CHAPTER IV

Polynomials

This chapter provides a continuation of Chapter II, §3. We prove standard properties of polynomials. Most readers will be acquainted with some of these properties, especially at the beginning for polynomials in one variable. However, one of our purposes is to show that some of these properties also hold over a commutative ring when properly formulated. The Gauss lemma and the reduction criterion for irreducibility will show the importance of working over rings. Chapter IX will give examples of the importance of working over the integers **Z** themselves to get universal relations. It happens that certain statements of algebra are universally true. To prove them, one proves them first for elements of a polynomial ring over **Z**, and then one obtains the statement in arbitrary fields (or commutative rings as the case may be) by specialization. The Cayley–Hamilton theorem of Chapter XV, for instance, can be proved in that way.

The last section on power series shows that the basic properties of polynomial rings can be formulated so as to hold for power series rings. I conclude this section with several examples showing the importance of power series in various parts of mathematics.

§1. BASIC PROPERTIES FOR POLYNOMIALS IN ONE VARIABLE

We start with the Euclidean algorithm.

Theorem 1.1. *Let A be a commutative ring, let $f, g \in A[X]$ be polynomials in one variable, of degrees $\geqq 0$, and assume that the leading*

coefficient of g is a unit in A. Then there exist unique polynomials $q, r \in A[X]$ such that

$$f = gq + r$$

and $\deg r < \deg g$.

Proof. Write

$$f(X) = a_n X^n + \cdots + a_0,$$
$$g(X) = b_d X^d + \cdots + b_0,$$

where $n = \deg f$, $d = \deg g$ so that a_n, $b_d \neq 0$ and b_d is a unit in A. We use induction on n.

If $n = 0$, and $\deg g > \deg f$, we let $q = 0$, $r = f$. If $\deg g = \deg f = 0$, then we let $r = 0$ and $q = a_n b_d^{-1}$.

Assume the theorem proved for polynomials of degree $< n$ (with $n > 0$). We may assume $\deg g \leqq \deg f$ (otherwise, take $q = 0$ and $r = f$). Then

$$f(X) = a_n b_d^{-1} X^{n-d} g(X) + f_1(X),$$

where $f_1(X)$ has degree $< n$. By induction, we can find q_1, r such that

$$f(X) = a_n b_d^{-1} X^{n-d} g(X) + q_1(X) g(X) + r(X)$$

and $\deg r < \deg g$. Then we let

$$q(X) = a_n b_d^{-1} X^{n-d} + q_1(X)$$

to conclude the proof of existence for q, r.

As for uniqueness, suppose that

$$f = q_1 g + r_1 = q_2 g + r_2$$

with $\deg r_1 < \deg g$ and $\deg r_2 < \deg g$. Subtracting yields

$$(q_1 - q_2) g = r_2 - r_1.$$

Since the leading coefficient of g is assumed to be a unit, we have

$$\deg(q_1 - q_2) g = \deg(q_1 - q_2) + \deg g.$$

Since $\deg(r_2 - r_1) < \deg g$, this relation can hold only if $q_1 - q_2 = 0$, i.e. $q_1 = q_2$, and hence finally $r_1 = r_2$ as was to be shown.

Theorem 1.2. *Let k be a field. Then the polynomial ring in one variable $k[X]$ is principal.*

Proof. Let $\mathfrak{a}$ be an ideal of $k[X]$, and assume $\mathfrak{a} \neq 0$. Let g be an element of $\mathfrak{a}$ of smallest degree $\geqq 0$. Let f be any element of $\mathfrak{a}$ such that $f \neq 0$. By the Euclidean algorithm we can find $q, r \in k[X]$ such that

$$f = qg + r$$

and $\deg r < \deg g$. But $r = f - qg$, whence r is in $\mathfrak{a}$. Since g had minimal degree $\geqq 0$ it follows that $r = 0$, hence that $\mathfrak{a}$ consists of all polynomials qg (with $q \in k[X]$). This proves our theorem. By Theorem 5.2 of Chapter II we get:

Corollary 1.3. *The ring $k[X]$ is factorial.*

If k is a field then every non-zero element of k is a unit in k, and one sees immediately that the units of $k[X]$ are simply the units of k. (No polynomial of degree $\geqq 1$ can be a unit because of the addition formula for the degree of a product.)

A polynomial $f(X) \in k[X]$ is called **irreducible** if it has degree $\geqq 1$, and if one cannot write $f(X)$ as a product

$$f(X) = g(X)h(X)$$

with $g, h \in k[X]$, and both $g, h \notin k$. Elements of k are usually called **constant polynomials**, so we can also say that in such a factorization, one of g or h must be constant. A polynomial is called **monic** if it has leading coefficient 1.

Let A be a commutative ring and $f(X)$ a polynomial in $A[X]$. Let A be a subring of B. An element $b \in B$ is called a **root** or a **zero** of f in B if $f(b) = 0$. Similarly, if (X) is an n-tuple of variables, an n-tuple (b) is called a zero of f if $f(b) = 0$.

Theorem 1.4. *Let k be a field and f a polynomial in one variable X in $k[X]$, of degree $n \geqq 0$. Then f has at most n roots in k, and if a is a root of f in k, then $X - a$ divides $f(X)$.*

Proof. Suppose $f(a) = 0$. Find q, r such that

$$f(X) = q(X)(X - a) + r(X)$$

and $\deg r < 1$. Then

$$0 = f(a) = r(a).$$

Since $r = 0$ or r is a non-zero constant, we must have $r = 0$, whence $X - a$ divides $f(X)$. If $a_1, \ldots, a_m$ are distinct roots of f in k, then inductively we see that the product

$$(X - a_1) \cdots (X - a_m)$$

divides $f(X)$, whence $m \leqq n$, thereby proving the theorem. The next corollaries give applications of Theorem 1.4 to polynomial functions.

Corollary 1.5. *Let k be a field and T an infinite subset of k. Let $f(X) \in k[X]$ be a polynomial in one variable. If $f(a) = 0$ for all $a \in T$, then $f = 0$, i.e. f induces the zero function.*

Corollary 1.6. *Let k be a field, and let $S_1, \ldots, S_n$ be infinite subsets of k. Let $f(X_1, \ldots, X_n)$ be a polynomial in n variables over k. If $f(a_1, \ldots, a_n) = 0$ for all $a_i \in S_i$ $(i = 1, \ldots, n)$, then $f = 0$.*

Proof. By induction. We have just seen the result is true for one variable. Let $n \geqq 2$, and write

$$f(X_1, \ldots, X_n) = \sum_j f_j(X_1, \ldots, X_{n-1}) X_n^j$$

as a polynomial in X_n with coefficients in $k[X_1, \ldots, X_{n-1}]$. If there exists

$$(b_1, \ldots, b_{n-1}) \in S_1 \times \cdots \times S_{n-1}$$

such that for some j we have $f_j(b_1, \ldots, b_{n-1}) \neq 0$, then

$$f(b_1, \ldots, b_{n-1}, X_n)$$

is a non-zero polynomial in $k[X_n]$ which takes on the value 0 for the infinite set of elements S_n. This is impossible. Hence f_j induces the zero function on $S_1 \times \cdots \times S_{n-1}$ for all j, and by induction we have $f_j = 0$ for all j. Hence $f = 0$, as was to be shown.

