INTRODUCTION TO HIGHER ALGEBRA

CHAPTER I

POLYNOMIALS AND THEIR MOST FUNDAMENTAL PROPERTIES

1. Polynomials in One Variable. By an integral rational function of x, or, as we shall say for brevity, a polynomial in x, is meant a function of x determined by an expression of the form

$$c_1x^{\alpha_1} + c_2x^{\alpha_2} + \cdots + c_kx^{\alpha_k}, \tag{1}$$

where the α's are integers positive or zero, while the c's are any constants, real or imaginary. We may without loss of generality assume that no two of the α's are equal. This being the case, the expressions $c_ix^{\alpha_i}$ are called the *terms* of the polynomial, c_i is called the *coefficient* of this term, and α_i is called its *degree*. The highest degree of any term whose coefficient is not zero is called the *degree of the polynomial.*

It should be noticed that the conceptions just defined — terms, coefficients, degree — apply not to the polynomial itself, but to the particular *expression* (1) which we use to determine the polynomial, and it would be quite conceivable that one and the same function of x might be given by either one of two wholly different expressions of the form (1). We shall presently see (cf. Theorem 5 below) that this cannot be the case except for the obvious fact that we may insert in or remove from (1) any terms we please with zero coefficients.

By arranging the terms in (1) in the order of decreasing α's and supplying, if necessary, certain missing terms with zero coefficients, we may write the polynomial in the normal form

$$a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n. \tag{2}$$

It should, however, constantly be borne in mind that a polynomial in this form is not necessarily of the nth degree; but will be of the nth degree when and only when $a_0 \neq 0$.

DEFINITION. *Two polynomials, $f_1(x)$ and $f_2(x)$, are said to be identically equal ($f_1 \equiv f_2$) if they are equal for all values of x. A polynomial $f(x)$ is said to vanish identically ($f \equiv 0$) if it vanishes for all values of x.*

We learn in elementary algebra how to add, subtract, and multiply * polynomials; that is, when two polynomials $f_1(x)$ and $f_2(x)$ are given, to form new polynomials equal to the sum, difference, and product of these two.

THEOREM 1. *If the polynomial*

$$f(x) \equiv a_0 x^n + a_1 x^{n-1} + \cdots + a_n$$

vanishes when $x = \alpha$, there exists another polynomial

$$\phi_1(x) \equiv a_0 x^{n-1} + a_1' x^{n-2} + \cdots + a_{n-1}',$$

such that $$f(x) \equiv (x-\alpha)\phi_1(x).$$

For since by hypothesis $f(\alpha) = 0$, we have

$$f(x) \equiv f(x) - f(\alpha) \equiv a_0(x^n - \alpha^n) + a_1(x^{n-1} - \alpha^{n-1}) + \cdots + a_{n-1}(x-\alpha).$$

Now by the rule of elementary algebra for multiplying together two polynomials we have

$$x^k - \alpha^k \equiv (x-\alpha)(x^{k-1} + \alpha x^{k-2} + \cdots + \alpha^{k-1}).$$

Hence

$$f(x) \equiv (x-\alpha)[a_0(x^{n-1} + \alpha x^{n-2} + \cdots + \alpha^{n-1}) + a_1(x^{n-2} + \alpha x^{n-3} + \cdots + \alpha^{n-2}) + \cdots + a_{n-1}].$$

If we take as $\phi_1(x)$ the polynomial in brackets, our theorem is proved.

Suppose now that β is another value of x distinct from α for which $f(x)$ is zero. Then

$$f(\beta) = (\beta - \alpha)\phi_1(\beta) = 0;$$

* The question of division is somewhat more complicated and will be considered in § 63.

and since $\beta - \alpha \neq 0$, $\phi_1(\beta) = 0$. We can therefore apply the theorem just proved to the polynomial $\phi_1(x)$, thus getting a new polynomial

$$\phi_2(x) \equiv a_0 x^{n-2} + a_1'' x^{n-3} + \cdots + a_{n-2}'',$$

such that $$\phi_1(x) \equiv (x-\beta)\phi_2(x),$$

and therefore $$f(x) \equiv (x-\alpha)(x-\beta)\phi_2(x).$$

Proceeding in this way, we get the following general result:

THEOREM 2. *If $\alpha_1, \alpha_2, \cdots \alpha_k$ are k distinct constants, and if*

$$f(x) \equiv a_0 x^n + a_1 x^{n-1} + \cdots + a_n \qquad (n \geq k),$$

and $$f(\alpha_1) = f(\alpha_2) = \cdots = f(\alpha_k) = 0,$$

then $$f(x) \equiv (x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_k)\phi(x),$$

where $$\phi(x) \equiv a_0 x^{n-k} + b_1 x^{n-k-1} + \cdots + b_{n-k}.$$

Applying this theorem in particular to the case $n = k$, we see that if the polynomial $f(x)$ vanishes for n distinct values $\alpha_1, \alpha_2, \cdots \alpha_n$ of x, then

$$f(x) \equiv a_0(x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_n).$$

Accordingly, if $a_0 \neq 0$, there can be no value of x other than $\alpha_1, \cdots \alpha_n$ for which $f(x) = 0$. We have thus proved

THEOREM 3. *A polynomial of the nth degree in x cannot vanish for more than n distinct values of x.*

Since the only polynomials which have no degree are those all of whose coefficients are zero, and since such polynomials obviously vanish identically, we get the fundamental result:

THEOREM 4. *A necessary and sufficient condition that a polynomial in x vanish identically is that all its coefficients be zero.*

Since two polynomials in x are identically equal when and only when their difference vanishes identically, we have

THEOREM 5. *A necessary and sufficient condition that two polynomials in x be identically equal is that they have the same coefficients.*

This theorem shows, as was said above, that the terms, coefficients, and degree of a polynomial depend merely on the polynomial itself, not on the special way in which it is expressed.

