$ terms $R_k$. Then $n_\beta = n_\gamma$, and a relation

$$\sum_{i=1}^{n_\alpha} p_i P_i + \sum_{k=1}^{n_\gamma} r_k R_k = 0$$

where the $p_i$ and $r_k$ are rational numbers implies that all the $p_i$ and $r_k$ are zero.

Let $Q_j$ be any $\beta$ term mentioned in the statement of the lemma. By Lemma 1.1 we have

$$Q_j = \sum_{i} p_{ji} P_i [\Sigma], \quad i = 1, \ldots, n_\beta.$$  

Therefore by Lemma 1.2 we have

$$Q_j = \sum_{k} r_{kj} R_k + \sum_{i} p_{ji} P_i, \quad i = 1, \ldots, n_\beta.$$  

If $n_\gamma$ were less than $n_\beta$ some linear combination of the $Q_j$ with rational coefficients would be a similar linear combination of the $P_i$. This is impossible, so that $n_\gamma \geq n_\beta$. Applying Lemma 1.3 we see that $n_\gamma = n_\beta$.

22. Every $R_k$ is by definition a homogeneous isobaric polynomial, so that we have

$$R_k = \sum_{i} a_{ki} P_i + \sum_{j} b_{kj} Q_j, \quad k = 1, \ldots, n_\gamma.$$  

Substituting the right member of (1.6) for $R_k$ in (1.5), we obtain the identities $Q_j = Q_j$, $j = 1, \ldots, n_\beta$. Thus $|r_{jk}| : |b_{kj}| = 1$ and it follows that both $|r_{jk}|$ and $|b_{kj}|$ are different from zero.

From (1.6) and the fact that $|b_{kj}| \neq 0$ we see that any linear combination of the $R_k$ with constant coefficients not all zero must equal a similar linear combination of the $P_i$ and $Q_j$ which involves some $Q_j$ effectively. If such a linear combination of terms $R_k$ were expressible as a linear combination of terms $P_i$, we would have the absurd result that a linear combination of the $Q_j$ with constant coefficients not all zero was a linear combination of the $P_i$.

23. **Theorem 1.1.** Let $F$ be any element of $\mathcal{R}$. $F$ is expressible in the form

$$F = \sum_{i} p_i P_i + \sum_{k} r_k R_k.$$  

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
where the $P_i$ are $\alpha$ terms, the $R_k$ are $\gamma$ terms and the $p_i, r_k$ are rational numbers\(^{(3)}\). For each $F$ there is only one such expression.

Let $F$ be split up into a sum of homogeneous isobaric polynomials $F_h$ in such a way that any two such polynomials have either different weights or different degrees. Then

$$F = \sum_h F_h$$

and the $F_h$ are uniquely determined by $F$.

By means of equation (1.5) we have

$$F_h = \sum_i p_{hi}P_i + \sum_k r_{hk}R_k$$

where the $d$ and $w$ of Lemma 1.4 are the degree and weight of $F_h$. By adding the $F_h$ we obtain

$$F = \sum_h \sum_i p_{hi}P_i + \sum_k \sum_k r_{hk}R_k,$$

and this is the desired equation.

24. To show that $F$ does not have two distinct representations of this type we need only show that if

$$\sum_i p_iP_i + \sum_k r_kR_k = 0$$

(1.7)

then the $p_i$ and $r_k$ are all zero.

Let some $p_i$ or $r_k$ be different from zero. The term which possesses such a coefficient must be cancelled in (1.7) by a sum of other terms of the same weight and degree. This contradiction to Lemma 1.4 establishes our result.

25. **Corollary.** No linear combination of $\alpha$ terms with rational\(^{(4)}\) coefficients is in $\Sigma$.

Let $d, w, n_\alpha$ be as in the statement of Lemma 1.4. The quantity $n$ depends on $d$ and $w$.

**Corollary.** The number of linearly independent (mod $\Sigma$) elements of $\mathcal{R}$ which are homogeneous of degree $d$ and isobaric of weight $w$ is $n_\alpha$.

26. $\mathcal{R}$ may be considered as an abelian group with operators, where the group "multiplication" is ordinary addition, and the operators are rational numbers. Theorem 1.1 implies that $\mathcal{R}$ considered in this way is the direct sum of two groups. One of them is $\Sigma$. The other is the additive group generated by

\(^{(3)}\) The same conclusion can be drawn if these symbols stand for any constants, or more generally if they are any elements of a domain of integrity which contains the rational numbers.

\(^{(4)}\) See the note to Theorem 1.1.
the totality of $\alpha$ terms. A linearly independent basis for the first group is the totality of $\gamma$ terms; the totality of $\alpha$ terms forms such a basis for the other.

27. We now determine circumstances under which the $n_\alpha$ of the corollary to Theorem 1.1 is zero. If for a given $d$ and $w$ this number is zero, then every homogeneous isobaric element of $\mathcal{R}$ having this degree and weight is in $\Sigma$. In settling this question we consequently develop a method for establishing the membership in $\Sigma$ of certain elements of $\mathcal{R}$ based entirely on an examination of the weights and degrees of their constituent terms.

If $d$ is less than $p$, every power product of degree $d$ is an $\alpha$ term. To treat the case for which $d$ is not less than $p$ we write

$$S = y_{p-1} y_{p-2} \cdots y_{p-k} y_{2k+2} \cdots.$$ 

$S$ is a formal infinite product whose status in this discussion is that of a visual aid. Let $\delta$ be a positive integer and write $d = a(p-1) + b$ $(a$ and $b$ non-negative integers with $0 < b \leq p - 1)$. Let

$$S_d = y_{p-1} y_{p-2} \cdots y_{2a-2} y_{2a}.$$ 

$S_d$ is an $\alpha$ term of degree $d$. It is obtained by taking the first $d$ letters of $S$ and multiplying them together. We denote the weight of $S_d$ by $w(p, d)$. We have

$$w(p, d) = a(a - 1)(p - 1) + 2a\delta.$$ 

$w(p, d)$ is defined for all integers $p$ greater than unity, and for all positive integers $d$. Its values are always positive integers or zero. An easy calculation shows that $w(p, d)$ satisfies the difference equation $w(p, d) + 2d = w(p, d + (p - 1))$, and this fact is used in proving the following result.

28. Theorem 1.2. Let $d$, $w$ and $n_\alpha$ be as in the statement of Lemma 1.4. A necessary and sufficient condition that $n_\alpha > 0$ is that $w \leq w(p, d)$.

In view of our earlier results this is equivalent to asserting that every power product of degree $d$ and weight $w < w(p, d)$ is in $\Sigma$ and not every power product of degree $d$ and weight $w \geq w(p, d)$ is in $\Sigma$.

29. The sufficiency proof is quickly disposed of. $S_d$ is an $\alpha$ term of degree $d$ and weight $w(p, d)$. Let

$$S_d^{(r)} = y_{p-1} y_{p-2} \cdots y_{2a-2} y_{2a} y_{2a+r}, \quad r > 0.$$ 

(4) In the remainder of Part I it is assumed that $p$ exceeds unity. The two results enunciated there are seen to be trivially true for $p$ equal to unity, if the weight function introduced at the end of §27 is defined to be plus infinity for $p$ equal to unity and for all positive integral values of $d$.
$S_d^\alpha$ is an $\alpha$ term of degree $d$ and weight $w(p, d) + r$. Therefore we see that for any integer $d$ and integer $w \geq w(p, d)$ there are $\alpha$ terms of degree $d$ and weight $w$.

30. We begin the necessity proof by observing that when $d$ is less than $p$, $w(p, d) = 0$. Consequently there are no power products of degree $d < p$ and weight $w < w(p, d)$. If our theorem were false there would be an integer $d \geq p$, and an $\alpha$ term whose degree was $d$ and whose weight was less than $w(p, d)$. We assume this to be the case and force a contradiction.

Let $d \geq p$ be the smallest integer for which there are $\alpha$ terms whose degree is $d$ and whose weight is less than $w(p, d)$. Let $P$ be an $\alpha$ term of degree $d$ and weight $w$, where $w$ is some integer such that $0 \leq w < w(p, d)$. Let $P = EP'$, where $E$ is that factor of $P$ of degree $p - 1$ which is higher than any other such factor. Then $P'$ is of positive degree and is an $\alpha$ term. Furthermore, since $P$ is an $\alpha$ term the definition of $E$ insures that $P'$ is free of $y$ and $y_j$. Let the degree of $P'$ be denoted by $d'$ and its weight by $w'$. Clearly $w' \leq w$ and $d' < d$. Let $P''$ be obtained from $P'$ by replacing each letter $y_j$ effectively present in $P'$ by $y_{j-2}$. $P''$ is an $\alpha$ term whose degree is $d'$ and whose weight is $w' - 2d'$. Our assumption about the minimal character of $d$, when applied to $P''$ implies

$$w(p, d') \leq w' - 2d'.$$

Using the difference equation satisfied by $w(p, d')$ this last inequality yields $w(p, d' + p - 1) \leq w'$. Since $d' + p - 1 = d$ and $w' \leq w$ we now have $w(p, d) \leq w$. This contradiction completes the proof.

The expression for a power product in the ideal of $y^p$

31. Having established the fact that certain power products are in $\Sigma$ we may naturally inquire as to the number of derivatives of $A$ needed to express them. This question may be precisely formulated in the following way. Let a power product $P$ be in $\Sigma$. It is a linear combination of $\gamma$ terms. Let the lowest of these be the $\gamma$ term $EA_{i_0} \cdots A_{i_n}$. It is required to determine an upper bound for $i_0$. This question arises in the following section for a special class of power products. We settle it now for this special class.

Corollary. Let $r$ be a positive integer, let $d = (r+1)p - 1$ and let $w$ be a non-negative integer which does not exceed $rd$. Then every power product $P$ of degree $d$ and weight $w$ is in $\Sigma$ and is a linear combination, with forms for coefficients, of $A$ and its derivatives of orders not exceeding $rp$.

Let $d' = (r+1)(p-1) + 1$. Certainly $d' \leq d$. We first extract from $P$ a factor $P'$ of degree $d'$ whose weight does not exceed $rd'$. This is made possible by the fact that the weight of $P$ does not exceed $rd$. We then show that $P'$, and hence $P$, is in $\Sigma$. Evaluation of $w(p, d')$ yields $r(r+1)(p-1) + 2(r+1)$ which exceeds $rd'$. This shows that $P'$ is in $\Sigma$. Actually for large $r$ the weight of $P$ is considerably smaller than $w(p, d)$. This additional restriction makes it pos-
sible to estimate relatively easily the number of derivatives of $A$ required to express $P$.