Corollary 1.7. *Let k be an infinite field and f a polynomial in n variables over k. If f induces the zero function on $k^{(n)}$, then $f = 0$.*

We shall now consider the case of finite fields. Let k be a finite field with q elements. Let $f(X_1, \ldots, X_n)$ be a polynomial in n variables over k. Write

$$f(X_1, \ldots, X_n) = \sum a_{(\nu)} X_1^{\nu_1} \cdots X_n^{\nu_n}.$$

If $a_{(\nu)} \neq 0$, we recall that the monomial $M_{(\nu)}(X)$ **occurs** in f. Suppose this is the case, and that in this monomial $M_{(\nu)}(X)$, some variable X_i occurs with an exponent $\nu_i \geqq q$. We can write

$$X_i^{\nu_i} = X_i^{q+\mu}, \qquad \mu = \text{integer} \geqq 0.$$

If we now replace $X_i^{\nu_i}$ by $X_i^{\mu+1}$ in this monomial, then we obtain a new polynomial which gives rise to the same function as f. The degree of this new polynomial is at most equal to the degree of f.

Performing the above operation a finite number of times, for all the monomials occurring in f and all the variables $X_1, \ldots, X_n$ we obtain some polynomial f^* giving rise to the same function as f, but whose degree in each variable is $< q$.

Corollary 1.8. *Let k be a finite field with q elements. Let f be a polynomial in n variables over k such that the degree of f in each variable is $< q$. If f induces the zero function on $k^{(n)}$, then $f = 0$.*

Proof. By induction. If $n = 1$, then the degree of f is $< q$, and hence f cannot have q roots unless it is 0. The inductive step is carried out just as we did for the proof of Corollary 1.6 above.

Let f be a polynomial in n variables over the finite field k. A polynomial g whose degree in each variable is $< q$ will be said to be **reduced**. We have shown above that there exists a reduced polynomial f^* which gives the same function as f on $k^{(n)}$. Theorem 1.8 now shows that *this reduced polynomial is unique.* Indeed, if g_1, g_2 are reduced polynomials giving the same function, then $g_1 - g_2$ is reduced and gives the zero function. Hence $g_1 - g_2 = 0$ and $g_1 = g_2$.

We shall give one more application of Theorem 1.4. Let k be a field. By a **multiplicative subgroup** of k we shall mean a subgroup of the group k^* (non-zero elements of k).

Theorem 1.9. *Let k be a field and let U be a finite multiplicative subgroup of k. Then U is cyclic.*

Proof. Write U as a product of subgroups $U(p)$ for each prime p, where $U(p)$ is a p-group. By Proposition 4.3(v) of Chapter I, it will suffice to prove that $U(p)$ is cyclic for each p. Let a be an element of $U(p)$ of maximal period p^r for some integer r. Then $x^{p^r} = 1$ for every element $x \in U(p)$, and hence all elements of $U(p)$ are roots of the polynomial

$$X^{p^r} - 1.$$

The cyclic group generated by a has p^r elements. If this cyclic group is not equal to $U(p)$, then our polynomial has more than p^r roots, which is impossible. Hence a generates $U(p)$, and our theorem is proved.

Corollary 1.10. *If k is a finite field, then k^* is cyclic.*

An element ζ in a field k such that there exists an integer $n \geqq 1$ such that $\zeta^n = 1$ is called a **root of unity**, or more precisely an n-th root of unity. Thus the set of n-th roots of unity is the set of roots of the polynomial $X^n - 1$. There are at most n such roots, and they obviously form a group, which is

cyclic by Theorem 1.9. We shall study roots of unity in greater detail later. A generator for the group of n-th roots of unity is called a **primitive** n-th root of unity. For example, in the complex numbers, $e^{2\pi i/n}$ is a primitive n-th root of unity, and the n-th roots of unity are of type $e^{2\pi iv/n}$ with $1 \leqq v \leqq n$.

The group of roots of unity is denoted by $\boldsymbol{\mu}$. The group of roots of unity in a field K is denoted by $\boldsymbol{\mu}(K)$.

A field k is said to be **algebraically closed** if every polynomial in $k[X]$ of degree $\geqq 1$ has a root in k. In books on analysis, it is proved that the complex numbers are algebraically closed. In Chapter V we shall prove that a field k is always contained in some algebraically closed field. If k is algebraically closed then the irreducible polynomials in $k[X]$ are the polynomials of degree 1. In such a case, the unique factorization of a polynomial f of degree $\geqq 0$ can be written in the form

$$f(X) = c \prod_{i=1}^{r} (X - \alpha_i)^{m_i}$$

with $c \in k$, $c \neq 0$ and distinct roots $\alpha_1, \ldots, \alpha_r$. We next develop a test when $m_i > 1$.

Let A be a commutative ring. We define a map

$$D\colon A[X] \to A[X]$$

of the polynomial ring into itself. If $f(X) = a_n X^n + \cdots + a_0$ with $a_i \in A$, we define the **derivative**

$$Df(X) = f'(X) = \sum_{v=1}^{n} v a_v X^{v-1} = n a_n X^{n-1} + \cdots + a_1.$$

One verifies easily that if f, g are polynomials in $A[X]$, then

$$(f + g)' = f' + g', \qquad (fg)' = f'g + fg',$$

and if $a \in A$, then

$$(af)' = af'.$$

Let K be a field and f a non-zero polynomial in $K[X]$. Let a be a root of f in K. We can write

$$f(X) = (X - a)^m g(X)$$

with some polynomial $g(X)$ relatively prime to $X - a$ (and hence such that $g(a) \neq 0$). We call m the **multiplicity** of a in f, and say that a is a **multiple root** if $m > 1$.

Proposition 1.11. *Let K, f be as above. The element a of K is a multiple root of f if and only if it is a root and $f'(a) = 0$.*

Proof. Factoring f as above, we get

$$f'(X) = (X - a)^m g'(X) + m(X - a)^{m-1} g(X).$$

If $m > 1$, then obviously $f'(a) = 0$. Conversely, if $m = 1$ then

$$f'(X) = (X - a)g'(X) + g(X),$$

whence $f'(a) = g(a) \neq 0$. Hence if $f'(a) = 0$ we must have $m > 1$, as desired.

Proposition 1.12. *Let $f \in K[X]$. If K has characteristic 0, and f has degree $\geqq 1$, then $f' \neq 0$. Let K have characteristic $p > 0$ and f have degree $\geqq 1$. Then $f' = 0$ if and only if, in the expression for $f(X)$ given by*

$$f(X) = \sum_{v=0}^{n} a_v X^v,$$

p divides each integer v such that $a_v \neq 0$.

Proof. If K has characteristic 0, then the derivative of a monomial $a_v X^v$ such that $v \geqq 1$ and $a_v \neq 0$ is not zero since it is $v a_v X^{v-1}$. If K has characteristic $p > 0$, then the derivative of such a monomial is 0 if and only if $p | v$, as contended.

Let K have characteristic $p > 0$, and let f be written as above, and be such that $f'(X) = 0$. Then one can write

$$f(X) = \sum_{\mu=0}^{d} b_\mu X^{p\mu}$$

with $b_\mu \in K$.