2. Polynomials in More than One Variable. A function of (x, y) is called a polynomial if it is given by an expression of the form

$$c_1x^{\alpha_1}y^{\beta_1} + c_2x^{\alpha_2}y^{\beta_2} + \cdots + c_kx^{\alpha_k}y^{\beta_k},$$

where the α's and β's are integers positive or zero.

More generally, a function of $(x_1, x_2, \cdots x_n)$ is called a polynomial if it is determined by an expression of the form

$$(1) \qquad c_1x_1^{\alpha_1}x_2^{\beta_1}\cdots x_n^{\nu_1} + c_2x_1^{\alpha_2}x_2^{\beta_2}\cdots x_n^{\nu_2} + \cdots + c_kx_1^{\alpha_k}x_2^{\beta_k}\cdots x_n^{\nu_k},$$

where the α's, β's, $\cdots$ ν's are integers positive or zero.

Here we may assume without loss of generality that in no two terms are the exponents of the various x's the same; that is, that if

$$\alpha_i = \alpha_j,\ \beta_i = \beta_j,\ \cdots\ \mu_i = \mu_j,$$

then

$$\nu_i \neq \nu_j.$$

This assumption being made, $c_ix_1^{\alpha_i}x_2^{\beta_i}\cdots x_n^{\nu_i}$ is called a *term* of the polynomial, c_i its *coefficient*, α_i the *degree of the term in* x_1, β_i in x_2, etc., and $\alpha_i + \beta_i + \cdots + \nu_i$ the *total degree*, or simply the degree, of the term. The highest degree in x_i of any term in the polynomial whose coefficient is not zero is called the *degree of the polynomial in* x_i, and the highest total degree of any term whose coefficient is not zero is called the *degree of the polynomial.*

Here, as in § 1, the conceptions just defined apply for the present not to the function itself but to the special method of representing it by an expression of the form (1). We shall see presently, however, that this method is unique.

Before going farther, we note explicitly that *according to the definition we have given, a polynomial all of whose coefficients are zero has no degree.*

When we speak of a polynomial in n variables, we do not necessarily mean that all n variables are actually present. One or more of them may have the exponent zero in every term, and hence not appear at all. Thus a polynomial in one variable, or even a constant, may be regarded as a special case of a polynomial in any larger number of variables.

A polynomial all of whose terms are of the same degree is said to be *homogeneous*. Such polynomials we will speak of as *forms*,*

* There is diversity of usage here. Some writers, following Kronecker, apply the term *form* to all polynomials. On the other hand, homogeneous polynomials are often spoken of as *quantics* by English writers.

distinguishing between *binary*, *ternary*, *quaternary*, and in general *n-ary* forms according to the number of variables involved, binary forms involving two, ternary three, etc.

Another method of classifying forms is according to their degree. We speak here of linear forms, quadratic forms, cubic forms, etc., according as the degree is 1, 2, 3, etc. We will, however, agree that a polynomial all of whose coefficients are zero may also be spoken of indifferently as a linear form, quadratic form, cubic form, etc., in spite of the fact that it has no degree.

If all the coefficients of a polynomial are real, it is called a *real polynomial* even though, in the course of our work, we attribute imaginary values to the variables.

It is frequently convenient to have a polynomial in more than one variable arranged according to the descending powers of some one of the variables. Thus a normal form in which we may write a polynomial in n variables is

$$\phi_0(x_2, \cdots x_n)x_1^m + \phi_1(x_2, \cdots x_n)x_1^{m-1} + \cdots + \phi_m(x_2, \cdots x_n),$$

the ϕ's being polynomials in the $n-1$ variables $(x_2, \cdots x_n)$.

We learn in elementary algebra how to add, subtract, and multiply polynomials, getting as the result new polynomials.

DEFINITION. *Two polynomials in any number of variables are said to be identically equal if they are equal for all values of the variables. A polynomial is said to vanish identically if it vanishes for all values of the variables.*

THEOREM 1. *A necessary and sufficient condition that a polynomial in any number of variables vanish identically is that all its coefficients be zero.*

That this is a sufficient condition is at once obvious. To prove that it is a necessary condition we use the method of mathematical induction. Since we know that the theorem is true in the case of one variable (Theorem 4, § 1), the theorem will be completely proved if we can show that if it is true for a certain number $n-1$ of variables, it is true for n variables.