The proof is by induction. When $r = 1$, our assertion is that no more than $p$ derivatives of $A$ are required to obtain $P$. Any $\gamma$ term of degree $2p - 1$ is of the form $A_i E$. If $i$ is zero no discussion is required. If $i$ is greater than zero, $E$ must be a power product of degree $p - 1$ in the letters $y_1, y_2, \cdots$. The weight of $E$ is then at least $p - 1$. If the weight of the $\gamma$ term is not to exceed $2p - 1$ it must be that $i \leq p$. Assume now that the result is established for all integers less than some fixed integer $r$. Let $G = EA_{i_0}A_{i_1} \cdots A_{i_r}$ be a $\gamma$ term of degree $d = (r + 1)p - 1$ and weight not greater than $rd$. It is to be shown that $i_0 \leq rp$. We need only consider the case in which $i_0 > 0$. For this case $E$ must be free of $y$, since $G$ is a $\gamma$ term. Let $E'$ be obtained from $E$ by diminishing by unity the subscript of each $y_i$ which appears in $E$. If $s > 0$, consider $G' = E'A_{i_0-p} \cdots A_{i_s-p}$. $G'$ is evidently a $\gamma$ term of degree $d' = d - p$ and weight $w' = w - i_0 - d + p$. If $i_0$ exceeded $rp$ we should then have, using $w \leq rd$, $w' < rd - rp - d + p$ or $w' < (r - 1)d'$. Our induction hypothesis then applies to $G'$ and shows that $i_1 - p \leq (r - 1)p$ whence $i_1 \leq rp$. Since $i_0 \leq i_1$, the assumption $i_0 > rp$ leads to a contradiction. If in $G$ the integer $s$ is zero, so that $G = EA_{i_0}$, a different procedure is required. $E$ must be an $\alpha$ term in the letters $y_1, y_2, \cdots$, of degree $d - p$ and weight $w - i_0$. Consequently $E'$ is an $\alpha$ term of degree $d' = d - p$ and weight $w' = w - i_0 - d + p$. Again assume that $i_0 > rp$. Using $w \leq rd$ and $i_0 > rp$ we have $w' < (r - 1)d'$. Since $d'$ is $rp - 1$ it follows as in the outset of this proof that the weight of $E'$ is too small for $E'$ to be an $\alpha$ term. The hypothesis $i_0 > rp$ must then be discarded and the induction is carried out.

**Part II. Some theorems on the structure of differential polynomials**

**The low power theorem**

32. Let $\mathfrak{S}$ be any differential domain of integrity which contains the rational numbers. Throughout this section when we refer to a form in the unknowns $u, v, \cdots, w$, we shall mean a differential polynomial in $u, v, \cdots, w$ whose coefficients are in $\mathfrak{S}$. Indeed the coefficients actually used are for the most part rational numbers. Our work involves auxiliary unknowns which may be specialized with great freedom, and it is with such specialization in mind that the above remarks are made.

33. **Theorem 2.1.** Let

$$F = \lambda y^p - \sum_{i=1}^{k} u_i B_i$$

be a form in $y, \lambda, u_1, \cdots, u_k$ where $p$ is a positive integer and the $B_i$ are power products in $y$ and its derivatives of degree $p + 1$. Then there is a positive integer $s$ and a form

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
\[ D = \lambda^s + H \]

where every term of \( H \) contains \( y \) or one of its derivatives effectively and where \( D \) is homogeneous of degree \( s \) in \( \lambda, u_1, \ldots, u_k \) and their derivatives, such that for some positive integer \( d \),

\[ y^d D \equiv 0 \ [F]. \]

The questions as to how large \( d \) and \( s \) need be, and how many derivatives of \( F \) are required to obtain \( y^d D \) are not answered precisely, but in the proof explicit upper bounds are given for each of these numbers.

34. This theorem is the abstract counterpart of certain results obtained by Ritt. It might be appropriate to discuss this relation before taking up the proof. If \( y, \lambda \) and the \( u_i \) are replaced by forms \( Y, L, U_i \) in the unknowns \( v_1, \ldots, v_n \), then \( F \) and \( D \) go over into forms \( F' \) and \( D' \) in the \( v_i \) and \( y^d D' = 0 [F'] \). Let us take \( Y \) to be an algebraically irreducible form whose order in \( v_n \) is \( h \), \( L \) to be a nonzero form not divisible by \( Y \) whose order in \( v_n \) does not exceed \( h \), and the \( U_i \) to be any forms. Then, when \( 3 \) is a ring of analytic functions, the relation \( y^d D' = 0 [F'] \) shows that the general solution of \( Y \) is an essential irreducible manifold in the manifold of \( F'(6) \). A somewhat different result is obtained by specializing \( y \) as \( v_i, \lambda \) as a form \( 1 + L \) where \( L \) vanishes for \( v_i = 0, i = 1, \ldots, n \), and the \( u_i \) as any forms in the \( v_j \). \( 3 \) is again a ring of analytic functions. The relation \( v_i^d D' = 0 [F'] \) shows that the solution \( v_i = 0, i = 1, \ldots, n \), of \( F' \) is not contained in any irreducible manifold held by \( F' \) but not by \( v_1(7) \).

35. We now take up the proof. Let \( r \) be the maximum of the weights of the \( B_i \). If \( r \) is zero or unity, each \( B_i \) is divisible by \( y^p \), and \( F \) itself may be factored into a product \( y^p (\lambda + H) \) of the required type. Assuming now that \( r > 2 \), let \( d \) be an integer such that the set of all power products in \( y \) and its derivatives of degree \( d \) whose weight does not exceed \( (r - 1)d \) is in the differential ideal generated by \( A = y^p \). The work of the preceding section proves that there are such integers \( d \). Let these power products be denoted by \( P_1, \ldots, P_m \), and let the weight of \( P_f \) be denoted by \( w_f, f = 1, \ldots, m \). Suppose all the \( P_f \) may be expressed as linear combinations of \( A \) and its derivatives of order not exceeding \( t \). Then

\[ P_f = \sum_{j=0}^t C_{fj} A_j, \quad f = 1, 2, \ldots, m, \]

where the \( C_{fj} \) are forms in \( y \). They are homogeneous of degree \( d - p \) and isobaric of weight \( w_f - j \). Using the fact that \( \lambda^{s+1} A_q \) is in the differential ideal


\(^{(7)}\) J. F. Ritt, *On certain points in the theory of algebraic differential equations*, American Journal of Mathematics, vol. 60 (1938), p. 9. This paper will be referred to as OCP.
generated by $\lambda A$ and that it is a linear combination of $\lambda A$ and its first $q$ derivatives\(^{(8)}\), we have for each $P_f$

$$\lambda^{t+1}P_f = \sum_{j=0}^{t} \sum_{g=0}^{q} C_{fq}L_g(\lambda A)_{j-g}.$$ 

Here the symbol $(\lambda A)_{j-g}$ means the $(j-g)$th derivative of $\lambda A$, and $L_g$ is a homogeneous form in $\lambda$ of degree $t$. Referring to (2.1) we have for each $P_f$

$$\lambda^{t+1}P_f = \sum_{j=0}^{t} \sum_{g=0}^{q} C_{fq}L_g \left( \sum_{i=1}^{k} u_iB_i \right)_{j-g} \ [F]$$

where $(\sum u_iB_i)_{j-g}$ means the $(j-g)$th derivative of the form inside the parentheses. The right member of this congruence is a linear combination of forms $T = C_{kq}(B_i)_{k}, h \leq j$, whose coefficients are homogeneous forms in $\lambda$ and the $u_i$. These forms $T$ are homogeneous of degree $d+1$, since the $B_i$ are all of degree $p+1$. The weight of each $(B_i)_k$ does not exceed $r+h$, and therefore the weight of each $T$ does not exceed $w_f - j + r + h$. Using $w_f \leq (r-1)d$ and $h \leq j$, it follows that the weight of each $T$ does not exceed $(r-1)d + r$. The forms $T$ are thus linear combinations of power products $\gamma_iP_f$ where again the $P_f$ are power products in the $y_i$ of degree $d$ and weight not exceeding $(r-1)d$. We need merely choose $c \geq r$ if such a $\gamma_i$ appears in a term of the $T$, while if no such $\gamma_i$ is present effectively, then for any choice of $\gamma_i$ the statement is true.

36. What we have shown is that

$$\lambda^{t+1}P_f = \sum_{j=1}^{m} E_{fj}P_j \ [F], \quad f = 1, \ldots, m,$$

where the $E_{fj}$ are forms in $y$, and the $u_i$. $E_{fj}$ is homogeneous of degree $t$ in $\lambda$, homogeneous of degree unity in the $u_i$ and homogeneous of degree unity in the $y_i$. Transposing, we have a system of $m$ linear congruences for the $P_f$. In the $i$th congruence the coefficients of $P_j$ with $j \neq i$ is $-E_{ij}$ while the coefficient of $P_i$ is $\lambda^t - E_{ii}$. It follows that

$$P_fD = 0 \ [F], \quad f = 1, 2, \ldots, m,$$

where $D$ is the determinant of the system of congruences. Clearly $D$ is of the form $\lambda^{m(t+1)} + H$ where $H$ vanishes for $y = 0$. Since $y^d$ is one of the $P_f$ we have our result. Observe that $H$ is homogeneous of degree $m(t+1)$ in $\lambda$ and the $u_i$.

37. As a supplement to the proof the size of $t$ and $d$ will be examined. Reference to the work of Part I reveals that if $d = rp - 1$ then $t$ may be taken as $(r-1)p$. If, on the other hand, $d = r(p-1) + 1$, the proof may also be car-

\( ^{(8)} \) This is obvious for $q = 0$. An easy induction establishes it for all positive integral values of $q$, for the product of $\lambda^s$ with the $g$th derivative of $A$ is a form $\lambda^{s+1}A_4$ plus terms all of which contain a factor $\lambda^sA_i$ with $i < q$. 

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
ried out with possibly a larger value of \( t \). Note that this value of \( d \) is in general smaller than the previous one.

**Generalization for one unknown**

38. This theorem may be extended in the following way. Let

\[
F = \lambda y^{p_0} y_1^{p_1} \cdots y_r^{p_r} + \sum u_i B_i
\]

be a form in \( y, \lambda \) and a finite number of unknowns \( u_i \), where \( p_0, p_1, \ldots, p_r \) are non-negative integers with \( p_r \neq 0 \), and where the \( B_i \) are power products in the \( y_i \). If the degree of each \( B_i \) in the letters \( y_r, y_{r+1}, \ldots \), exceeds \( p_r \), then Theorem 2.1 may be applied to \( F \). It yields a relation

\[
y_r^d ((\lambda y^{p_0} y_1^{p_1} \cdots y_{r-1}^{p_{r-1}}) + H) = 0 \ [F],
\]

where \( H \) vanishes for \( y_{r+i} = 0, i = 0, 1, 2, \ldots \). It is obvious that \( F \) admits the solution \( y = 0 \). This relation shows that every irreducible manifold held by \( F \) which contains \( y = 0 \) and is not held by \( y_r \) must be held by the form in the outer parentheses. We are going to show how additional hypotheses on the \( B_i \) make it possible to draw a stronger conclusion.