Since the binomial coefficients $\binom{p}{v}$ are divisible by p for $1 \leqq v \leqq p - 1$ we see that if K has characteristic p, then for $a, b \in K$ we have

$$(a + b)^p = a^p + b^p.$$

Since obviously $(ab)^p = a^p b^p$, the map

$$x \mapsto x^p$$

is a homomorphism of K into itself, which has trivial kernel, hence is injective. Iterating, we conclude that for each integer $r \geqq 1$, the map $x \mapsto x^{p^r}$

is an endomorphism of K, called the **Frobenius endomorphism**. Inductively, if $c_1, \ldots, c_n$ are elements of K, then

$$(c_1 + \cdots + c_n)^p = c_1^p + \cdots + c_n^p.$$

Applying these remarks to polynomials, we see that for any element $a \in K$ we have

$$(X - a)^{p^r} = X^{p^r} - a^{p^r}.$$

If $c \in K$ and the polynomial

$$X^{p^r} - c$$

has one root a in K, then $a^{p^r} = c$ and

$$X^{p^r} - c = (X - a)^{p^r}.$$

Hence our polynomial has precisely one root, of multiplicity p^r. For instance, $(X - 1)^{p^r} = X^{p^r} - 1$.

§2. POLYNOMIALS OVER A FACTORIAL RING

Let A be a factorial ring, and K its quotient field. Let $a \in K$, $a \neq 0$. We can write a as a quotient of elements in A, having no prime factor in common. If p is a prime element of A, then we can write

$$a = p^r b,$$

where $b \in K$, r is an integer, and p does not divide the numerator or denominator of b. Using the unique factorization in A, we see at once that r is uniquely determined by a, and we call r the **order of a at p** (and write $r = \operatorname{ord}_p a$). If $a = 0$, we define its order at p to be ∞.

If $a, a' \in K$ and $aa' \neq 0$, then

$$\operatorname{ord}_p(aa') = \operatorname{ord}_p a + \operatorname{ord}_p a'.$$

This is obvious.

Let $f(X) \in K[X]$ be a polynomial in one variable, written

$$f(X) = a_0 + a_1 X + \cdots + a_n X^n.$$

If $f = 0$, we define $\operatorname{ord}_p f$ to be ∞. If $f \neq 0$, we define $\operatorname{ord}_p f$ to be

$$\operatorname{ord}_p f = \min \operatorname{ord}_p a_i,$$

the minimum being taken over all those i such that $a_i \neq 0$.

If $r = \operatorname{ord}_p f$, we call up^r a **p-content** for f, if u is any unit of A. We define the **content** of f to be the product.

$$\prod p^{\operatorname{ord}_p f},$$

the product being taken over all p such that $\operatorname{ord}_p f \neq 0$, or any multiple of this product by a unit of A. Thus the content is well defined up to multiplication by a unit of A. We abbreviate **content** by **cont**.

If $b \in K$, $b \neq 0$, then $\operatorname{cont}(bf) = b \operatorname{cont}(f)$. This is clear. Hence we can write

$$f(X) = c \cdot f_1(X)$$

where $c = \operatorname{cont}(f)$, and $f_1(X)$ has content 1. In particular, all coefficients of f_1 lie in A, and their g.c.d. is 1. We define a polynomial with content 1 to be a **primitive polynomial**.

Theorem 2.1. (Gauss Lemma). *Let A be a factorial ring, and let K be its quotient field. Let $f, g \in K[X]$ be polynomials in one variable. Then*

$$\operatorname{cont}(fg) = \operatorname{cont}(f) \operatorname{cont}(g).$$

Proof. Writing $f = cf_1$ and $g = dg_1$ where $c = \operatorname{cont}(f)$ and $d = \operatorname{cont}(g)$, we see that it suffices to prove: If f, g have content 1, then fg also has content 1, and for this, it suffices to prove that for each prime p, $\operatorname{ord}_p(fg) = 0$. Let

$$f(X) = a_n X^n + \cdots + a_0, \qquad a_n \neq 0,$$

$$g(X) = b_m X^m + \cdots + b_0, \qquad b_m \neq 0,$$

be polynomials of content 1. Let p be a prime of A. It will suffice to prove that p does not divide all coefficients of fg. Let r be the largest integer such that $0 \leqq r \leqq n$, $a_r \neq 0$, and p does not divide a_r. Similarly, let b_s be the coefficient of g farthest to the left, $b_s \neq 0$, such that p does not divide b_s. Consider the coefficient of X^{r+s} in $f(X)g(X)$. This coefficient is equal to

$$c = a_r b_s + a_{r+1} b_{s-1} + \cdots + a_{r-1} b_{s+1} + \cdots$$

and $p \nmid a_r b_s$. However, p divides every other non-zero term in this sum since in each term there will be some coefficient a_i to the left of a_r or some coefficient b_j to the left of b_s. Hence p does not divide c, and our lemma is proved.

We shall now give another proof for the key step in the above argument, namely the statement:

If $f, g \in A[X]$ are primitive (i.e. have content 1) then fg is primitive.

Proof. We have to prove that a given prime p does not divide all the coefficients of fg. Consider reduction mod p, namely the canonical homomorphism $A \to A/(p) = \bar{A}$. Denote the image of a polynomial by a bar, so $f \mapsto \bar{f}$ and $g \mapsto \bar{g}$ under the reduction homomorphism. Then

$$\overline{fg} = \bar{f}\bar{g}.$$

By hypothesis, $\bar{f} \neq 0$ and $\bar{g} \neq 0$. Since $\bar{A}$ is entire, it follows that $\overline{fg} \neq 0$, as was to be shown.

Corollary 2.2. *Let $f(X) \in A[X]$ have a factorization $f(X) = g(X)h(X)$ in $K[X]$. If $c_g = \operatorname{cont}(g)$, $c_h = \operatorname{cont}(h)$, and $g = c_g g_1$, $h = c_h h_1$, then*

$$f(X) = c_g c_h g_1(X) h_1(X),$$

and $c_g c_h$ is an element of A. In particular, if $f, g \in A[X]$ have content 1, then $h \in A[X]$ also.

Proof. The only thing to be proved is $c_g c_h \in A$. But

$$\operatorname{cont}(f) = c_g c_h \operatorname{cont}(g_1 h_1) = c_g c_h,$$

whence our assertion follows.

Theorem 2.3. *Let A be a factorial ring. Then the polynomial ring $A[X]$ in one variable is factorial. Its prime elements are the primes of A and polynomials in $A[X]$ which are irreducible in $K[X]$ and have content 1.*

Proof. Let $f \in A[X]$, $f \neq 0$. Using the unique factorization in $K[X]$ and the preceding corollary, we can find a factorization

$$f(X) = c \cdot p_1(X) \cdots p_r(X)$$

where $c \in A$, and $p_1, \ldots, p_r$ are polynomials in $A[X]$ which are irreducible in $K[X]$. Extracting their contents, we may assume without loss of generality that the content of p_i is 1 for each i. Then $c = \operatorname{cont}(f)$ by the Gauss lemma. This gives us the existence of the factorization. It follows that each $p_i(X)$ is irreducible in $A[X]$. If we have another such factorization, say

$$f(X) = d \cdot q_1(X) \cdots q_s(X),$$

then from the unique factorization in $K[X]$ we conclude that $r = s$, and after a permutation of the factors we have

$$p_i = a_i q_i$$

with elements $a_i \in K$. Since both p_i, q_i are assumed to have content 1, it follows that a_i in fact lies in A and is a unit. This proves our theorem.