Suppose, then, that

$$f(x_1, \cdots x_n) \equiv \phi_0(x_2, \cdots x_n)x_1^m + \phi_1(x_2, \cdots x_n)x_1^{m-1} + \cdots + \phi_m(x_2, \cdots x_n)$$

vanishes identically. If we assign to $(x_2, \cdots x_n)$ any fixed values $(x_2', \cdots x_n')$, f becomes a polynomial in x_1 alone, which, by hypothesis, vanishes for all values of x_1. Hence its coefficients must, by Theorem 4, § 1, all be zero:

$$\phi_i(x_2', \cdots x_n') = 0 \qquad (i = 0, 1, \cdots m).$$

That is, the polynomials $\phi_0, \phi_1, \cdots \phi_m$ vanish for all values of the variables, since $(x_2', \cdots x_n')$ was *any* set of values. Accordingly, by the assumption we have made that our theorem is true for polynomials in $n-1$ variables, all the coefficients of all the polynomials $\phi_0, \phi_1, \cdots \phi_m$ are zero. These, however, are simply the coefficients of f. Thus our theorem is proved.

Since two polynomials are identically equal when and only when their difference is identically zero, we infer now at once the further theorem:

THEOREM 2. *A necessary and sufficient condition that two polynomials be identically equal is that the coefficients of their corresponding terms be equal.*

We come next to

THEOREM 3. *If f_1 and f_2 are polynomials in any number of variables of degrees m_1 and m_2 respectively, the product $f_1 f_2$ will be of degree $m_1 + m_2$.*

This theorem is obviously true in the case of polynomials in one variable. If, then, assuming it to be true for polynomials in $n-1$ variables we can prove it to be true for polynomials in n variables, the proof of our theorem by the method of mathematical induction will be complete.

Let us look first at the special case in which both polynomials are homogeneous. Here every term we get by multiplying them together by the method of elementary algebra is of degree $m_1 + m_2$. Our theorem will therefore be proved if we can show that there is at least one term in the product whose coefficient is not zero. For this purpose, let us arrange the two polynomials f_1 and f_2 according to descending powers of x_1,

$$f_1(x_1, \cdots x_n) \equiv \phi_0'(x_2, \cdots x_n)x_1^{k_1} + \phi_1'(x_2, \cdots x_n)x_1^{k_1-1} + \cdots,$$

$$f_2(x_1, \cdots x_n) \equiv \phi_0''(x_2, \cdots x_n)x_1^{k_2} + \phi_1''(x_2, \cdots x_n)x_1^{k_2-1} + \cdots.$$

Here we may assume that neither ϕ_0' nor ϕ_0'' vanishes identically. Since f_1 and f_2 are homogeneous, ϕ_0' and ϕ_0'' will also be homogeneous

of degrees $m_1 - k_1$ and $m_2 - k_2$ respectively. In the product $f_1 f_2$ the terms of highest degree in x_1 will be those in the product

$$\phi_0'(x_2, \cdots x_n)\phi_0''(x_2, \cdots x_n)x_1^{k_1+k_2},$$

and since we assume our theorem to hold for polynomials in $n-1$ variables, $\phi_0'\phi_0''$ will be a polynomial of degree $m_1 + m_2 - k_1 - k_2$. Any term in this product whose coefficient is not zero gives us when multiplied by $x_1^{k_1+k_2}$ a term of the product $f_1 f_2$ of degree $m_1 + m_2$ whose coefficient is not zero. Thus our theorem is proved for the case of homogeneous polynomials.

Let us now, in the general case, write f_1 and f_2 in the forms

$$f_1(x_1, \cdots x_n) \equiv \phi_{m_1}'(x_1, \cdots x_n) + \phi_{m_1-1}'(x_1, \cdots x_n) + \cdots,$$

$$f_2(x_1, \cdots x_n) \equiv \phi_{m_2}''(x_1, \cdots x_n) + \phi_{m_2-1}''(x_1, \cdots x_n) + \cdots,$$

where ϕ_i' and ϕ_j'' are homogeneous polynomials which are either of degrees i and j respectively, or which vanish identically. Since, by hypothesis, f_1 and f_2 are of degrees m_1 and m_2 respectively, ϕ_{m_1}' and ϕ_{m_2}'' will not vanish identically, but will be of degrees m_1 and m_2.

The terms of highest degree in the product $f_1 f_2$ will therefore be the terms of the product $\phi_{m_1}'\phi_{m_2}''$, and this being a product of homogeneous polynomials comes under the case just treated and is therefore of degree $m_1 + m_2$. The same is therefore true of the product $f_1 f_2$, and our theorem is proved.

By a successive application of this theorem we infer

COROLLARY. *If k polynomials are of degrees $m_1, m_2, \cdots m_k$ respectively, their product is of degree $m_1 + m_2 + \cdots + m_k$.*

We mention further, on account of their great importance, the two rather obvious results:

THEOREM 4. *If the product of two or more polynomials is identically zero, at least one of the factors must be identically zero.*

For if none of them were identically zero, they would all have definite degrees, and therefore their product would, by Theorem 3, have a definite degree, and would therefore not vanish identically.

It is from this theorem that we draw our justification for cancelling out from an identity a factor which we know to be not identically zero.