**Theorem 2.2.** Let the \( F \) be given by (2.2) and let the \( B_i \) be such that for each integer \( f, f = 0, 1, \ldots, r \), the degree of each \( B_i \) in the letters \( y_f, y_{f+1}, y_{f+2}, \ldots \) exceeds \( p_f + p_{f+1} + \cdots + p_r \). Then there is a positive integer \( s \) and a form

\[
D = \lambda s + H
\]

where every term of \( H \) contains \( y \) or one of its derivatives effectively and where \( D \) is homogeneous of degree \( s \) in \( \lambda \) and the \( u_i \) and their derivatives, such that for some positive integer \( d \)

\[
y^d D = 0 \ [F].
\]

The stronger conclusion drawn here is that if \( \lambda \) is specialized as any non-zero element of \( \mathfrak{A} \), then the solution \( y = 0 \) of \( F \) is contained in no irreducible manifold held by \( F \) but not by \( y_r \).

39. The proof is by induction on \( r \). For \( r = 0 \) this result is practically identical with Theorem 2.1. The only difference is that the \( B_i \) may be of degree greater than \( p_0 + 1 \) in the \( y_i \). However, it is easy to see that by incorporating superfluous factors of the \( B_i \) into their coefficients we obtain a form for which Theorem 2.1 may be invoked. We assume the theorem established for integers from zero to \( r - 1 \) inclusive and prove it for \( r \).

40. We introduce new letters \( \lambda', u_i' \) and a new form \( F' \) in \( y_1, \lambda', u_i' \) in the following way. Let \( Q_i \) be that factor of \( B_i \) of degree \( p_0 \) which is higher than any other such factor (if \( p_0 = 0 \) then \( Q_i \) is unity). Let \( B_i' \) be defined by \( B_i = B_i' Q_i \). Let
\[ F' = \lambda' y_1^{p_1} y_2^{p_2} \cdots y_r^{p_r} + \sum u'_i B'_i. \]

\( F' \) goes over into \( F \) when \( \lambda' \) is replaced in \( F' \) by \( \lambda y^{p_0} \) and \( u'_i \) by \( u_i Q_i \). It will be shown that \( F' \) satisfies the induction hypothesis for \( r - 1 \). To do this it suffices to show that for all integers \( f, f = 1, 2, \ldots, r \), the degree of each \( B'_i \) in \( y_f, y_{f+1}, \ldots \) exceeds \( p_f + p_{f+1} + \cdots + p_r \). Suppose there were a \( B'_i \) and an \( f \) for which this were not so. Because of the way \( B'_i \) was defined, our hypothesis would then require that \( Q_i \) be of positive degree in some \( y_t \) with \( t \geq f \). Since \( Q_i \) is the highest factor of \( B_i \) of degree \( p_0 \), it follows that \( B'_i \) is free of \( y, y_1, \ldots, y_{f-1} \). Thus the degree of \( B_i \) is simply \( p_0 \) plus the degree of \( B'_i \) in \( y_f, y_{f+1}, \ldots \). Clearly the statement that the degree of \( B'_i \) in \( y_f, y_{f+1}, \ldots \) does not exceed \( p_f + p_{f+1} + \cdots + p_r \) conflicts with our assumption that the degree of \( B_i \) exceeds \( p_0 + p_f + \cdots + p_r \).

41. Under our induction hypothesis there is a form \( D' = \lambda' + H' \) and an integer \( d_1 \) such that

\[ y_1^{d_1} D' = 0 \quad [F'] \]

where \( H' \) is homogeneous in \( \lambda' \) and the \( u'_i \) of degree \( t \) and vanishes for \( y_1 = 0 \). Let \( D' \) become \( D_1 \) when \( \lambda' \) and the \( u'_i \) are replaced as above. Then \( y_1^{d_1} D_1 \) is in \([F]\). \( D_1 \) is of the form \( \lambda y^{p_0} + H_1 \) where \( H_1 \) is homogeneous in \( \lambda \) and the \( u_i \) of degree \( t \). Every term of \( H \) is of degree at least \( tp_0 + 1 \) in the \( y_i \).

It follows from Theorem 2.1 that there is a form \( D_2 = \lambda^w + H_2 \), \( H_2 \) homogeneous in \( \lambda \) and the \( u_i \) of degree \( w \) and vanishing for \( y = 0 \), and an integer \( k \) such that \( y^k D_2 = 0 \quad [D_1] \). By the result stated in the footnote to Theorem 1.1 we find that there are integers \( h \) and \( g \) such that \( y^g D_2^g = 0 \quad [D_1] \). This same result used again shows that for sufficiently large \( d \), \( y^d D_2^d \) is in \([y_1^{d_1} D_1]\). This completes the proof, since \( y_1^{d_1} D_1 \) is in \([F]\), and the form \( D = D_2^d \) and the integer \( d \) have the required properties.

A special case

**Theorem 2.3.** Let \( p \) be a positive integer, let \( A = y^p \), and let \( A_i \) be the \( i \)th derivative of \( A \). Let

\[ F = A - \sum C_{ij} y_i A_j \]

where \( i \) and \( j \) have some definite range and where the \( C_{ij} \) are forms in \( y \). Then there is a form \( D = 1 + H \), where \( H \) is a form in \( y \) which vanishes for \( y = 0 \), such that

\[ y^p D = 0 \quad [F]. \]

The distinctive feature of this result is that the integer \( p \) is available as the integer \( d \) of our previous work. For \( p = 1 \) this theorem is identical with a result obtained in a paper by Ritt and Kolchin(9).


License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
The device of continued substitution used in the first part of our proof is borrowed from this paper.

42. By hypothesis

\[ A = \sum_{i,j} C_{ij} y_i A_j [F]. \]  

If in the right member of (2.3) \( A_j \) is replaced by the \( j \)th derivative of the whole right member, the result is a congruence

\[ A = \sum_{i,j,k} C_{ijk} y_i y_j A_k [F]. \]

For each \( C_{ij} \) of (2.3) which is different from zero, let the sum \( i+j \) be computed. Suppose \( r \) is the largest of these sums. Then no sum \( i+j+k \) for which there is a nonzero \( C_{ijk} \) exceeds \( 2r \). If in the right member of (2.4) the \( k \)th derivative of the whole right member of (2.3) is substituted for \( A_k \), a new congruence for \( A \) is obtained. After \( s-1 \) iterations of this substitution, a congruence

\[ A = \sum C_{i_1 i_2 \cdots i_{s+1}} y_{i_1} y_{i_2} \cdots y_{i_s} A_{i_{s+1}} [F] \]

is obtained, where no sum \( i_1+i_2+\cdots+i_s \) with nonzero \( C_{i_1 \cdots i_s} \) exceeds \( sr \). If \( s = (r+1)(p-1)+1 \) then by our frequently used criterion, \( y_{i_1} y_{i_2} \cdots y_{i_s} \equiv 0 [A] \). Thus

\[ A = \sum D_{ij} A_i A_j [F] \]

where \( D_{ij} \) is a form in \( y \).

43. We now compute the sum \( i+j \) for each \( D_{ij} \) effectively present in (2.5). Let \( s \) be the largest of these sums. If (2.5) is differentiated \( s \) times, we obtain \( s+1 \) congruences, expressing \( A_k, k=0, 1, \ldots, s \), as linear combinations of products \( A_i A_j \) whose coefficients are forms in \( y \). Each product so obtained must contain an \( A_i \) with \( i \leq s \), for the differentiation introduced an increment of at most \( s \) to the sums \( i+j \) and they did not exceed \( s \) at the outset. We have shown, then, that

\[ A_i = \sum_{j=0}^s K_{ij} A_j [F], \quad i = 0, 1, \ldots, s, \]

where \( K_{ij} \) vanishes for \( y=0 \). Transposing, we have a system of \( s+1 \) linear homogeneous congruences for the \( A_i \). In the \( i \)th congruence the coefficient of \( A_j \) with \( j \neq i \) is \(-K_{ij}\) while the coefficient of \( A_i \) is \((1-K_{ii})\). \( D \), the determinant of the system, is of the form \( 1+H \) where \( H \) vanishes for \( y=0 \). Since \( A_i D \equiv 0 [F]; i=0, 1, \ldots, s \) we have our result.

**Generalizations for several unknowns**

44. Our next result concerns systems of forms in the unknowns \( y_1, \ldots, y_n, \lambda \) and a finite number of unknowns \( u_{ij} \). As is customary, we shall denote the \( j \)th
derivative of $y_i$ with the symbol $y_{ij}$ and the $i$th derivative of $y_i$ by $\lambda_i$. The second subscript of the $u_{ij}$ will not mean differentiation, but will simply indicate how these unknowns are displayed in rows and columns.

45. Let $s$ be any positive integer not greater than $n$. We consider subsets $(i_1, i_2, \ldots, i_s)$ of $(1, 2, \ldots, n)$ where in each subset the numbers $i_k$ are all different. If the binomial coefficient $C_{n,s}$ is denoted by $q$, there are exactly $q$ such subsets. We suppose a number $j$, $j = 1, \ldots, q$, assigned in any univocal manner to each such subset. We shall consider the system

\begin{equation}
F_j = \lambda y_{i_1}^{p_{j_1}} y_{i_2}^{p_{j_2}} \cdots y_{i_s}^{p_{j_s}} + \sum u_{ij} B_{ij},
\end{equation}

where the $p_{j_i}$ are positive integers, and the $B_{ij}$ power products in $y_1, \cdots, y_n$ and their derivatives. We call $\lambda y_{i_1}^{p_{j_1}} y_{i_2}^{p_{j_2}} \cdots y_{i_s}^{p_{j_s}}$ the first term of $F_j$. For each form $F_j$ we make the following assumptions concerning the degree of the $B_{ij}$ in the unknowns $y_1, \cdots, y_n$ and their derivatives. These assumptions describe a relation whereby the $B_{ij}$ dominate the first terms of the $F_j$ and it should be understood that each $B_{ij}$ is qualified in this way only by the first term of that form $F_j$ which contains it. Let $(i_a, i_b, \cdots, i_j)$ be any (proper or improper) subset of $(i_1, i_2, \cdots, i_s)$. It is required of $B_{ij}$ that

1. either

   (1a) $B_{ij}$ be divisible by $y_{i_a}^{p_{j_a}} y_{i_b}^{p_{j_b}} \cdots y_{i_j}^{p_{j_j}}$

   or

   (1b) the degree of $B_{ij}$ in the unknowns $y_{i_a}, y_{i_b}, \cdots, y_{i_j}$, in the unknowns $y_k$ not in the first term of $F_j$, and in the derivatives of all these unknowns exceeds $p_{j_a} + p_{j_b} + \cdots + p_{j_j}$;

2. the total degree of each $B_{ij}$ must exceed

   $p_{j_1} + p_{j_2} + \cdots + p_{j_s}$.