Corollary 2.4. *Let A be a factorial ring. Then the ring of polynomials in n variables $A[X_1, \ldots, X_n]$ is factorial. Its units are precisely the units of A, and its prime elements are either primes of A or polynomials which are irreducible in $K[X]$ and have content* 1.

Proof. Induction.

In view of Theorem 2.3, when we deal with polynomials over a factorial ring and having content 1, it is not necessary to specify whether such polynomials are irreducible over A or over the quotient field K. The two notions are equivalent.

Remark 1. The polynomial ring $K[X_1, \ldots, X_n]$ over a field K is not principal when $n \geqq 2$. For instance, the ideal generated by $X_1, \ldots, X_n$ is not principal (trivial proof).

Remark 2. It is usually not too easy to decide when a given polynomial (say in one variable) is irreducible. For instance, the polynomial $X^4 + 4$ is *reducible* over the rational numbers, because

$$X^4 + 4 = (X^2 - 2X + 2)(X^2 + 2X + 2).$$

Later in this book we shall give a precise criterion when a polynomial $X^n - a$ is irreducible. Other criteria are given in the next section.

§3. CRITERIA FOR IRREDUCIBILITY

The first criterion is:

Theorem 3.1. (Eisenstein's Criterion). *Let A be a factorial ring. Let K be its quotient field. Let $f(X) = a_n X^n + \cdots + a_0$ be a polynomial of degree $n \geqq 1$ in $A[X]$. Let p be a prime of A, and assume:*

$$a_n \not\equiv 0 \pmod{p}, \qquad a_i \equiv 0 \pmod{p} \qquad \text{for all} \quad i < n,$$

$$a_0 \not\equiv 0 \pmod{p^2}.$$

Then $f(X)$ is irreducible in $K[X]$.

Proof. Extracting a g.c.d. for the coefficients of f, we may assume without loss of generality that the content of f is 1. If there exists a factorization into factors of degree $\geqq 1$ in $K[X]$, then by the corollary of Gauss' lemma there exists a factorization in $A[X]$, say $f(X) = g(X)h(X)$,

$$g(X) = b_d X^d + \cdots + b_0,$$

$$h(X) = c_m X^m + \cdots + c_0,$$

with $d, m \geqq 1$ and $b_d c_m \neq 0$. Since $b_0 c_0 = a_0$ is divisible by p but not p^2, it follows that one of b_0, c_0 is not divisible by p, say b_0. Then $p | c_0$. Since $c_m b_d = a_n$ is not divisible by p, it follows that p does not divide c_m. Let c_r be the coefficient of h furthest to the right such that $c_r \not\equiv 0 \pmod{p}$. Then

$$a_r = b_0 c_r + b_1 c_{r-1} + \cdots .$$

Since $p \nmid b_0 c_r$ but p divides every other term in this sum, we conclude that $p \nmid a_r$, a contradiction which proves our theorem.

Example. Let a be a non-zero square-free integer $\neq \pm 1$. Then for any integer $n \geqq 1$, the polynomial $X^n - a$ is irreducible over $\mathbf{Q}$. The polynomials $3X^5 - 15$ and $2X^{10} - 21$ are irreducible over $\mathbf{Q}$.

There are some cases in which a polynomial does not satisfy Eisenstein's criterion, but a simple transform of it does.

Example. Let p be a prime number. Then the polynomial

$$f(X) = X^{p-1} + \cdots + 1$$

is irreducible over $\mathbf{Q}$.

Proof. It will suffice to prove that the polynomial $f(X + 1)$ is irreducible over $\mathbf{Q}$. We note that the binomial coefficients

$$\binom{p}{v} = \frac{p!}{v!(p-v)!}, \qquad 1 \leqq v \leqq p - 1,$$

are divisible by p (because the numerator is divisible by p and the denominator is not, and the coefficient is an integer). We have

$$f(X+1) = \frac{(X+1)^p - 1}{(X+1) - 1} = \frac{X^p + pX^{p-1} + \cdots + pX}{X}$$

from which one sees that $f(X + 1)$ satisfies Eisenstein's criterion.

Example. Let E be a field and t an element of some field containing E such that t is transcendental over E. Let K be the quotient field of $E[t]$.

For any integer $n \geqq 1$ the polynomial $X^n - t$ is irreducible in $K[X]$. This comes from the fact that the ring $A = E[t]$ is factorial and that t is a prime in it.

Theorem 3.2. (Reduction Criterion). *Let A, B be entire rings, and let*

$$\varphi: A \to B$$

be a homomorphism. Let K, L be the quotient fields of A and B respectively. Let $f \in A[X]$ be such that $\varphi f \neq 0$ and $\deg \varphi f = \deg f$. If φf is irreducible in $L[X]$, then f does not have a factorization $f(X) = g(X)h(X)$ with

$$g, h \in A[X] \qquad \text{and} \qquad \deg g, \deg h \geqq 1.$$

Proof. Suppose f has such a factorization. Then $\varphi f = (\varphi g)(\varphi h)$. Since $\deg \varphi g \leqq \deg g$ and $\deg \varphi h \leqq \deg h$, our hypothesis implies that we must have equality in these degree relations. Hence from the irreducibility in $L[X]$ we conclude that g or h is an element of A, as desired.

In the preceding criterion, suppose that A is a local ring, i.e. a ring having a unique maximal ideal $\mathfrak{p}$, and that $\mathfrak{p}$ is the kernel of φ. Then from the irreducibility of φf in $L[X]$ we conclude the irreducibility of f in $A[X]$. Indeed, any element of A which does not lie in $\mathfrak{p}$ must be a unit in A, so our last conclusion in the proof can be strengthened to the statement that g or h is a unit in A.

One can also apply the criterion when A is factorial, and in that case deduce the irreducibility of f in $K[X]$.

Example. Let p be a prime number. It will be shown later that $X^p - X - 1$ is irreducible over the field $\mathbf{Z}/p\mathbf{Z}$. Hence $X^p - X - 1$ is irreducible over $\mathbf{Q}$. Similarly,

$$X^5 - 5X^4 - 6X - 1$$

is irreducible over $\mathbf{Q}$.

There is also a routine elementary school test whether a polynomial has a root or not.

Proposition 3.3. (Integral Root Test). *Let A be a factorial ring and K its quotient field. Let*

$$f(X) = a_n X^n + \cdots + a_0 \in A[X].$$

Let $\alpha \in K$ be a root of f, with $\alpha = b/d$ expressed with b, $d \in A$ and b, d relatively prime. Then $b | a_0$ and $d | a_n$. In particular, if the leading coefficient a_n is 1, then a root α must lie in A and divides a_0.

We leave the proof to the reader, who should be used to this one from way back. As an irreducibility test, the test is useful especially for a polynomial of degree 2 or 3, when reducibility is equivalent with the existence of a root in the given field.

§4. HILBERT'S THEOREM

This section proves a basic theorem of Hilbert concerning the ideals of a polynomial ring. We define a commutative ring A to be **Noetherian** if every ideal is finitely generated.