THEOREM 5. *If $f(x_1, \cdots x_n)$ is a polynomial which is not identically zero, and if $\phi(x_1, \cdots x_n)$ vanishes at all points where f does not vanish, then ϕ vanishes identically.*

This follows from Theorem 4 when we notice that $f\phi \equiv 0$.

EXERCISES

1. If f and ϕ are polynomials in any number of variables, what can be inferred from the identity $f^2 \equiv \phi^2$ concerning the relation between the polynomials f and ϕ?

2. If f_1 and f_2 are polynomials in $(x_1, \cdots x_n)$ which are of degrees m_1 and m_2 respectively in x_1, prove that their product is of degree $m_1 + m_2$ in x_1.

3. Geometric Interpretations. In dealing with functions of a single real variable, the different values which the variable may take on may be represented geometrically by the points of a line; it being understood that when we speak of a point x we mean the point which is situated on the line at a distance of x units (to the right or left according as x is positive or negative) from a certain fixed origin O, on the line. Similarly, in the case of functions of two real variables, the sets of values of the variables may be pictured geometrically by the points of a plane, and in the case of three real variables, by the points of space; the set of values represented by a point being, in each case, the *rectangular* coördinates of that point. When we come to functions of four or more variables, however, this geometric representation is impossible.

The complex variable $x = \xi + \eta i$ depends on the two independent real variables ξ and η in such a way that to every pair of real values (ξ, η) there corresponds one and only one value of x. The different values which a single complex variable may take on may, therefore, be represented by the points of a plane in which (ξ, η) are used as cartesian coördinates. In dealing with functions of more than one complex variable, however, this geometric representation is impossible, since even two complex variables $x = \xi + \eta i$, $y = \xi_1 + \eta_1 i$ are equivalent to four real variables $(\xi, \eta, \xi_1, \eta_1)$.

By the *neighborhood of a point* $x = a$ we mean that part of the line between the points $x = a - \alpha$ and $x = a + \alpha$ (α being an arbitrary positive constant, large or small), or what is the same thing, all points whose coördinates x satisfy the inequality $|x - a| < \alpha$.*

* We use the symbol $|Z|$ to denote the absolute value of Z, *i.e.* the numerical value of Z if Z is real, the modulus of Z if Z is imaginary.

Similarly, by the neighborhood of a point (a, b) in a plane, we shall mean all points whose coördinates (x, y) satisfy the inequalities

$$|x - a| < \alpha, \quad |y - b| < \beta,$$

where α and β are positive constants. This neighborhood thus consists of the interior of a rectangle of which (a, b) is the center and whose sides are parallel to the coördinate axes.

By the neighborhood of a point (a, b, c) in space we mean all points whose coördinates (x, y, z) satisfy the inequalities

$$|x - a| < \alpha, \quad |y - b| < \beta, \quad |z - c| < \gamma.$$

In all these cases it will be noticed that the neighborhood may be large or small according to the choice of the constants α, β, γ.

If we are dealing with a single complex variable $x = \xi + \eta i$, we understand by the neighborhood of a point a all points in the plane of complex quantities whose complex coördinate x satisfies the inequality $|x - a| < \alpha$, α being as before a real positive constant. Since $|x - a|$ is equal to the distance between x and a, the neighborhood of a now consists of the interior of a circle of radius α described about a as center.

It is found convenient to extend the geometric terminology we have here introduced to the case of any number of real or complex variables. Thus if we are dealing with n independent variables $(x_1, x_2, \cdots x_n)$, we speak of any particular set of values of these variables as a *point in space of n dimensions*. Here we have to distinguish between *real points*, that is sets of values of the x's which are all real, and *imaginary points* in which this is not the case. In using these terms we do not propose even to raise the question whether in any geometric sense there is such a thing as space of more than three dimensions. We merely use these terms in a wholly conventional algebraic sense because on the one hand they have the advantage of conciseness over the ordinary algebraic terms, and on the other hand, by calling up in our minds the geometric pictures of three dimensions or less, this terminology is often suggestive of new relations which might otherwise not present themselves to us so readily.

By the neighborhood of the point $(a_1, a_2, \cdots a_n)$ we understand all points which satisfy the inequalities

$$|x_1 - a_1| < \alpha_1, \quad |x_2 - a_2| < \alpha_2, \cdots |x_n - a_n| < \alpha_n,$$

where $\alpha_1, \alpha_2, \cdots \alpha_n$ are real positive constants.

If, in particular, $(a_1, a_2, \cdots a_n)$ is a real point, we may speak of the real neighborhood of this point, meaning thereby all real points $(x_1, x_2, \cdots x_n)$ which satisfy the above inequalities.

As an illustration of the use to which the conception of the neighborhood of a point can be put in algebra, we will prove the following important theorem:

THEOREM 1. *A necessary and sufficient condition that a polynomial $f(x_1, \cdots x_n)$ vanish identically is that it vanish throughout the neighborhood of a point $(a_1, \cdots a_n)$.*

That this is a necessary condition is obvious. To prove that it is sufficient we begin with the case $n = 1$.