Theorem 2.4. Let $\Sigma$ be the differential ideal generated by the $F_j$. Then there exists a form $D = \lambda^c + H$, where $H$ is a form in the $y_i$, $\lambda$ and the $u_{ij}$, which vanishes for $y_i = 0$, $i = 1, \ldots, n$, and an integer $t$ such that for every form $V_j = y_{i_1} y_{i_2} \cdots y_{i_s}$

\begin{equation}
V_j D = 0 \quad [\Sigma].
\end{equation}

$H$ is homogeneous of degree $c$ in $\lambda$ and the $u_{ij}$.

46. Here, too, our result has considerable contact with Ritt's work in differential equations. Before taking up the proof a few remarks might be made concerning the content of this theorem from the standpoint of differential equations. The unknowns $\lambda$ and the $u_{ij}$ have been introduced as auxiliaries to facilitate the proof. For purposes of illustration we may suppose $\lambda$ replaced by unity and the $u_{ij}$ by any forms in the $y_i$. Equations (2.7) then have the appearance
Obviously $\Sigma$ admits the solution $y_i = 0$, $i = 1, \cdots, n$. These equations show that any irreducible manifold held by $\Sigma$ which contains $y_i = 0$, $i = 1, \cdots, n$, must be held by the system $V_j$, $j = 1, \cdots, q$. Each irreducible manifold in this latter system is found by letting some $n-s+1$ of the unknowns $y_i$ be zero, the remaining $s-1$ unknowns being arbitrary. In one extreme case, with $s = 1$, the manifold of the system $V_j$ is precisely $y_i = 0$, $i = 1, \cdots, n$. The essentiality of this solution in $\Sigma$ was shown by Ritt(10). He also treated the case $s = n$, obtaining(11) the above conclusion as a consequence of an approximation theorem. The intermediate cases, that is, those in which $1 < s < n$, appear here as new results, both from the abstract viewpoint and that of differential equations. The extreme cases owe their novelty to the fact that their proof is abstract. For the extreme case $s = n$, we reverse the procedure followed by Ritt. This result will be established first and then the analogue of Ritt's approximation theorem will be shown to follow from it. For $s = n$ and $\lambda = 1$ we have a single form

$$F = y_1^{p_1} y_2^{p_2} \cdots y_n^{p_n} + \sum u_i B_i.$$  

Our hypothesis now reduces to the statements that for each $k$ each $B_i$ is either divisible by $y_k^{p_k}$ or its degree in the $y_{kj}$ exceeds $p_k$. The total degree of each $B_i$ exceeds $p_1 + \cdots + p_n$. The conclusion is that there is a congruence

$$(y_1 \cdots y_n)'(1 + H) \equiv 0 \ [F]$$

where $H$ vanishes for $y_i = 0$, $i = 1, \cdots, n$. This special case of Theorem 2.4 will be used in our later work.

47. We prove Theorem 2.4 by induction on $s$, beginning with $s = 1$ and $n$ arbitrary. For this case our hypothesis states that $\Sigma$ contains the forms

$$F_j = \lambda y_j^{p_j} + \sum u_{ji} B_{ji}, \quad = 1, \cdots, n,$$

the $B_{ji}$ being power products in the $y_i$ and their derivatives whose total degree exceeds $p_j$. Let $r$ be the maximum of the weights of the $B_{ji}$. We understand that the weight of $y_{jk}$ is $k$. Let $p$ be the maximum of the $p_j$. Let $d = n(r+1)(p-1)+1$. Then every power product $P_h$ in the $y_{ij}$ of degree $d$ whose weight $w_h$ does not exceed $r((r+1)(p-1)+1)$ is in some $[y_i^{p_i}]$, for $P_h$ is of degree at least $(r+1)(p-1)+1$ in at least one $y_j$, and then our earlier result applies. We now follow the procedure used in the proof of Theorem 2.1. We first multiply each $P_h$ by $\lambda^p$ so that the product is in some $[\lambda y_i^{p_i}]$ and after substitutions and re-
arrangements similar to those used in the proof of that theorem we obtain the congruences
\[ \lambda^e P_h \equiv \sum_{i=1}^{m} E_{hi} P_i [\Sigma], \quad h = 1, 2, \ldots, m, \]
m being the number of \( P_h \). The only difference is that now the \( E_{hi} \) are forms in \( y_1, \ldots, y_n, \lambda \) and the \( u_{ji} \) which need not be homogeneous in the \( y_i \). It is still true that every term of each \( E_{hi} \) involves some \( y_k \) effectively. \( E_{hi} \) is homogeneous of degree \( g-1 \) in \( \lambda \) and is homogeneous of degree unity in the \( u_{ji} \). After transposing, we see that
\[ P_h D = 0 [\Sigma], \quad h = 1, \ldots, m, \]
where \( D \), the determinant of the transposed system, is of the type described in the statement of this theorem. Since \( y_1^\alpha, y_2^\alpha, \ldots, y_n^\alpha \) are all to be found among the \( P_h \) we have our result.

48. Continuing with the proof we suppose that the theorem holds for all values of \( s \) from unity to some fixed integer, and proceed to show that it holds for the next integer. We shall denote this latter integer by \( s \). In our induction assumptions the only restriction on \( n \) is that it be sufficiently large to insure that the statement of the theorem makes sense, that is, \( n \) is never less than \( s \). It is otherwise arbitrary.

Referring to (2.6) select all those forms \( F_j \) whose first terms contain the letter \( y_1 \) effectively. Let \( p \) be the maximum of the exponents of \( y_1 \) in the first terms of these \( F_j \). By multiplying certain of these \( F_j \) by a suitable power of \( y_1 \) we obtain a system of forms \( G_j \) each of whose first terms contains the letter \( y_1 \) exactly to the \( p \)th power. The set of forms \( G_j \), being composed of some of these \( F_j \) and multiples by a power of \( y_1 \) of the others, is certainly in \( \Sigma \). Furthermore the terms of the \( G_j \) which are not first terms satisfy the conditions of our hypothesis relative to the first term of the form which contains them. We now introduce new letters \( \lambda' \), \( u_{ji}' \), and new forms \( G_j' \) in these letters and \( y_2, y_3, \ldots, y_n \). The \( G_j' \) will be defined so that their first terms will contain only \( s-1 \) letters. They will fulfill the conditions of our hypothesis, and will go over into the \( G_j \) when appropriate replacements are made for \( \lambda' \) and the \( u_{ji}' \). We begin with
\[ G_1 = \lambda y_1^{p_1} y_2^{p_{12}} \cdots y_s^{p_{1s}} + \sum u_{1i} B_{1i}. \]
The \( B_{1i} \) are either the original \( B_{1i} \) or multiples of them by a power of \( y_1 \). As a first step we divide the \( B_{1i} \) into two classes, those which are divisible by \( y_2^{p_{12}} \cdots y_s^{p_{1s}} \) and those which are not. A \( B_{1i} \) which is divisible by this term may be written as \( L_{1i} y_2^{p_{12}} \cdots y_s^{p_{1s}} \). Our hypothesis reveals that the degree of \( L_{1i} \) in \( y_1, \ldots, y_n \) and their derivatives exceeds \( p \). We now have
\[ G_1 = \left( \lambda y_1^{p} + \sum_k u_{1k} L_{1k} \right) y_2^{p_{12}} \cdots y_s^{p_{1s}} + \sum u_{1i} B_{1i}. \]
where the second summation is performed over all those \( B_{i'i} \) which are in the second class. We further subdivide the \( B_{i'i} \) of the second class into those which contain a factor in \( y_1 \) and its derivatives of degree \( p \) and those which do not. For a \( B_{i'i} \) of the first kind let \( Q_{i'i} \) be any such factor and let \( B_{i'i} = Q_{i'i}B_{i'i}' \). For \( B_{i'i} \) of the second kind the total multiplicity of \( y_1 \) and its derivatives in \( B_{i'i} \) is some number \( q_{i'i} < p \). Let all these letters be split off from \( B_{i'i} \) and multiplied by any factor of \( B_{i'i} \) which contains only the unknowns \( y_{i+1}, \ldots, y_n \) and whose degree in these unknowns and their derivatives is \( p - q_{i'i} \). Our hypothesis permits us to construct such a product for each \( B_{i'i} \) of the second kind. Denoting this product by \( Q_{i'i} \) we have in \( Q_{i'i} \) a power product in \( y_1, y_{i+1}, \ldots, y_n \) and their derivatives of degree \( p \). Here too we let \( B_{i'i} = Q_{i'i}B_{i'i}' \). Let

\[
G'_i = \lambda_1 y_{i+1}^{p_{i+1}} \cdots y_n^{p_n} + \sum u_{i'i} B_{i'i}'.
\]

When \( \lambda_1 \) is replaced by \( \lambda_1 + \sum u_{i'i} L_{i'i} \) and \( u_{i'i} \) by \( u_{i'i} Q_{i'i} \) then \( G'_i \) goes over into \( G_i \).

49. It is readily seen that if \( (a, b, \ldots, f) \) is any subset of \( (2, 3, \ldots, s) \) then for each \( B_{i'i}' \) either

(a) \( B_{i'i}' \) is divisible by \( y_2 y_3 \cdots y_f \) or

(b) the degree of \( B_{i'i}' \) in \( y_2, y_3, \ldots, y_f, y_{i+1}, \ldots, y_n \) and their derivatives exceeds \( p_{i+1} + p_{i+2} + \cdots + p_f \).