Theorem 4.1. *Let A be a commutative Noetherian ring. Then the polynomial ring $A[X]$ is also Noetherian.*

Proof. Let $\mathfrak{A}$ be an ideal of $A[X]$. Let $\mathfrak{a}_i$ consist of 0 and the set of elements $a \in A$ appearing as leading coefficient in some polynomial

$$a_0 + a_1 X + \cdots + aX^i$$

lying in $\mathfrak{A}$. Then it is clear that $\mathfrak{a}_i$ is an ideal. (If a, b are in $\mathfrak{a}_i$, then $a \pm b$ is in $\mathfrak{a}_i$ as one sees by taking the sum and difference of the corresponding polynomials. If $x \in A$, then $xa \in \mathfrak{a}_i$ as one sees by multiplying the corresponding polynomial by x.) Furthermore we have

$$\mathfrak{a}_0 \subset \mathfrak{a}_1 \subset \mathfrak{a}_2 \subset \cdots,$$

in other words, our sequence of ideals $\{\mathfrak{a}_i\}$ is increasing. Indeed, to see this multiply the above polynomial by X to see that $a \in \mathfrak{a}_{i+1}$.

By criterion (2) of Chapter X, §1, the sequence of ideals $\{\mathfrak{a}_i\}$ stops, say at $\mathfrak{a}_r$:

$$\mathfrak{a}_0 \subset \mathfrak{a}_1 \subset \mathfrak{a}_2 \subset \cdots \subset \mathfrak{a}_r = \mathfrak{a}_{r+1} = \cdots.$$

Let

$$a_{01}, \ldots, a_{0n_0} \quad \text{be generators for} \quad \mathfrak{a}_0,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$a_{r1}, \ldots, a_{rn_r} \quad \text{be generators for} \quad \mathfrak{a}_r.$$

For each $i = 0, \ldots, r$ and $j = 1, \ldots, n_i$ let f_{ij} be a polynomial in $\mathfrak{A}$, of degree i, with leading coefficient a_{ij}. We contend that the polynomials f_{ij} are a set of generators for $\mathfrak{A}$.

Let f be a polynomial of degree d in $\mathfrak{A}$. We shall prove that f is in the ideal generated by the f_{ij}, by induction on d. Say $d \geqq 0$. If $d > r$, then we

note that the leading coefficients of

$$X^{d-r}f_{r1}, \ldots, X^{d-r}f_{rn_r}$$

generate $\mathfrak{a}_d$. Hence there exist elements $c_1, \ldots, c_{n_r} \in A$ such that the polynomial

$$f - c_1X^{d-r}f_{r1} - \cdots - c_{n_r}X^{d-r}f_{rn_r}$$

has degree $< d$, and this polynomial also lies in $\mathfrak{A}$. If $d \leqq r$, we can subtract a linear combination

$$f - c_1f_{d1} - \cdots - c_{n_d}f_{dn_d}$$

to get a polynomial of degree $< d$, also lying in $\mathfrak{A}$. We note that the polynomial we have subtracted from f lies in the ideal generated by the f_{ij}. By induction, we can subtract a polynomial g in the ideal generated by the f_{ij} such that $f - g = 0$, thereby proving our theorem.

We note that if $\varphi: A \to B$ is a surjective homomorphism of commutative rings and A is Noetherian, so is B. Indeed, let $\mathfrak{b}$ be an ideal of B, so $\varphi^{-1}(\mathfrak{b})$ is an ideal of A. Then there is a finite number of generators $(a_1, \ldots, a_n)$ for $\varphi^{-1}(\mathfrak{b})$, and it follows since φ is surjective that $\mathfrak{b} = \varphi(\varphi^{-1}(\mathfrak{b}))$ is generated by $\varphi(a_1), \ldots, \varphi(a_n)$, as desired. As an application, we obtain:

Corollary 4.2. *Let A be a Noetherian commutative ring, and let $B = A[x_1, \ldots, x_m]$ be a commutative ring finitely generated over A. Then B is Noetherian.*

Proof. Use Theorem 4.1 and the preceding remark, representing B as a factor ring of a polynomial ring.

Ideals in polynomial rings will be studied more deeply in Chapter IX. The theory of Noetherian rings and modules will be developed in Chapter X.

§5. PARTIAL FRACTIONS

In this section, we analyze the quotient field of a principal ring, using the factoriality of the ring.

Theorem 5.1. *Let A be a principal entire ring, and let P be a set of representatives for its irreducible elements. Let K be the quotient field of A, and let $\alpha \in K$. For each $p \in P$ there exists an element $\alpha_p \in A$ and an integer $j(p) \geqq 0$, such that $j(p) = 0$ for almost all $p \in P$, α_p and $p^{j(p)}$ are*

relatively prime, and

$$\alpha = \sum_{p \in P} \frac{\alpha_p}{p^{j(p)}}.$$

If we have another such expression

$$\alpha = \sum_{p \in P} \frac{\beta_p}{p^{i(p)}},$$

then $j(p) = i(p)$ *for all* p, *and* $\alpha_p \equiv \beta_p \bmod p^{j(p)}$ *for all* p.

Proof. We first prove existence, in a special case. Let a, b be relatively prime non-zero elements of A. Then there exists x, $y \in A$ such that $xa + yb = 1$. Hence

$$\frac{1}{ab} = \frac{x}{b} + \frac{y}{a}.$$

Hence any fraction c/ab with $c \in A$ can be decomposed into a sum of two fractions (namely cx/b and cy/a) whose denominators divide b and a respectively. By induction, it now follows that any $\alpha \in K$ has an expression as stated in the theorem, except possibly for the fact that p may divide α_p. Canceling the greatest common divisor yields an expression satisfying all the desired conditions.

As for uniqueness, suppose that α has two expressions as stated in the theorem. Let q be a fixed prime in P. Then

$$\frac{\alpha_q}{q^{j(q)}} - \frac{\beta_q}{q^{i(q)}} = \sum_{p \neq q} \frac{\beta_p}{p^{i(p)}} - \frac{\alpha_p}{p^{j(p)}}.$$

If $j(q) = i(q) = 0$, our conditions concerning q are satisfied. Suppose one of $j(q)$ or $i(q) > 0$, say $j(q)$, and say $j(q) \geqq i(q)$. Let d be a least common multiple for all powers $p^{j(p)}$ and $p^{i(p)}$ such that $p \neq q$. Multiply the above equation by $dq^{j(q)}$. We get

$$d(\alpha_q - q^{j(q)-i(q)}\beta_q) = q^{j(q)}\beta$$

for some $\beta \in A$. Furthermore, q does not divide d. If $i(q) < j(q)$ then q divides α_q, which is impossible. Hence $i(q) = j(q)$. We now see that $q^{j(q)}$ divides $\alpha_q - \beta_q$, thereby proving the theorem.

We apply Theorem 5.1 to the polynomial ring $k[X]$ over a field k. We let P be the set of irreducible polynomials, normalized so as to have leading coefficient equal to 1. Then P is a set of representatives for all the irreducible elements of $k[X]$. In the expression given for α in Theorem 5.1, we can now divide α_p by $p^{j(p)}$, i.e. use the Euclidean algorithm, if $\deg \alpha_p \geqq \deg p^{j(p)}$. We denote the quotient field of $k[X]$ by $k(X)$, and call its elements **rational functions**.