Suppose then that $f(x)$ vanishes throughout a certain neighborhood of the point $x = a$. If $f(x)$ did not vanish identically, it would be of some definite degree, say k, and therefore could not vanish at more than k points (cf. Theorem 3, § 1). This, however, is not the case, since it vanishes at an infinite number of points, namely all points in the neighborhood of $x = a$. Thus our theorem is proved in the case $n = 1$.

Turning now to the case $n = 2$, let

$$f(x, y) \equiv \phi_0(y)x^k + \phi_1(y)x^{k-1} + \cdots + \phi_k(y)$$

be a polynomial which vanishes throughout a certain neighborhood of the point (a, b), say when

$$|x - a| < \alpha, \quad |y - b| < \beta.$$

Let y_0 be any constant satisfying the inequality

$$|y_0 - b| < \beta.$$

Then $f(x, y_0)$ is a polynominal in x alone which vanishes whenever $|x - a| < \alpha$. Hence, by the case $n = 1$ of our theorem, $f(x, y_0) \equiv 0$. That is,

$$\phi_0(y_0) = \phi_1(y_0) = \cdots = \phi_k(y_0) = 0.$$

Thus all these polynomials ϕ vanish at every point y_0 in the neighborhood of $y = b$, and therefore, by the case $n = 1$ of our theorem, they are all identically zero. From this it follows that for every value of x, $f(x, y)$ vanishes for all values of y, that is $f \equiv 0$, and our theorem is proved.

We leave to the reader the obvious extension of this method of proof to the case of n variables by the use of mathematical induction.

From the theorem just proved we can infer at once the following:

THEOREM 2. *A necessary and sufficient condition that two polynomials in the variables $(x_1, \cdots x_n)$ be identically equal is that they be equal throughout the neighborhood of a point $(a_1, \cdots a_n)$.*

EXERCISES

1. Theorem 3, § 1 may be stated as follows: If f is a polynomial in one variable which is known not to be of degree higher than n, then if f vanishes at $n + 1$ distinct points, it vanishes identically.

Establish the following generalization of this theorem:

If f is a polynomial in (x, y) which is known not to be of higher degree than n in x, and not of higher degree than m in y, then, if f vanishes at the $(n+1)(m+1)$ distinct points:

$$(x_i, y_j) \qquad \begin{pmatrix} i = 1, 2, \cdots n+1 \\ j = 1, 2, \cdots m+1 \end{pmatrix},$$

it vanishes identically.

2. Generalize the theorem of Exercise 1 to polynomials in any number of variables.

3. Prove Theorem 4, § 2 by means of Theorem 1 of the present section; and from this result deduce Theorem 3, § 2.

4. Do Theorems 1 and 2 of this section hold if we consider only real polynomials and the real neighborhoods of real points?

4. Homogeneous Coördinates. Though only two quantities are necessary in order to locate the position of a point in a plane, it is frequently more convenient to use three, the precise values of the quantities being of no consequence, but only their ratios. We will represent these three quantities by x, y, t, and define their ratios by the equations

$$\frac{x}{t} = X, \qquad \frac{y}{t} = Y,$$

where X and Y are the cartesian coördinates of a point in a plane. Thus (2, 3, 5) will represent the point whose abscissa is $\frac{2}{5}$ and whose ordinate is $\frac{3}{5}$. Any set of three numbers which are proportional to

(2, 3, 5) will represent the same point. So that, while to every set of three numbers (with certain exceptions to be noted below) there corresponds one and only one point, to each point there correspond an infinite number of different sets of three numbers, all of which, however, are proportional.

When $t=0$ our definition is meaningless; but if we consider the points (2, 3, 1), (2, 3, 0.1), (2, 3, 0.01), (2, 3, 0.001), $\cdots$, which are, in cartesian coördinates, the points (2, 3), (20, 30), (200, 300), (2000, 3000), $\cdots$, we see that they all lie on the straight line through the origin whose slope is $\frac{3}{2}$. Thus as t approaches zero, x and y remaining fixed but not both zero, the point (x, y, t) moves away along a straight line through the origin whose slope is y/x. Hence it is natural to speak of $(x, y, 0)$ as *the point at infinity* on the line whose slope is y/x. If t approaches zero through negative values, the point will move off along *the same* line, but in the opposite direction. We will not distinguish between these two cases, but will speak of *only one* point at infinity on any particular line. It can be easily verified that if a point moves to infinity along any line parallel to the one just considered, its homogeneous coördinates may be made to approach the same values $(x, y, 0)$ as those just obtained. It is therefore natural to speak of the point at infinity in a certain direction rather than on a definite line. Finally we will agree that two points at infinity whose coördinates are proportional shall be regarded as coinciding, since these coördinates may be regarded as the limits of the coördinates of one and the same point which moves further and further off.*

If $x=y=t=0$, we will not say that we have a point at all, since the coördinates of *any point whatever* may be taken as small as we please, and so (0, 0, 0) might be regarded as the limits of the coördinates of any fixed or variable point.

* It should be noticed that in speaking of points at infinity we are, considering the matter from a purely logical point of view, doing exactly the same thing that we did in § 3 in speaking of imaginary points, or points in space of n dimensions; that is, we are speaking of a set of quantities as a "point" which are not the coördinates of any point. The only difference between the two cases is that the coördinates of our "point at infinity" are the limits of the coördinates of a true point.