For a \( B_{i'i}' \) obtained from a \( B_{i'i} \) of the first kind this is part of our initial hypothesis, since in this case \( B_{i'i}' \) is the same as \( B_{i'i} \) as far as the unknowns \( y_2, y_3, \ldots, y_n \) are concerned. As for a \( B_{i'i}' \) obtained from a \( B_{i'i} \) of the second kind, recall that \( B_{i'i} = Q_{i'i}B_{i'i}' \) where \( Q_{i'i} \) is of degree \( p \) in \( y_1, y_{i+1}, \ldots, y_n \) and their derivatives and where \( B_{i'i}' \) is free of \( y_1 \) and its derivatives. Since \( B_{i'i} \) is not divisible by \( y_2 y_3 \cdots y_f \) (because its degree in \( y_1 \) is less than \( p \)) our original hypothesis requires that the degree of \( B_{i'i} \) in \( y_1, y_2, y_3, \ldots, y_i, y_{i+1}, \ldots, y_n \) and their derivatives exceed \( p + p_{i+1} + p_{i+2} + \cdots + p_f \). Since \( Q_{i'i} \) is of degree \( p \) it follows that the degree of \( B_{i'i}' \) exceeds \( p_{i+1} + p_{i+2} + \cdots + p_f \). Since in addition \( B_{i'i}' \) is free of \( y_1 \) and its derivatives we have our result. To complete the task of showing that \( G'_i \) satisfies our hypothesis for the unknowns \( y_2, \ldots, y_n \) we must also dispose of the requirement that the degree of each \( B_{i'i}' \) in \( y_2, \ldots, y_n \) exceed \( p_{i+1} + p_{i+2} + \cdots + p_f \). Of the original \( B_{i'i} \) the only ones which need not meet this requirement are those divisible by \( y_2 y_3 \cdots y_f \). These have been removed from consideration by our choice of \( \lambda_1 \). Thus, referring to the argument just given, if \( (a, b, \ldots, f) \) coincides with \( (2, 3, \ldots, s) \) the alternative (a) is excluded; alternative (b) is what was to have been established.

50. Proceeding in this way we obtain a form \( G'_j \) for each \( F_j \) whose first term contains \( y_1 \) effectively. We have

\[
G'_j = \lambda_j y_{i+1}^{p_{i+1}} y_i^{p_i} \cdots y_n^{p_n} + \sum u_{ji} B_{ji}'.
\]
where \((i_2, i_3, \ldots, i_s)\) is a subset of \(s - 1\) distinct numbers of the set \((2, 3, \ldots, n)\) and where each \(B_n^i\) is a power product in \(y_2, y_3, \ldots, y_n\) and their derivatives which dominates the first term of \(G'_j\) in the required way. Each \(G'_j\) goes over into \(G_j\) when \(\lambda_j\) is replaced by \(\lambda y_j^i + \sum u_{jk} L_{jk}\) and the \(u^i_{ji}\) are replaced by \(u_{ji} Q_{ji}\). Let the number of these forms be \(q_1\).

51. We now introduce new unknowns \(\lambda''\) and \(u''_{ji}\) and new forms

\[ G'_j = \lambda'' y_{i_2}^{p_{i_2}} \cdots y_{i_s}^{p_{i_s}} + \sum u''_{ji} B_{ji}, \quad j = 1, \ldots, q_1. \]

When in \(G'_j\) the unknown \(\lambda''\) is replaced by \(\lambda_1 \lambda_2 \cdots \lambda_{q_1}\) and each \(u''_{ji}\) by \(\lambda_1 \cdots \lambda_{j-i} \lambda_{j+1} \cdots \lambda_{q_1} u'_{ji}\) then \(G'_j\) goes over into \(\lambda_1 \cdots \lambda_{j-i} \lambda_{j+1} \cdots \lambda_{q_1} G_j^i\).

Let \(\Sigma''\) be the differential ideal generated by the \(G'_j\). Under the terms of the induction hypothesis we conclude that \(G'' = \lambda''^{w_1} + H''\) and an integer \(t_1\), such that for every form \(y_{i_2} y_{i_3} \cdots y_{i_s}\) we have

\[ (y_{i_2} y_{i_3} \cdots y_{i_s})^t D'' = 0 [\Sigma'']. \]

\(H''\) vanishes for \(y_i = 0, i = 1, \ldots, n\), which is to say that every term of \(H''\) contains some \(y_{ij}\) effectively. \(H''\) is homogeneous of degree \(w_1\) in \(\lambda''\) and the \(u''_{ji}\). When \(\lambda''\) and the \(u''_{ji}\) are replaced in \(D''\) by the forms indicated above, \(D''\) goes over into a form

\[ D'' = (\lambda_1 \lambda_2 \cdots \lambda_{q_1})^{w_1} + H'. \]

\(D''\) is a form in \(y_1, \ldots, y_n, \lambda_1, \ldots, \lambda_{q_1}\) and the \(u'_{ji}\). It is homogeneous of degree \(q_1 w_1\) in the \(\lambda'\) and the \(u'_{ji}\). Now in \(D''\) let \(\lambda_{ij}\) be replaced by \(\lambda y_j^i + \sum u_{jk} L_{jk}\) and \(u'_{ji}\) by \(u_{ji} Q_{ji}\). Let \(p_1 = p w_1 q_1\). Then \(D''\) goes over into a form

\[ D'' = \lambda^{q_1 w_1} y_j^i + H', \]

\(D_1\) is a form in \(y_1, \ldots, y_n, \lambda\) and the \(u_{ji}\). It is homogeneous of degree \(q_1 w_1\) in \(\lambda\) and the \(u_{ji}\). We are going to show that the degree in the \(y_{ij}\) of each term of \(H_1\) exceeds \(p_1\). To this end we consider two types of terms of \(H_1\), those arising from \(H'_1\) and those from \(L = (\lambda_1 \cdots \lambda_{q_1})^{w_1}\). Each \(\lambda_{ij}\) was replaced by \(\lambda y_j^i + \sum u_{jk} L_{jk}\) where the degree of each \(L_{jk}\) in \(y_1, \ldots, y_n\) and their derivatives exceeded \(p\). Therefore the degree in the \(y_{ij}\) of every term of \(L\) except \(\lambda^{q_1 w_1} y_j^i\) exceeds \(p_1\). This accounts for terms of \(H_1\) arising from \(L\). As for those arising from \(H'_1\) recall that \(H'_1\) is homogeneous of degree \(q_1 w_1\) in the \(\lambda_i\) and the \(u_{ji}\). Since each \(\lambda_i\) contributes at least \(p\) to the degree of \(H_1\) in the \(y_{ij}\) and each \(u_{ji}\) contributes exactly \(p\) to this degree, the \(\lambda_i\) and \(u_{ji}\) contribute at least \(p_1 = p w_1 q_1\). Because each term of \(H'_1\) was of positive degree in the \(y_{ij}\) it follows that the terms of \(H_1\) arising from \(H'_1\) also have a degree in the \(y_{ij}\) which exceeds \(p_1\). This verifies our assertion about \(D_1\).

After these replacements are made, equation (2.8) becomes

\[ (y_{i_2} y_{i_3} \cdots y_{i_s})^t D_1 = 0 [\Sigma]. \]
Consequently, for every $V_j = y_{i_1} y_{i_2} \cdots y_{i_r}$,

$$V_j^i D_1 \equiv 0 \ [\Sigma].$$

By singling out all of the original $F_j$ whose first terms contain effectively $y_k$, $k = 2, 3, \cdots, n$, and repeating the above procedure for each $k$, we obtain finally $n$ forms

$$D_k = \lambda^{q_k w_k} y_k^{p_k} + H_k, \quad k = 1, \cdots, n,$$

where $H_k$ is homogeneous of degree $q_k w_k$ in $\lambda$ and the $u_{ji}$ and every term of $H_k$ contains a power product in the $y_{ij}$ whose degree exceeds $p_k$. There is an integer $t$ such that

$$(2.9) \quad V_j^t D_k \equiv 0 \ [\Sigma], \quad k = 1, 2, \cdots, n; j = 1, 2, \cdots, q.$$ 

Let $w$ be the maximum of the numbers $w_k q_k$ and let each $D_k$ be multiplied by $\lambda^{w-w_k}$. Using the same symbols to denote the modified forms, we see that equations (2.9) still hold, and that now the $D_k$ are homogeneous of degree $w$ in $\lambda$ and the $u_{ji}$. What we have accomplished by this alteration is to obtain a set of forms $D_k = \lambda^w y_k^{p_k} + H_k, k = 1, \cdots, n$, to which we may apply our result for the case $s = 1$. It follows that there is a form $D = \lambda^e + H$ and an integer $a$ such that

$$(2.10) \quad y_i \lambda^a D \equiv 0 \ [D_1, D_2, \cdots, D_n], \quad i = 1, 2, \cdots, n.$$ 

Every term of $H$ contains some $y_{ij}$ effectively, and $H$ is homogeneous of degree $e$ in $\lambda$ and the $u_{ji}$. (Actually the conclusion that the case $s = 1$ entitles us to draw is that $D = (\lambda^w)^b + H$ where $H$ is homogeneous of degree $b$ in $\lambda^w$ and the coefficients of the $H_k$. Since these coefficients are themselves homogeneous of degree $w$ in $\lambda$ and the $u_{ji}$, the above conclusion is justified.)

52. We now show that there is an integer $t$ such that

$$V_j^t D \equiv 0 \ [\Sigma], \quad j = 1, 2, \cdots, q.$$ 

This will complete the proof, since $D$ meets all our other requirements.

We know that for each $D_k$ there is a power of $V_j$ such that its product with $D_k$ is in $\Sigma$. This is likewise true of any derivative of the $D_k$. We chose the integer $h$ sufficiently large so that the product of $V_j$ with any $D_k$ or with any derivative of a $D_k$ which appears effectively in the right member of some congruence (2.10) shall be in $\Sigma$. A single $h$ serves for all $j$. Clearly

$$V_j^h y_{i}^a D \equiv 0 \ [\Sigma], \quad j = 1, \cdots, q; i = 1, \cdots, n.$$ 

The integer $h+a$ thus has the required property; for every $j$, $V_j^{h+a} D$ is in $\Sigma$.  

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
An approximation theorem

53. Our next theorem is an application of Theorem 2.4. Its statement and proof presuppose Raudenbush's theory of perfect differential ideals. The coefficient domain is an arbitrary differential field of characteristic zero.

Theorem 2.5. Let $\Sigma$ be a prime ideal of forms in $y_1, \ldots, y_n$ which does not contain $V = y_1 y_2 \cdots y_n$ and which is such that every form of $\Sigma$ vanishes for $y_i = 0$, $i = 1, \ldots, n$. Let $r$ be a positive integer and let $u_1, \ldots, u_n$ be unknowns. If in each form of $\Sigma$, $y_{ij}$ is replaced by $(u_i^j)^{1/r}$, $i = 1, \ldots, n$, $\Sigma$ goes over into a system $\sigma$ of forms in $u_1, \ldots, u_n$. Let $\{\sigma\}$ be the perfect ideal generated by $\sigma$ and let it be the intersection of prime ideals $\Omega_1, \ldots, \Omega_r$. Then there is an $\Omega_i$ which does not contain $W = u_1 u_2 \cdots u_n$ and every form of $\Omega_i$ vanishes for $u_i = 0$, $i = 1, \ldots, n$.

What this amounts to in the theory of differential equations is that the approximation theorem which holds for $r = 1$ holds for any positive integral value of $r$. In this form our theorem has been established by Ritt(12). What we shall prove is thus the abstract counterpart of Ritt's approximation theorem relative to the $r$th roots of the functions constituting a solution of an irreducible system of differential equations. Our proof is indirect. We assume the theorem false and force a contradiction.