Theorem 5.2. *Let $A = k[X]$ be the polynomial ring in one variable over a field k. Let P be the set of irreducible polynomials in $k[X]$ with leading coefficient 1. Then any element f of $k(X)$ has a unique expression*

$$f(X) = \sum_{p \in P} \frac{f_p(X)}{p(X)^{j(p)}} + g(X),$$

where f_p, g are polynomials, $f_p = 0$ if $j(p) = 0$, f_p is relatively prime to p if $j(p) > 0$, and $\deg f_p < \deg p^{j(p)}$ if $j(p) > 0$.

Proof. The existence follows at once from our previous remarks. The uniqueness follows from the fact that if we have two expressions, with elements f_p and φ_p respectively, and polynomials g, h, then $p^{j(p)}$ divides $f_p - \varphi_p$, whence $f_p - \varphi_p = 0$, and therefore $f_p = \varphi_p$, $g = h$.

One can further decompose the term $f_p/p^{j(p)}$ by expanding f_p according to powers of p. One can in fact do something more general.

Theorem 5.3. *Let k be a field and $k[X]$ the polynomial ring in one variable. Let f, $g \in k[X]$, and assume $\deg g \geqq 1$. Then there exist unique polynomials*

$$f_0, f_1, \ldots, f_d \in k[X]$$

such that $\deg f_i < \deg g$ and such that

$$f = f_0 + f_1 g + \cdots + f_d g^d.$$

Proof. We first prove existence. If $\deg g > \deg f$, then we take $f_0 = f$ and $f_i = 0$ for $i > 0$. Suppose $\deg g \leqq \deg f$. We can find polynomials q, r with $\deg r < \deg g$ such that

$$f = qg + r,$$

and since $\deg g \geqq 1$ we have $\deg q < \deg f$. Inductively, there exist polynomials $h_0, h_1, \ldots, h_s$ such that

$$q = h_0 + h_1 g + \cdots + h_s g^s,$$

and hence

$$f = r + h_0 g + \cdots + h_s g^{s+1},$$

thereby proving existence.

As for uniqueness, let

$$f = f_0 + f_1 g + \cdots + f_d g^d = \varphi_0 + \varphi_1 g + \cdots + \varphi_m g^m$$

be two expressions satisfying the conditions of the theorem. Adding terms

equal to 0 to either side, we may assume that $m = d$. Subtracting, we get

$$0 = (f_0 - \varphi_0) + \cdots + (f_d - \varphi_d)g^d.$$

Hence g divides $f_0 - \varphi_0$, and since $\deg(f_0 - \varphi_0) < \deg g$ we see that $f_0 = \varphi_0$. Inductively, take the smallest integer i such that $f_i \neq \varphi_i$ (if such i exists). Dividing the above expression by g^i we find that g divides $f_i - \varphi_i$ and hence that such i cannot exist. This proves uniqueness.

We shall call the expression for f in terms of g in Theorem 5.3 the ***g*-adic expansion** of f. If $g(X) = X$, then the g-adic expansion is the usual expression of f as a polynomial.

Remark. In some sense, Theorem 5.2 redoes what was done in Theorem 8.1 of Chapter I for $\mathbf{Q}/\mathbf{Z}$; that is, express explicitly an element of K/A as a direct sum of its p-components.

§6. SYMMETRIC POLYNOMIALS

Let A be a commutative ring and let $t_1, \ldots, t_n$ be algebraically independent elements over A. Let X be a variable over $A[t_1, \ldots, t_n]$. We form the polynomial

$$\begin{aligned} F(X) &= (X - t_1) \cdots (X - t_n) \\ &= X^n - s_1 X^{n-1} + \cdots + (-1)^n s_n, \end{aligned}$$

where each $s_i = s_i(t_1, \ldots, t_n)$ is a polynomial in $t_1, \ldots, t_n$. Then for instance

$$s_1 = t_1 + \cdots + t_n \qquad \text{and} \qquad s_n = t_1 \cdots t_n.$$

The polynomials $s_1, \ldots, s_n$ are called the **elementary symmetric polynomials** of $t_1, \ldots, t_n$.

We leave it as an easy exercise to verify that s_i **is homogeneous of degree** i **in** $t_1, \ldots, t_n$.

Let σ be a permutation of the integers $(1, \ldots, n)$. Given a polynomial $f(t) \in A[t] = A[t_1, \ldots, t_n]$, we define σf to be

$$\sigma f(t_1, \ldots t_n) = f(t_{\sigma(1)}, \ldots, t_{\sigma(n)}).$$

If σ, τ are two permutations, then $\sigma\tau f = \sigma(\tau f)$ and hence the symmetric group G on n letters operates on the polynomial ring $A[t]$. A polynomial is called **symmetric** if $\sigma f = f$ for all $\sigma \in G$. It is clear that the set of symmetric polynomials is a subring of $A[t]$, which contains the constant polynomials

(i.e. A itself) and also contains the elementary symmetric polynomials $s_1, \ldots, s_n$. We shall see below that $A[s_1, \ldots, s_n]$ is the ring of symmetric polynomials.

Let $X_1, \ldots, X_n$ be variables. We define the **weight** of a monomial

$$X_1^{v_1} \cdots X_n^{v_n}$$

to be $v_1 + 2v_2 + \cdots + nv_n$. We define the weight of a polynomial $g(X_1, \ldots, X_n)$ to be the maximum of the weights of the monomials occurring in g.

Theorem 6.1. *Let $f(t) \in A[t_1, \ldots, t_n]$ be symmetric of degree d. Then there exists a polynomial $g(X_1, \ldots, X_n)$ of weight $\leqq d$ such that*

$$f(t) = g(s_1, \ldots, s_n).$$

If f is homogeneous of degree d, then every monomial occurring in g has weight d.

Proof. By induction on n. The theorem is obvious if $n = 1$, because $s_1 = t_1$. Assume the theorem proved for polynomials in $n - 1$ variables.

If we substitute $t_n = 0$ in the expression for $F(X)$, we find

$$(X - t_1) \cdots (X - t_{n-1})X = X^n - (s_1)_0 X^{n-1} + \cdots + (-1)^{n-1}(s_{n-1})_0 X,$$

where $(s_i)_0$ is the expression obtained by substituting $t_n = 0$ in s_i. We see that $(s_1)_0, \ldots, (s_{n-1})_0$ are precisely the elementary symmetric polynomials in $t_1, \ldots, t_{n-1}$.

We now carry out induction on d. If $d = 0$, our assertion is trivial. Assume $d > 0$, and assume our assertion proved for polynomials of degree $< d$. Let $f(t_1, \ldots, t_n)$ have degree d. There exists a polynomial $g_1(X_1, \ldots, X_{n-1})$ of weight $\leqq d$ such that

$$f(t_1, \ldots, t_{n-1}, 0) = g_1((s_1)_0, \ldots, (s_{n-1})_0).$$

We note that $g_1(s_1, \ldots, s_{n-1})$ has degree $\leqq d$ in $t_1, \ldots, t_n$. The polynomial

$$f_1(t_1, \ldots, t_n) = f(t_1, \ldots, t_n) - g_1(s_1, \ldots, s_{n-1})$$

has degree $\leqq d$ (in $t_1, \ldots, t_n$) and is symmetric. We have

$$f_1(t_1, \ldots, t_{n-1}, 0) = 0.$$

Hence f_1 is divisible by t_n, i.e. contains t_n as a factor. Since f_1 is symmetric, it contains $t_1 \cdots t_n$ as a factor. Hence

$$f_1 = s_n f_2(t_1, \ldots, t_n)$$

for some polynomial f_2, which must be symmetric, and whose degree is

$\leqq d - n < d$. By induction, there exists a polynomial g_2 in n variables and weight $\leqq d - n$ such that

$$f_2(t_1, \ldots, t_n) = g_2(s_1, \ldots, s_n).$$

We obtain

$$f(t) = g_1(s_1, \ldots, s_{n-1}) + s_n g_2(s_1, \ldots, s_n),$$

and each term on the right has weight $\leqq d$. This proves our theorem, except for the last statement which will be left to the reader.