Thus, in particular, it is a pure convention, though a desirable and convenient one, when we say that two points at infinity shall be regarded as coincident when and only when their coördinates are proportional. We might, if we chose, regard *all* points at infinity as coincident. There is no logical compulsion in the matter.

The equation

$$AX^2+BXY+CY^2+DX+EY+F=0$$

becomes, in homogeneous coördinates,

$$A\frac{x^2}{t^2}+B\frac{xy}{t^2}+C\frac{y^2}{t^2}+D\frac{x}{t}+E\frac{y}{t}+F=0,$$

or

$$Ax^2+Bxy+Cy^2+Dxt+Eyt+Ft^2=0,$$

a homogeneous equation of the second degree; and it is evident that if the coördinates X, Y in any algebraic equation be replaced by the coördinates x, y, t, the resulting equation will be homogeneous, and of the same degree as the original equation. It is to this fact that the system owes its name, as well as one of its chief advantages.

The equation

$$Ax+By+Ct=0$$

represents, in general, a line, but if $A=B=0$, $C\neq 0$, it has no true geometric locus. It is, in this case, satisfied by the coördinates of all points at infinity, and by the coördinates of no other point. We shall therefore speak of it as the equation of *the line at infinity*. The reader may easily verify, by using the equation of a line in terms of its intercepts, that if a straight line move further and further away, its homogeneous equation will approach more and more nearly the form $t=0$.

In space of three dimensions we will represent the point whose cartesian coördinates are X, Y, Z by the four homogeneous coördinates x, y, z, t, whose ratios are defined by the equations

$$\frac{x}{t}=X,\quad \frac{y}{t}=Y,\quad \frac{z}{t}=Z.$$

We will speak of $(x, y, z, 0)$ as "the point at infinity" on a line whose direction cosines are

$$\frac{x}{\sqrt{x^2+y^2+z^2}},\quad \frac{y}{\sqrt{x^2+y^2+z^2}},\quad \frac{z}{\sqrt{x^2+y^2+z^2}}.$$

(0, 0, 0, 0) will be excluded, and $t=0$ will be spoken of as the equation of *the plane at infinity*.

Extending the same terminology to the general case, we shall sometimes find it convenient to speak of $(x_1, x_2, \cdots x_n)$ not as a point in space of n dimensions, but as a point represented by its homogeneous coördinates in space of $n-1$ dimensions. Two points whose coördinates are proportional will be spoken of as identical, a point whose last coördinate is zero will be spoken of as a point at infinity, and the case $x_1 = \cdots = x_n = 0$ will not be spoken of as a point at all. This terminology will be adopted only in connection with homogeneous polynomials, and even then it must be clearly understood that we are perfectly free to adopt whichever terminology we find most convenient. Thus, for instance, if $f(x_1, x_2, x_3)$ is a homogeneous polynomial of the second degree, the equation $f = 0$ may be regarded either as determining a conic in a plane (x_1, x_2, x_3 being homogeneous coördinates) or a quadric cone in space (x_1, x_2, x_3 being ordinary cartesian coördinates).

Homogeneous coördinates may also be used in space of one dimension. We should then determine the points on a line by two coördinates x, t whose ratio x/t is the non-homogeneous coördinate X, *i.e.* the distance of the point from the origin. It is this representation that is commonly made use of in connection with the theory of binary forms.

5. The Continuity of Polynomials.

DEFINITION. *A function $f(x_1, \cdots x_n)$ is said to be continuous at the point $(c_1, \cdots c_n)$ if, no matter how small a positive quantity ϵ be chosen, a neighborhood of the point $(c_1, \cdots c_n)$ can be found so small that the difference between the value of the function at any point of this neighborhood and its value at the point $(c_1, \cdots c_n)$ is in absolute value less than ϵ.*

That is, f is continuous at $(c_1, \cdots c_n)$ if, having chosen a positive quantity ϵ, it is possible to determine a positive δ such that

$$|f(x_1, \cdots x_n) - f(c_1, \cdots c_n)| < \epsilon$$

for all values of $(x_1 \cdots x_n)$ which satisfy the inequalities,

$$|x_1 - c_1| < \delta,\ |x_2 - c_2| < \delta,\ \cdots\ |x_n - c_n| < \delta.$$

THEOREM 1. *If two functions are continuous at a point, their sum is continuous at this point.*

Let f_1 and f_2 be two functions continuous at the point $(c_1, \cdots c_n)$ and let k_1 and k_2 be their respective values at this point. Then, no matter how small the positive quantity ϵ may be chosen, we may take δ_1 and δ_2 so small that

$$|f_1 - k_1| < \tfrac{1}{2}\epsilon \qquad \text{when } |x_i - c_i| < \delta_1,$$
$$|f_2 - k_2| < \tfrac{1}{2}\epsilon \qquad \text{when } |x_i - c_i| < \delta_2.$$

Accordingly

$$|f_1 - k_1| + |f_2 - k_2| < \epsilon \qquad \text{when } |x_i - c_i| < \delta,$$

where δ is the smaller of the two quantities δ_1 and δ_2; and, since $|A| + |B| \geqq |A + B|$, we have

$$|f_1 - k_1 + f_2 - k_2| = |(f_1 + f_2) - (k_1 + k_2)| < \epsilon \quad \text{when } |x_i - c_i| < \delta.$$

Hence $f_1 + f_2$ is continuous at the point $(c_1, \cdots c_n)$.