54. If the theorem were false, every $\Omega_i$ which did not contain $W$ would contain a form $1 + B_i$ where $B_i$ vanishes for $u_i = 0, i = 1, \ldots, n$. Clearly $W$ is not in each $\Omega_i$ for then $V$ would be in $\Sigma$. Our assumption that the theorem is false implies that there actually are such forms $1 + B_i$. Let their product be $1 + B$. $B$ vanishes for $u_i = 0, i = 1, \ldots, n$, and $W/(1 + B)$ is in $\{\sigma\}$. Let $F = W/(1 + B)$. There is an integer $s$ such that $F^s = 0$. We work back from $F^s$ to a form of $\Sigma$.

The forms of $\sigma$ were obtained from those of $\Sigma$ by the transformation $y_{ij} = (u_i^j)^{1/r}$. Thus while $\sigma$ is not a different ideal it is closed with respect to differentiation. The inverse of the above transformation may be obtained from the formulas $u_i^j = ((u_i y_i)/(ry_i))_{j-1}$ where the subscript outside the parentheses denotes differentiation. These formulas show that $u_i^j, j = 1, 2, 3, \ldots$, may be expressed as the product of $u_i$ with a polynomial (rational coefficients) in $y_i/y_i$ and its first $j - 1$ derivatives. Each term of these expressions for the $u_i^j$ is the quotient of a polynomial in $u_i$ and the $y_i$ by a power of $y_i$. The total degree of the numerator in $u_i$ and $y_i$ exceeds the degree of the denominator.

55. We examine the effect of making the above replacements for the derivatives of the $u_i$ in a form in the $u_i$ no term of which is free of all the $u_i$. We obtain a rational function of the $u_i$ and the $y_{ij}$ whose least common denominator is a power product $y_1^{p_1} y_2^{p_2} \cdots y_n^{p_n}$. When the rational function is written in the form $P/(y_1^{p_1} y_2^{p_2} \cdots y_n^{p_n})$ with $P$ a polynomial in the $u_i$ and the

(12) OCP, pp. 7–14.
then for each \( i \) for which \( p_i \neq 0 \) each term of \( P \) not divisible by \( y_i^{p_i} \) is of degree greater than \( p_i \) in \( yi \) and its derivatives. In addition, the total degree of every term of \( P \) in the \( u_i, y_i \) exceeds \( p_1 + p_2 + \cdots + p_n \). If, for some \( k, p_k = 0 \), then \( P \) may be free of the letters \( u_k, y_k \).

Let us suppose these replacements made in \( F^t \). An expression \( W^t(1 + T) \) is obtained which involves the \( u_i \) and the \( y_i \). The expression \( T \) is a rational function whose numerator is a polynomial \( L \) in the \( u_i \) and the \( y_i \) and whose denominator is a power product \( y_1^{a_1}y_2^{a_2}\cdots y_n^{a_n} \). The remarks made above about the degree of \( P \) hold also for \( L \) relative to the exponents \( a_1, a_2, \cdots, a_n \). \( F^t \) belongs to \( [\sigma] \) and is a linear combination of forms of \( \sigma \) whose coefficients are forms in \( u_1, u_2, \cdots, u_n \). When the above replacements are made in a form of \( \sigma \), what is obtained is a form in \( u_i^t \) and the \( y_i \) which, when \( u_i^t \) is replaced by \( y_i \), becomes a form of \( \Sigma \). The fractions with which we deal are produced only by the coefficients which figure in the linear combination. Let \( W^t(1 + T) \) be multiplied by the least common denominator \( y_1^{b_1}y_2^{b_2}\cdots y_n^{b_n} \) of these coefficients. A form

\[
G = W^t(y_1^{b_1}y_2^{b_2}\cdots y_n^{b_n} + K)
\]

is obtained. \( G \) is a polynomial in the \( u_i \) and the \( y_i \). It is a linear combination, with polynomials in these letters for coefficients, of polynomials in \( u_i^t \) and the \( y_i \), these latter polynomials having the property that when \( u_i^t \) is replaced by \( y_i \) in them, forms of \( \Sigma \) are obtained. Each term of \( K \) dominates \( y_1^{b_1}y_2^{b_2}\cdots y_n^{b_n} \) in the same way that each term of \( L \) dominates \( y_1^{a_1}y_2^{a_2}\cdots y_n^{a_n} \).

56. Let \( \omega \) be a primitive \( r \)th root of unity. Let \( G', G'', \cdots, G^{(m)} \) denote the forms obtained by replacing \( u_i \) by \( \omega^k u_i \) in \( G \); \( k = 1, \cdots, r - 1; i = 1, \cdots, n \), the replacements being made independently on all the \( u_i \) in all possible ways. The product \( GG' \cdots G^{(m)} \) is a form in \( u_i^t \) and the \( y_i \). Our final substitution is to replace \( u_i^t \) by \( y_i \), \( i = 1, \cdots, n \), in this product. We denote this form in the \( y_i \) by \( P \). We have

\[
P = V^q(y_1^{q_1}y_2^{q_2}\cdots y_n^{q_n} + H)
\]

where \( H \) is a form in the \( y_i \) such that the total degree of every one of its terms exceeds \( q_1 + q_2 + \cdots + q_n \). In addition, every term of \( H \) not divisible by \( y_i^q \) has a total degree in \( y_i \) and its derivatives which exceeds \( q \), provided that \( q_i \neq 0 \). Clearly \( P \) is in \( \Sigma \). We are going to show that its presence in \( \Sigma \) contradicts our hypothesis.

Since \( \Sigma \) is prime and does not contain \( V \) we must have

\[
y_1^{q_1}y_2^{q_2}\cdots y_n^{q_n} + H \equiv 0 [\Sigma].
\]

Let \( N \) denote the form constituting the left member of this congruence. We recall that every form of \( \Sigma \) vanishes for \( y_i = 0, i = 1, \cdots, n \), and that \( H \) also
has this property. Consequently not all the exponents \( q_i \) of the first term of \( N \) are zero. Renumbering the letters if necessary, suppose that \( q_1, q_2, \ldots, q_t \) are not zero, while \( q_{t+1} = q_{t+2} = \cdots = q_n = 0 \). We have \( 1 \leq t \leq n \). We consider \( N \) as a form in \( y_1, \ldots, y_t \). It follows from Theorem 2.4 that there is a form \( 1 + R, R \) vanishing for \( y_i = 0, i = 1, \ldots, t \), and an integer \( a \) such that

\[
(y_1 y_2 \cdots y_t)^a (1 + R) \equiv 0 \pmod{n}.
\]

Since \( N \) and, consequently, \((y_1 y_2 \cdots y_t)^a (1 + R)\) are in \( \Sigma \) we have the following contradiction. \( \Sigma \) is prime, so it must contain either \( y_1 y_2 \cdots y_t \) or \( 1 + R \). By hypothesis \( \Sigma \) does not contain \( y_1 y_2 \cdots y_t \). By hypothesis every form of \( \Sigma \) vanishes for \( y_i = 0, i = 1, \ldots, n \), and since \( 1 + R \) does not have this property it cannot be in \( \Sigma \).

**Part III. The differential ideal generated by \( uv \)**

**The form \( uv \) and its derivatives**

57. Let \( u \) and \( v \) be unknowns. We investigate the differential ideal \( \Omega \) generated by the form \( X = uv \). For most of the discussion, the underlying ring will be that of forms in the unknowns \( u \) and \( v \) whose coefficients are rational numbers. Results obtained under these circumstances carry over readily to more general ones.

Our arguments follow the pattern of those used in the discussion of the ideal generated by \( y^p \). We begin with some conventions concerning power products in \( u \) and \( v \) and their derivatives. We retain for power products in the \( u_i \) alone the definitions concerning weight, degree and order already made for the unknown \( y \); and likewise for power products in the \( v_i \) alone. A power product \( P \) in both the \( u_i \) and the \( v_i \) may be written in the form \( UV \) where \( U \) involves only the \( u_i \) and \( V \) only the \( v_i \). The signature of \( P \) is defined as an ordered pair of numbers \((d_1, d_2)\) where \( d_1 \) and \( d_2 \) are the respective degrees of \( U \) and \( V \).

We can without fear of confusion describe as homogeneous a form all of whose terms have the same signature. The weight of \( P \) is defined as the sum of the weights of \( U \) and \( V \). A power product \( P \) is defined as higher than \( P' = U' V' \), and \( P' \) as lower than \( P \), if (a) \( U \) is higher than \( U' \) or (b) \( U = U' \) and \( V \) is higher than \( V' \). Evidently if \( P \) is different from \( P' \) it must be either higher than \( P \) or lower than \( P' \). It is evident that our ordering is transitive. Furthermore, if \( P \) is higher than \( P' \) and \( G \) is any power product, then \( GP \) is higher than \( GP' \). A \( \beta \) term is defined in the following way. Let \( P = UV \) be of signature \((d_1, d_2)\). \( P \) is a \( \beta \) term if \( V \) effectively contains some \( v_k \) with \( k < d_1 \). This implies of course that \( d_1 > 0 \). All other power products are called \( \alpha \) terms.

In particular unity, any power product in the \( u_i \) alone and any power product in the \( v_i \) alone are \( \alpha \) terms. It will be noted that these definitions do not respect the symmetry of \( \Omega \).
58. We prove the following lemma.

**Lemma 3.1.** Let $d_1, d_2, w$ be any non-negative integers. Every $\beta$ term $P = UV$ of signature $(d_1, d_2)$ and weight $w$ is congruent mod $\Omega$ to a linear combination with rational coefficients of a terms of this weight and signature.

We need only show that $P$ is either in $\Omega$ or is a congruent mod $\Omega$ to a linear combination with rational coefficients of power products of signature $(d_1, d_2)$ and weight $w$, all of which are higher than $P$. Arguments identical with those used in the proof of Lemma 1.1 show how the proof may then be completed.

59. The proof is by induction on $d_1$, with $d_2$ and $w$ arbitrary, starting with $d_1 = 1$. For this case $P = u_k V$ where $k \geq 0$ and $V$ involves only the $v_i$. The fact that $P$ is a $\beta$ term implies that $v$ is effectively present in $P$ and we have $P = u_0 w V'$. If $k = 0$ then $P$ is in $\Omega$ and requires no further discussion. Assume now that $k > 0$ and consider $X_k$, the $k$th derivative of $X = uv$. It is a form

$$u_k v + \sum_{i=1}^{k} c_i u_{k-i} v_i$$

where the $c_i$ are positive integers. Thus

$$P = (X_k - \sum_{i=1}^{k} c_i u_{k-i} v_i) V'$$

$$= - (\sum_{i=1}^{k} c_i u_{k-i} v_i) V' [\Omega].$$

Evidently each term $u_{k-i} v_i V'$ is higher than $P$ and has the same weight and signature as $P$, so that the statement is verified for $d_1 = 1$. Observe that in comparing the terms in the right member of this congruence with $P$, an examination of their factors in the $u_i$ alone reveals that they are higher than $P$. We attach this fact to the induction hypothesis and assume that the lemma, with this additional restriction, is valid for all power products of signature $(d_1, d_2)$ and weight $w$, and $d_2$ and $w$ arbitrary, provided that $d_1$ is less than some integer $d > 0$. It will be shown that it likewise holds for power products of signature $(d, d_2)$ and weight $w$.