We shall now prove that the elementary symmetric polynomials $s_1, \ldots, s_n$ are algebraically independent over A.

If they are not, take a polynomial $f(X_1, \ldots, X_n) \in A[X]$ of least degree and not equal to 0 such that

$$f(s_1, \ldots, s_n) = 0.$$

Write f as a polynomial in X_n with coefficients in $A[X_1, \ldots, X_{n-1}]$,

$$f(X_1, \ldots, X_n) = f_0(X_1, \ldots, X_{n-1}) + \cdots + f_d(X_1, \ldots, X_{n-1})X_n^d.$$

Then $f_0 \neq 0$. Otherwise, we can write

$$f(X) = X_n \psi(X)$$

with some polynomial ψ, and hence $s_n \psi(s_1, \ldots, s_n) = 0$. From this it follows that $\psi(s_1, \ldots, s_n) = 0$, and ψ has degree smaller than the degree of f.

We substitute s_i for X_i in the above relation, and get

$$0 = f_0(s_1, \ldots, s_{n-1}) + \cdots + f_d(s_1, \ldots, s_{n-1})s_n^d.$$

This is a relation in $A[t_1, \ldots, t_n]$, and we substitute 0 for t_n in this relation. Then all terms become 0 except the first one, which gives

$$0 = f_0((s_1)_0, \ldots, (s_{n-1})_0),$$

using the same notation as in the proof of Theorem 6.1. This is a non-trivial relation between the elementary symmetric polynomials in $t_1, \ldots, t_{n-1}$, a contradiction.

Example. (The Discriminant). Let $f(X) = (X - t_1) \cdots (X - t_n)$. Consider the product

$$\delta(t) = \prod_{i<j} (t_i - t_j).$$

For any permutation σ of $(1, \ldots, n)$ we see at once that

$$\delta^\sigma(t) = \pm \delta(t).$$

Hence $\delta(t)^2$ is symmetric, and we call it the **discriminant**:

$$D_f = D(s_1, \ldots, s_n) = \prod_{i<j} (t_i - t_j)^2.$$

We thus view the discriminant as a polynomial in the elementary symmetric functions. For a continuation of the general theory, see §8. We shall now consider special cases.

Quadratic case. You should verify that for a quadratic polynomial $f(X) = X^2 + bX + c$, one has

$$D = b^2 - 4c.$$

Cubic case. Consider $f(X) = X^3 + aX + b$. We wish to prove that

$$D = -4a^3 - 27b^2.$$

Observe first that D is homogeneous of degree 6 in t_1, t_2. Furthermore, a is homogeneous of degree 2 and b is homogeneous of degree 3. By Theorem 6.1 we know that there exists some polynomial $g(X_2, X_3)$ of weight 6 such that $D = g(a, b)$. The only monomials $X_2^m X_3^n$ of weight 6, i.e. such that $2m + 3n = 6$ with integers $m, n \geqq 0$, are those for which $m = 3$, $n = 0$, or $m = 0$ and $n = 2$. Hence

$$g(X_2, X_3) = vX_2^3 + wX_3^2$$

where v, w are integers which must now be determined.

Observe that the integers v, w are universal, in the sense that for any special polynomial with special values of a, b its discriminant will be given by $g(a, b) = va^3 + wb^2$.

Consider the polynomial

$$f_1(X) = X(X - 1)(X + 1) = X^3 - X.$$

Then $a = -1$, $b = 0$, and $D = va^3 = -v$. But also $D = 4$ by using the definition of the discriminant of the product of the differences of the roots, squared. Hence we get $v = -4$. Next consider the polynomial

$$f_2(X) = X^3 - 1.$$

Then $a = 0$, $b = -1$, and $D = wb^2 = w$. But the three roots of f_2 are the cube roots of unity, namely

$$1, \frac{-1 + \sqrt{-3}}{2}, \frac{-1 - \sqrt{-3}}{2}.$$

Using the definition of the discriminant we find the value $D = -27$. Hence we get $w = -27$. This concludes the proof of the formula for the discriminant of the cubic when there is no X^2 term.

In general, consider a cubic polynomial

$$f(X) = X^3 - s_1X^2 + s_2X - s_3 = (X - t_1)(X - t_2)(X - t_3).$$

We find the value of the discriminant by reducing this case to the simpler case when there is no X^2 term. We make a translation, and let

$$Y = X - \tfrac{1}{3}s_1 \qquad \text{so} \quad X = Y + \tfrac{1}{3}s_1 = Y + \tfrac{1}{3}(t_1 + t_2 + t_3).$$

Then $f(X)$ becomes

$$f(X) = f^*(Y) = Y^3 + aY + b = (Y - u_1)(Y - u_2)(Y - u_3),$$

where $a = u_1u_2 + u_2u_3 + u_1u_3$ and $b = -u_1u_2u_3$, while $u_1 + u_2 + u_3 = 0$. We have

$$u_i = t_i - \tfrac{1}{3}s_1 \qquad \text{for} \quad i = 1, 2, 3,$$

and $u_i - u_j = t_i - t_j$ for all $i \neq j$, so the discriminant is unchanged, and you can easily get the formula in general. Do Exercise 12(b).

§7. MASON-STOTHERS THEOREM AND THE *abc* CONJECTURE

In the early 80s a new trend of thought about polynomials started with the discovery of an entirely new relation. Let $f(t)$ be a polynomial in one variable over the complex numbers if you wish (an algebraically closed field of characteristic 0 would do). We define

$$n_0(f) = \text{number of distinct roots of } f.$$

Thus $n_0(f)$ counts the zeros of f by giving each of them multiplicity 1, and $n_0(f)$ can be small even though $\deg f$ is large.

Theorem 7.1 (Mason-Stothers, [Mas 84], [Sto 81]). *Let $a(t)$, $b(t)$, $c(t)$ be relatively prime polynomials such that $a + b = c$. Then*

$$\max \deg\{a, b, c\} \leqq n_0(abc) - 1.$$

Proof. (Mason) Dividing by c, and letting $f = a/c$, $g = b/c$ we have

$$f + g = 1,$$

where f, g are rational functions. Differentiating we get $f' + g' = 0$, which we rewrite as

$$\frac{f'}{f}f + \frac{g'}{g}g = 0,$$

so that

$$\frac{b}{a} = \frac{g}{f} = -\frac{f'/f}{g'/g}.$$

Let

$$a(t) = c_1 \prod (t - \alpha_i)^{m_i}, \qquad b(t) = c_2 \prod (t - \beta_j)^{n_j}, \qquad c(t) = c_3 \prod (t - \gamma_k)^{r_k}.$$

Then by calculus algebraicized in Exercise 11(c), we get

$$\frac{b}{a} = -\frac{f'/f}{g'/g} = -\frac{\sum \frac{m_i}{t - \alpha_i} - \sum \frac{r_k}{t - \gamma_k}}{\sum \frac{n_j}{t - \beta_j} - \sum \frac{r_k}{t - \gamma_k}}.$$

A common denominator for f'/f and g'/g is given by the product

$$N_0 = \prod (t - \alpha_i) \prod (t - \beta_j) \prod (t - \gamma_k),$$

whose degree is $n_0(abc)$. Observe that $N_0 f'/f$ and $N_0 g'/g$ are both polynomials of degrees at most $n_0(abc) - 1$. From the relation

$$\frac{b}{a} = -\frac{N_0 f'/f}{N_0 g'/g},$$

and the fact that a, b are assumed relatively prime, we deduce the inequality in the theorem.