COROLLARY. *If a finite number of functions are continuous at a point, their sum is continuous at this point.*

THEOREM 2. *If two functions are continuous at a point, their product is continuous at this point.*

Let f_1 and f_2 be the two functions, and k_1 and k_2 their values at the point $(c_1, \cdots c_n)$ where they are assumed to be continuous. We have to prove that however small ϵ may be, δ can be chosen so small that

(1) $$|f_1 f_2 - k_1 k_2| < \epsilon \qquad \text{when } |x_i - c_i| < \delta.$$

Let η be a positive constant, which we shall ultimately restrict to a certain degree of smallness, and let us choose two positive constants δ_1 and δ_2 such that

$$|f_1 - k_1| < \eta \qquad \text{when } |x_i - c_i| < \delta_1,$$
$$|f_2 - k_2| < \eta \qquad \text{when } |x_i - c_i| < \delta_2.$$

Now take δ as the smaller of the two quantities δ_1 and δ_2. Then, when $|x_i - c_i| < \delta$,

$$\begin{aligned}|f_1 f_2 - k_1 k_2| &= |f_2(f_1 - k_1) + k_1(f_2 - k_2)| \\ &\leqq |f_2|\,|f_1 - k_1| + |k_1|\,|f_2 - k_2| \leqq \{|f_2| + |k_1|\}\,\eta.\end{aligned}$$

Accordingly since, when $|x_i - c_i| < \delta$,

$$|f_2| = |k_2 + (f_2 - k_2)| \leqq |k_2| + |f_2 - k_2| < |k_2| + \eta,$$

we may write

(2) $$|f_1 f_2 - k_1 k_2| < \{|k_1| + |k_2|\}\,\eta + \eta^2.$$

If k_1 and k_2 are not both zero, let us take η small enough to satisfy the two inequalities

$$\eta < \frac{\epsilon}{2\{|k_1| + |k_2|\}}, \qquad \eta < \sqrt{\frac{\epsilon}{2}}.$$

If $k_1 = k_2 = 0$, we will restrict η merely by the inequality

$$\eta < \sqrt{\epsilon}.$$

In either case, inequality (2) then reduces to the form (1) and our theorem is proved.

COROLLARY. *If a finite number of functions are continuous at a point, their product is continuous at this point.*

Referring now to our definition of continuity, we see that any constant may be regarded as a continuous function of $(x_1, \cdots x_n)$ for all values of these variables, and that the same is true of any one of these variables themselves. Hence by the last corollary any function of the form $Cx_1^{k_1} \cdots x_n^{k_n}$, where the k's are integers positive or zero, is continuous at every point. If we now refer to the corollary to Theorem 1, we arrive at the theorem:

THEOREM 3. *Any polynomial is a continuous function for all values of the variables.*

Finally, we give a simple application of this theorem.

THEOREM 4. *If $f(x_1, \cdots x_n)$ is a polynomial and $f(c_1, \cdots c_n) \neq 0$, it is possible to take a neighborhood of the point $(c_1, \cdots c_n)$ so small that f does not vanish at any point in this neighborhood.*

Let $k = f(c_1, \cdots c_n)$. Then, on account of the continuity of f at $(c_1, \cdots c_n)$, a positive quantity δ can be chosen so small that throughout the neighborhood $|x_i - c_i| < \delta$, the inequality

$$|f - k| < \tfrac{1}{2}|k|$$

is satisfied. In this neighborhood f cannot vanish; for at any point where it vanished we should have

$$|f - k| = |k| < \tfrac{1}{2}|k|,$$

which is impossible since by hypothesis $k \neq 0$.

6. The Fundamental Theorem of Algebra. Up to this point no use has been made of what is often known as the fundamental theorem of algebra, namely the proposition that every algebraic equation has a root. This fact we may state in more precise form as follows:

THEOREM 1. *If $f(x)$ is a polynomial of the nth degree where $n \geq 1$, there exists at least one value of x for which $f(x) = 0$.*

This theorem, fundamental though it is, is not necessary for most of the developments in this book. Moreover, the methods of proving the theorem are essentially not algebraic, or only in part algebraic. Accordingly, we will give no proof of the theorem here, but merely refer the reader who desires a formal proof to any of the text-books on the theory of functions of a complex variable. We shall, however, when we find it convenient to do so, assume the truth of this theorem. In this section we will deduce a few of its more immediate consequences.

THEOREM 2. *If $f(x)$ is a polynomial of the nth degree,*

$$f(x) \equiv a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n \qquad (a_0 \neq 0),$$

there exists one and only one set of constants, $\alpha_1, \alpha_2, \cdots \alpha_n$, such that

$$f(x) \equiv a_0(x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n).$$

This theorem is seen at once to be true for polynomials of the first degree. Let us then use the method of mathematical induction and assume the proposition true for all polynomials of degree less than n. If we can infer that the theorem is true for polynomials of the nth degree, it follows that being true for those of the first degree it is true for those of the second, hence for those of the third, etc.