60. Let $P$ be such a $\beta$ term. Then $P$ effectively contains some $v_k$ with $k < d$. Let $u_r$ be that derivative of $u$ of highest order which is effectively present in $P$. If $r = 0$, then $P$ is in $\Omega$ and needs no further discussion. Assuming $r > 0$, write

$$P = u_r v_k U' V'.$$

Now the $(r+k)$th derivative of $X$ is a form

$$X_{r+k} = c u_r v_k + \sum_{i=1}^{k} c_i u_{r+i} v_{k-i} + \sum_{i=1}^{r} c_i' u_{r-i} v_{k+i}$$

where $c$, the $c_i$ and the $c_i'$ are non-negative integers. Certainly $c$ and the $c_i'$ are different from zero and the $c_i$ are zero only if $k = 0$. Thus
(3.1) \[ P = \sum_{i=1}^{k} d_i u_{r+i} v_{k-i} U' V' + \sum_{i=1}^{r} d'_i u_{r-i} v_{k+i} U' V' \] 
where the \( d_i, d'_i \) are rational numbers. The terms \( u_{r+i} v_{k-i} U' V' \) are all of the same weight and signature as \( P \) and are all higher than \( P \) in the proper way. Consequently only the terms \( u_{r+i} v_{k+i} U' V' \) need be considered. Ignoring \( u_{r+i} \) for the moment, consider the terms \( v_{k+i} U' V' \). Since \( U' \) is of degree \( d-1 \) and \( d > k \) these are all \( \beta \) terms. The induction hypothesis applied to them shows that each is congruent to a linear combination of terms \( U'' V'' \) of the same weight and signature, and with \( U'' \) higher than \( U' \). Since \( U'' \) has the same degree as \( U' \), and \( U' \) involves only the letters \( u, \ u_1, \ldots, u_r \), it follows that for some \( t < r \) (\( t \) is a non-negative integer) the exponent of \( u_t \) in \( U'' \) exceeds that of \( u_t \) in \( U' \), while for all non-negative integers \( s < t \) the exponents of \( u_s \) in \( U'' \) and in \( U' \) are the same. We see then that each \( u_{r+i} U'' \) is higher than \( U \). Since in the congruence (3.1) each \( u_{r+i} v_{k+i} U' V' \) may be replaced by a linear combination of terms \( u_{r+i} U'' V'' \) of the proper weight and signature, the result follows.

**Canonical representations**

61. A \( \gamma \) term is defined to be a form
\[ G = EX_{i_1} X_{i_2} \ldots X_{i_s} \]
where \( E \) is an \( \alpha \) term and \( X_{i_1} X_{i_2} \ldots X_{i_s} \) is any power product in the \( X_i \) of positive degree. Let \( E \) be of signature \( (d_1, d_2) \) and weight \( w \). The signature of \( G \) is defined to be \( (d_1+s, d_2+s) \) and its weight to be \( w+i_1+\cdots+i_s \). \( G \) is actually a homogeneous isobaric form of this signature and weight. The following lemma shows that the forms defined here as \( \gamma \) terms are entirely analogous to those so defined relative to the ideal \([\gamma^p]\).

**Lemma 3.2.** Let \( H \) be any homogeneous isobaric element of \( \Omega \). Then \( H \) may be expressed as a linear combination with rational coefficients of \( \gamma \) terms all of which we have the same weight and signature as \( H \).

\( H \) is a linear combination with rational coefficients of terms \( K X_{i_1} \ldots X_{i_s}, \)
\( K \) being some power product in the \( u_i \) and \( v_i \). If \( K \) is not an \( \alpha \) term Lemma 3.1 asserts the existence of a congruence
\[ K = \sum c_i K_i \] 
where the \( K_i \) are \( \alpha \) terms and the \( c_i \) are constants. This congruence may be written as an equality
\[ K = \sum_i c_i K_i + \sum_{i,j} c_{ij} K_{ij} X_j \]
where the \( c_{ij} \) are constants and the \( K_{ij} \) power products. If the right member of this equation is substituted for \( K \) in the term \( K X_{i_1} \ldots X_{i_s} \) in ques-
tion, a linear combination of \( \gamma \) terms plus a linear combination of terms 
\( K_{ij}X_jX_{i_1} \cdots X_{i_r} \) is obtained. These terms all have a degree in the \( X_i \) which exceeds that of the original one. Repetition of this process a finite number of times in the finitely many terms of \( H \) must yield an expression for \( H \) of the required nature.

**The fundamental lemma**

62. The following lemma is the counterpart of Lemma 1.3.

**Lemma 3.3.** Let \( d_1, d_2, w \) be non-negative integers. The number of \( \gamma \) terms of signature \((d_1, d_2)\) and weight \( w \) does not exceed the number of \( \beta \) terms of this weight and signature.

The plan of the proof is the same as that of Lemma 1.3. A definite \( \beta \) term of this weight and signature will be assigned to each \( \gamma \) term of this weight and signature in such a way that different \( \beta \) terms are assigned to different \( \gamma \) terms.

63. We consider expressions \( EX_h \) of the following description. If \( E \) is an \( \alpha \) term it may be completely arbitrary. If \( E \) is a \( \beta \) term it is restricted by \( h \), the exact statement of the restriction requiring that \( E \) be written out explicitly. In this case let

\[
E = u_{i_1}^{a_1} \cdots u_{i_r}^{a_r} v_{j_1}^{b_1} \cdots v_{j_s}^{b_s}
\]

where the \( a_i \) and \( b_i \) are positive integers. We might state explicitly that the subscripts satisfy the relations \( i_1 < \cdots < i_r \) and \( j_1 < \cdots < j_s \). Let \( t \) be the smallest integer for which \( a_1 + a_2 + \cdots + a_t > j_1 \). Our restriction on \( E \) is that \( i_t + j_1 \geq h \).

A procedure for associating a \( \beta \) term with such an expression will now be described, the \( \beta \) term to have the same weight and signature as \( EX_h \). The \( \beta \) term will have the general appearance \( u_{\alpha} v_{\beta} E \) where \( a+b = h \). Thus requirements of weight and signature will evidently be met. Some preliminary calculations must be made before the \( \beta \) term can actually be produced.

64. Let \( E \) be given by equation (3.2). Let \( e_i = a_1 + a_2 + \cdots + a_i \), \( i = 1, \cdots, r \). Let \( f_1 = i_1 \) and \( f_{2n+1} = i_{n+1} + e_n, n = 1, \cdots, r - 1 \). Let \( f_{2n} = i_n + e_n, n = 1, \cdots, r \). Finally let \( f_{2r+1} \) be an integer which exceeds both \( h \) and \( f_{2r} \). Obviously \( f_1 < f_2 < \cdots < f_{2r+1} \). Then either \( h \leq f_1 \) or there is a positive integer \( c \) such that \( f_c < h \leq f_{c+1} \). Three cases are treated, depending on the relation \( h \) bears to the \( f_i \). The cases are

(i) \( h \leq f_1 \). \( E' \) is defined by

\[
E' = u_{h} v_{E}.
\]

(ii) The integer \( c \) mentioned above is odd, say \( 2d - 1 \), so that \( f_{2d-1} < h \leq f_{2d} \). Then \( E' \) is defined by

\[
E' = u_{i_d} v_{h-i_d} E.
\]
(iii) The integer \( c \) is even, say \( 2d \), so that \( f_{2d} < h < f_{2d+1} \). Then \( E' \) is defined by

\[
E' = u_{h-e_d}v_{e_d}E.
\]

**Note.** We admit the possibility that \( E \) contains no \( u_i \), effectively so that the quantities \( e_i \) cannot be computed. For this case the quantity \( f_{2r+1} \) is simply \( f_1 \), and the case is covered by (i).

The problem now is to show that these assignments always lead to \( \beta \) terms and that distinct expressions are assigned to distinct \( \beta \) terms. It will first be shown that this procedure always leads to a \( \beta \) term, and then the following characterization of the \( u_a \) and \( v_b \) used as factors with the \( E \) will be obtained. The \( v_b \) will be shown to be such that no \( v_k \) is effectively present in \( E' \) with \( k < b \). The \( u_a \) will be shown to be such that for no \( k < a \) does the degree of \( E' \) in \( u, u_1, \ldots, u_k \) exceed \( b \). The three cases will be treated separately and for each one the validity of these remarks will be shown.

65. For case (i) the fact that \( E' \) contains the factor \( u_{h}v \) proves that \( E' \) is a \( \beta \) term. It is evident that \( E' \) contains no \( v_k \) with \( k < b \) because \( b \) in this case is zero. Finally the fact that \( h \leq i_1 \) and \( a = h \) shows that \( E' \) contains no \( u_j \) effectively with \( j < a \).

66. For case (ii) it is desirable to write out the inequality \( f_{2d-1} < h \leq f_{2d} \) in full. It states

\[
(3.3) \quad i_d + a_1 + \cdots + a_{d-1} < h \leq i_d + a_1 + \cdots + a_d.
\]

(If \( d \) is unity this is to mean \( i_1 < h \leq i_1 + a_1 \).) In this case \( E \) is multiplied by \( v_b \) with \( b = h - i_d \). The degree of \( E' \) in the \( u_i \) exceeds that of \( E \) in the \( u_i \) and the latter degree is certainly not less than \( a_1 + \cdots + a_d \). It is a consequence of (3.3) that \( a_1 + \cdots + a_d \) is not less than \( h - i_d \) so that the degree of \( E' \) in the \( u_i \) exceeds \( h - i_d \). Then \( E' \) is a \( \beta \) term since it contains \( v_b \) effectively with \( b \) less than the degree of \( E' \) in the \( u_i \). To show that \( E' \) contains no \( v_k \) with \( k < b \) observe that then \( E \) would also contain this \( v_k \). Such an integer would be less than \( a_1 + \cdots + a_d \) and this fact with the supposition \( k + i_d < b + i_d = h \) would mean that \( E \) did not obey the restriction imposed on it. To show that the factor \( u_a \) has the property described, note that it is a consequence of (3.3) that \( a_1 + \cdots + a_{d-1} < h - i_d = b \). Since \( E' \) is identical with \( E \) as far as the letters \( u, u_1, \ldots, u_{i_d-1} \) are concerned and \( a \) is \( i_d \), this inequality shows that for no \( k < a \) does the degree of \( E' \) in \( u, u_1, \ldots, u_k \) exceed \( b \).

67. We now turn to case (iii). The inequality \( f_{2d} < h < f_{2d+1} \) written out in full becomes

\[
(3.4) \quad i_d + a_1 + \cdots + a_d < h \leq i_{d+1} + a_1 + \cdots + a_d
\]

(for \( d = r \) this is to mean \( i_r + a_1 + \cdots + a_r < h \)).

The factor \( v_b \) used with \( E \) is in this case defined by \( b = a_1 + \cdots + a_d \). Since the degree of \( E' \) in the \( u_i \) exceeds that of \( E \), and since this latter degree is at
least $b$, it follows that $E'$ is a $\beta$ term because it contains $v_b$ effectively. To show that $E'$ contains no $v_k$ effectively with $k < b$ observe that then $E$ would also contain this $v_k$. It follows from (3.4) that $i_d + k < h$ and since $b = a_1 + \cdots + a_d > k$ the restriction imposed on $E$ could not be satisfied. We now show that $u_a$ is such that for all $k < a$ the degree of $E'$ in $u$, $u_1$, \ldots, $u_k$ does not exceed $b$. This degree is the same as that of $E$ in these letters. The integer $a$ is defined as $h - e_d$. It follows from (3.4) that $i_d + h - e_d \leq i_{d+1}$. The degree of $E'$ in the letters $u_k$ with $k < a$ is thus $a_1 + \cdots + a_d$. This number is precisely $b$, which verifies the statement.

68. It can now be shown that this procedure assigns distinct $\beta$ terms $E'$ to distinct expressions $EX_h$. Let $E'$ be any $\beta$ term. It must contain some of the $v_i$ effectively. Let $b$ be the smallest subscript for which $v_b$ is effectively present in $E'$. Since $E'$ is a $\beta$ term its degree in the $u_i$ exceeds $b$. Let $a$ be the smallest integer such that the degree of $E'$ in $u$, $u_1$, \ldots, $u_a$ exceeds $b$. Then if $EX_h$ led to $E'$ by the method described above, it must have been that $E' = u_av_bE$, $h = a + b$. Thus given $E'$ there is only one possibility for $EX_h$.

We need the additional fact that if $EX_h$ determines $E'$ as above, and if $g$ is any non-negative integer such that $g \leq h$, then $E'X_g$ is also an admissible expression. This means that if $k$ is such that the degree of $E'$ in $u$, $u_1$, \ldots, $u_k$ exceeds $t$ and $v_t$ actually appears in $E'$, then $k + t \geq g$. It has already been shown that $a$ and $b$ are the smallest relevant integers and since $a + b = h$ it follows that $a + b \geq g$.

69. We are now in a position to describe the way in which $\beta$ terms may be associated with $\gamma$ terms. Let $G = E_rX_{i_1}X_{i_2} \cdots X_{i_r}$ be a $\gamma$ term (it is assumed that $i_1 \leq i_2 \leq \cdots \leq i_r$, and that $r \geq 1$). $E_r$ is an $\alpha$ term, so that $EX_{i_r}$ is an expression of the type considered and determines a $\beta$ term $E_{r-1}$ of the same weight and signature. If $r = 1$ this $\beta$ term is associated with $G$. If $r > 1$ it has been shown that $E_{r-1}X_{i_{r-1}}$ is also an admissible expression and determines a $\beta$ term $E_{r-2}$. $E_{r-2}$ has the same weight and signature as $E_rX_{i_r}$. If $r = 2$ then $E_{r-2}$ is associated with $G$. If $r > 2$ we continue. In this way a sequence of $\beta$ terms $E_{r-1}, E_{r-2}, \cdots, E_0$ is obtained where each $E_{j-1}$ is the $\beta$ term determined by $E_jX_{i_j}$. $E_0$ has the same weight and signature as $G$. We associate it with $G$.

It will be shown that if $G$ and $G'$ are $\gamma$ terms associated with $E_0$ and $E_0'$, and if $E_0 = E_0'$, then $G$ and $G'$ are the same. Let $G' = E'_rX_{i_1}X_{i_2} \cdots X_{i_r}$ and $G$ are to be considered as identical if and only if $E_r = E'_r$, $r = s$ and $(i_1, \cdots, i_r) = (j_1, \cdots, j_s)$. Suppose $E_0 = E_0'$. $E_0$ was determined by $E_1X_{i_1}$ and $E_0'$ by $E'_1X_{j_1}$. Because of the uniqueness property of the procedure used, it follows that $E_1 = E'_1$ and $i_1 = j_1$. Suppose $r \leq s$. Then by reconstructing $E_2X_{i_1}, \cdots, E_rX_{i_r}$ it follows that $E_k = E'_k$ and $i_k = j_k$ for $k = 1, \cdots, r$. Now $E_r$ is an $\alpha$ term, so that $E'_r$ is also an $\alpha$ term. The only power product in the sequence $E_0'$, $E'_1$, \ldots, $E'_r$ which is an $\alpha$ term is $E'_s$. It follows that $s = r$ and thus $G = G'$.
It has been shown that every $\gamma$ term determines a $\beta$ term having the same weight and signature and that distinct $\gamma$ terms determine distinct $\beta$ terms. This completes the proof.

**The structure of the ideal of $uv$**

70. We now can prove the following lemma.

**Lemma 3.4.** Let $d_1, d_2, w$ be non-negative integers. Let $n_\alpha$ denote the number of $\alpha$ terms $A_i$ of signature $(d_1, d_2)$ and weight $w$, let $n_\beta$ denote the number of $\beta$ terms of this signature and weight, and let $n_\gamma$ denote the corresponding number of $\gamma$ terms $G_j$. Then $n_\beta = n_\gamma$ and a relation

$$\sum_{i=1}^{n_\alpha} a_i A_i + \sum_{j=1}^{n_\gamma} g_j G_j = 0$$

where the $a_i$ and $g_j$ are rational numbers implies that all the $a_i$ and $g_j$ are zero.

The proof of this lemma is identical with that of Lemma 1.4, so that no further argument will be given. It will be noted that if the $a_i$ and $g_j$ are elements of any differential domain of integrity which contains the rational numbers and over which $u$ and $v$ are unknowns, the same conclusion can be drawn.

71. The above lemmas combine to yield the following theorem.

**Theorem 3.1.** Let $3$ be any differential domain of integrity which contains the rational numbers. Let $F$ be any differential polynomial in the unknowns $u$ and $v$. Then $F$ is expressible in the form

$$F = \sum a_i A_i + \sum g_j G_j, \quad a_i, g_j \in 3,$$

where the $A_i$ are $\alpha$ terms and the $G$ are $\gamma$ terms. For each $F$ there is only one such expression.

For the proof of this theorem the reader is referred to the proof of Theorem 1.1.

**Corollary.** Let $d_1, d_2, w$ be as in the statement of Lemma 3.4. Let the $3$ considered above be a field. Then the number of linearly independent (mod $\Omega$) forms with coefficients in $3$ which are homogeneous and isobaric of this signature and weight is $n_\alpha$.

**Corollary.** No linear combination of $\alpha$ terms with coefficients in $3$ is in $\Omega$.

**Indecomposability of the ideal of $uv$**

72. The preceding work enables us to indicate a striking difference between ordinary (algebraic) ideals of polynomials in a finite number of unknowns and differential ideals of such polynomials. This difference is manifested by the differential ideal of $uv$, as we proceed to show. The mani-
fold of this ideal is reducible into the union of two irreducible manifolds, namely \( u = 0 \) with \( v \) arbitrary and \( v = 0 \) with \( u \) arbitrary. Nonetheless we are going to show that the differential ideal \([uv]\) has no representation as the intersection or product of two differential ideals whose manifolds are respectively the first and the second just described. Let \( \Sigma_1 \) be any differential ideal of differential polynomials in the unknowns \( u \) and \( v \) whose manifold is \( u = 0 \) with \( v \) arbitrary. Then(18) \( \Sigma_1 \) contains some power of \( u \), say \( u^r \). Again if \( \Sigma_2 \) is a similar differential ideal whose manifold is \( v = 0 \) with \( u \) arbitrary, then some power \( v^t \) belongs to \( \Sigma_2 \). Suppose that the ideal \([uv]\) had a representation as the intersection or the product of \( \Sigma_1 \) and \( \Sigma_2 \). Since \( \Sigma_2 \) contains \( v^t \), it contains some power \( v^t \), and the form \( u^r v^t \) is in the product and intersection of \( \Sigma_1 \) and \( \Sigma_2 \). This form is an \( \alpha \) term and is thus not in \([uv]\). This proves our contention.

### The power products in the ideal of \( uv \)

73. **Theorem 3.2.** Let \( d_1, d_2, w, n_\alpha \) be as in Lemma 3.4. A necessary and sufficient condition that \( n_\alpha > 0 \) is that \( w \geq d_1 d_2 \).

This is equivalent to the assertion that every power product of signature \((d_1, d_2)\) and weight \( w < d_1 d_2 \) is in \( \Omega \) and not every power product of this signature and weight \( w \leq d_1 d_2 \) is in \( \Omega \).

We need only investigate the circumstances under which \( \alpha \) terms exist. Let \( w = d_1 d_2 + h \) with \( h \geq 0 \). Then \( u^{d_1} v^{d_2 - 1} v_{d_1 + h} \) is an \( \alpha \) term of signature \((d_1, d_2)\) and weight \( w \). This disposes of the sufficiency condition. We now show that there are no \( \alpha \) terms of signature \((d_1, d_2)\) and weight less than \( d_1 d_2 \).

If \( d_1 d_2 = 0 \), there are no power products with this property and certainly no \( \alpha \) terms. If \( d_1 d_2 > 0 \), such an \( \alpha \) term would have to be such that every \( v_k \) effectively present in it would have a subscript not less than \( d_1 \). Since it must have \( d_2 \) such letters \( v_k \) its weight is at least \( d_1 d_2 \).

74. It might be pointed out that this discussion of \( \Omega \) applies to more general ideals, in that the \( X_i \) need not be the \( i \)th derivative of \( uv \). If each \( X_i \) is a homogeneous isobaric polynomial in the \( u_i \) and \( v_i \) of signature \((1, 1)\) and weight \( i \), and is such that every term of this signature and weight is present in \( X_i \) with a nonzero coefficient, then the whole discussion applies verbatim. The coefficients of the \( X_i \) must be confined to some field and may otherwise be arbitrary.

---

(18) Raudenbush, loc. cit.

Columbia University,
New York, N. Y.