By Theorem 1 we see that there is at least one value of x for which $f(x) = 0$. Call such a value α_1. By Theorem 1, § 1 we may write

$$f(x) \equiv (x - \alpha_1)\phi(x),$$

where

$$\phi(x) \equiv a_0x^{n-1} + b_1x^{n-2} + \cdots + b_{n-1}.$$

Since $\phi(x)$ is a polynomial of degree $n-1$, and since we are assuming our theorem to be true for all such polynomials, there exist $n-1$ constants $\alpha_2, \cdots \alpha_n$ such that

$$\phi(x) \equiv a_0(x - \alpha_2) \cdots (x - \alpha_n).$$

Hence

$$f(x) \equiv a_0(x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n).$$

Thus half of our theorem is proved.

Suppose now there were two such sets of constants, $\alpha_1, \cdots \alpha_n$ and $\beta_1, \cdots \beta_n$. We should then have

$$f(x) \equiv a_0(x-\alpha_1)\cdots(x-\alpha_n) \equiv a_0(x-\beta_1)\cdots(x-\beta_n). \tag{1}$$

Let $x = \alpha_1$ in this identity. This gives

$$a_0(\alpha_1-\beta_1)(\alpha_1-\beta_2)\cdots(\alpha_1-\beta_n)=0.$$

Accordingly, since $a_0 \neq 0$, α_1 must be equal to one of the quantities $\beta_1, \beta_2, \cdots \beta_n$. Let us suppose the β's to have been taken in such an order that $\alpha_1 = \beta_1$. Now in the identity (1) cancel out the factor $a_0(x-\alpha_1)$ (see Theorem 4, § 2). This gives

$$(x-\alpha_2)\cdots(x-\alpha_n) \equiv (x-\beta_2)\cdots(x-\beta_n).$$

Accordingly, since we have assumed the theorem we are proving to be true for polynomials of degree $n-1$, the constants $\beta_2, \cdots \beta_n$ are the same, except perhaps for the order, as the constants $\alpha_2, \cdots \alpha_n$, and our theorem is proved.

Definition. *The constants $\alpha_1, \cdots \alpha_n$ determined in the last theorem are called the roots of the polynomial $f(x)$, or of the equation $f(x)=0$. If k of these roots are equal to one another, but different from all the other roots, this root is called a k-fold root.*

It is at once seen by reference to Theorem 1, § 1 that these roots are the only points at which $f(x)$ vanishes.

Theorem 3. *If $f(x_1, \cdots x_n)$ is a polynomial which is not identically equal to a constant, there are an infinite number of points $(x_1, \cdots x_n)$ at which $f \neq 0$, and also an infinite number at which $f=0$, provided $n>1$.*

The truth of the first part of this theorem is at once obvious, for, since f is not identically zero, a point can be found at which it is not zero, and then a neighborhood of this point can be taken so small that f does not vanish in this neighborhood (Theorem 4, § 5). This neighborhood, of course, consists of an infinite number of points.

To prove that f vanishes at an infinite number of points, let us select one of the variables which enters into f to at least the first degree. Without loss of generality we may suppose this variable to be x_1. We may then write

$$f(x_1, \cdots x_n) \equiv F_0(x_2, \cdots x_n)x_1^k + F_1(x_2, \cdots x_n)x_1^{k-1} + \cdots + F_k(x_2, \cdots x_n),$$

where $k \geqq 1$ and F_0 is not identically zero. Let $(c_2, \cdots c_n)$ be any point at which F_0 is not zero. Then $f(x_1, c_2, \cdots c_n)$ is a polynomial of the kth degree in x_1 alone. Accordingly, by Theorem 1, there is at least one value of x_1 for which it vanishes. If c_1 is such a value, $f(c_1, c_2, \cdots c_n)=0$. Moreover, by the part of our theorem already proved, there are an infinite number of points where $F_0 \neq 0$, that is an infinite number of choices possible for the quantities $c_2, \cdots c_n$. Thus our theorem is completely proved.

Finally, we will state, without proof, for future reference, a theorem which says, in brief, that the roots of an algebraic equation are continuous functions of the coefficients:

Theorem 4. *If α is a root of the polynomial*

$$a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n \qquad \begin{pmatrix} a_0 \neq 0 \\ n > 0 \end{pmatrix}^*,$$

then no matter how small a neighborhood $|x-\alpha| < \epsilon$ of the point α we may consider, it is possible to take in space of $n+1$ dimensions a neighborhood of the point $(a_0, a_1, \cdots a_n)$ so small that, if $(b_0, b_1, \cdots b_n)$ is any point in this neighborhood, the polynomial

$$b_0x^n + b_1x^{n-1} + \cdots + b_{n-1}x + b_n$$

has at least one root β in the neighborhood $|x-\alpha| < \epsilon$ of the point α.

For a proof of this theorem we refer to Weber's *Algebra*, Vol. 1, § 44.

* The theorem remains true if we merely assume that the polynomial is of at least the first degree. That is, some of the first coefficients $a_0, a_1, \cdots$ may be zero.