As an application, let us prove **Fermat's theorem** for polynomials. Thus let $x(t)$, $y(t)$, $z(t)$ be relatively prime polynomials such that one of them has degree $\geqq 1$, and such that

$$x(t)^n + y(t)^n = z(t)^n.$$

We want to prove that $n \leqq 2$. By the Mason-Stothers theorem, we get

$$n \deg x = \deg x(t)^n \leqq \deg x(t) + \deg y(t) + \deg z(t) - 1,$$

and similarly replacing x by y and z on the left-hand side. Adding, we find

$$n(\deg x + \deg y + \deg z) \leqq 3(\deg x + \deg y + \deg z) - 3.$$

This yields a contradiction if $n \geqq 3$.

As another application in the same vein, one has:

Davenport's theorem. *Let f, g be non-constant polynomials such that $f^3 - g^2 \neq 0$. Then*

$$\deg(f^3 - g^2) \geqq \tfrac{1}{2} \deg f - 1.$$

See Exercise 13.

One of the most fruitful analogies in mathematics is that between the integers $\mathbf{Z}$ and the ring of polynomials $F[t]$ over a field F. Evolving from the insights of Mason [Ma 84], Frey [Fr 87], Szpiro, and others, Masser and Oesterle formulated the *abc* conjecture for integers as follows. Let m be a non-zero integer. Define the **radical** of m to be

$$N_0(m) = \prod_{p|m} p,$$

i.e. the product of all the primes dividing m, taken with multiplicity 1.

The *abc* conjecture. *Given $\varepsilon > 0$, there exists a positive number $C(\varepsilon)$ having the following property. For any non-zero relative prime integers a, b, c such that $a + b = c$, we have*

$$\max(|a|, |b|, |c|) \leqq C(\varepsilon) N_0(abc)^{1+\varepsilon}.$$

Observe that the inequality says that many prime factors of a, b, c occur to the first power, and that if "small" primes occur to high powers, then they have to be compensated by "large" primes occurring to the first power. For instance, one might consider the equation

$$2^n \pm 1 = m.$$

For m large, the *abc* conjecture would state that m has to be divisible by large primes to the first power. This phenomenon can be seen in the tables of [BLSTW 83].

Stewart–Tijdeman [ST 86] have shown that it is necessary to have the ε in the formulation of the conjecture. Subsequent examples were communicated to me by Wojtek Jastrzebowski and Dan Spielman as follows.

We have to give examples such that for all $C > 0$ there exist natural numbers a, b, c relatively prime such that $a + b = c$ and $|a| > CN_0(abc)$. But trivially,

$$2^n | (3^{2^n} - 1).$$

We consider the relations $a_n + b_n = c_n$ given by

$$3^{2^n} - 1 = c_n.$$

It is clear that these relations provide the desired examples. Other examples can be constructed similarly, since the role of 3 and 2 can be played by other integers. Replace 2 by some prime, and 3 by an integer $\equiv 1 \bmod p$.

The *abc* conjecture implies what we shall call the

Asymptotic Fermat Theorem. *For all n sufficiently large, the equation*

$$x^n + y^n = z^n$$

has no solution in relatively prime integers $\neq 0$.

The proof follows exactly the same pattern as for polynomials, except that we write things down multiplicatively, and there is a $1 + \varepsilon$ floating around. The extent to which the *abc* conjecture will be proved with an explicit constant $C(\varepsilon)$ (or say $C(1)$ to fix ideas) yields the corresponding explicit determination of the bound for n in the application. We now go into other applications.

Hall's conjecture [Ha 71]. *If u, v are relatively prime non-zero integers such that $u^3 - v^2 \neq 0$, then*

$$|u^3 - v^2| \gg |u|^{1/2-\varepsilon}.$$

The symbol $\gg$ means that the left-hand side is $\geqq$ the right-hand side times a constant depending only on ε. Again the proof is immediate from the *abc* conjecture. Actually, the hypothesis that u, v are relatively prime is not necessary; the general case can be reduced to the relatively prime case by extracting common factors, and Hall stated his conjecture in this more general way. However, he also stated it without the epsilon in the exponent, and that does not work, as was realized later. As in the polynomial case, Hall's conjecture describes how small $|u^3 - v^2|$ can be, and the answer is not too small, as described by the right-hand side.

The Hall conjecture can also be interpreted as giving a bound for integral relatively prime solutions of

$$v^2 = u^3 + b \quad \text{with integral } b.$$

Then we find

$$|u| \ll |b|^{2+\varepsilon}.$$

More generally, in line with conjectured inequalities from Lang–Waldschmidt [La 78], let us fix non-zero integers A, B and let u, v, k, m, n be variable, with u, v relatively prime and $mv > m + n$. Put

$$Au^m + Bv^n = k.$$

By the *abc* conjecture, one derives easily that

$$(1) \qquad |u| \ll N_0(k)^{\frac{n}{mn-(m+n)}(1+\varepsilon)} \qquad \text{and} \qquad |v| \ll N_0(k)^{\frac{m}{mn-(m+n)}(1+\varepsilon)}.$$

From this one gets

$$|k| \ll N_0(k)^{\frac{mn}{mn-(m+n)}(1+\varepsilon)}.$$

The Hall conjecture is a special case after we replace $N_0(k)$ with $|k|$, because $N_0(k) \leqq |k|$.

Next take $m = 3$ and $n = 2$, but take $A = 4$ and $B = -27$. In this case we write

$$D = 4u^3 - 27v^2$$

and we get

$$|u| \ll N_0(D)^{2+\varepsilon} \quad \text{and} \quad |v| \ll N_0(D)^{3+\varepsilon}. \tag{2}$$

These inequalities are supposed to hold at first for u, v relatively prime. Suppose we allow u, v to have some bounded common factor, say d. Write

$$u = u'd \quad \text{and} \quad v = v'd$$

with u', v' relatively prime. Then

$$D = 4d^3u'^3 - 27d^2v'^2.$$

Now we can apply inequality (1) with $A = 4d^3$ and $B = -27d^2$, and we find the same inequalities (2), with the constant implicit in the sign $\ll$ depending also on d, or on some fixed bound for such a common factor. Under these circumstances, we call inequalities (2) the **generalized Szpiro conjecture**.

The original Szpiro conjecture was stated in a more sophisticated situation, cf. [La 90] for an exposition, and Szpiro's inequality was stated in the form

$$|D| \ll N(D)^{6+\varepsilon},$$

where $N(D)$ is a more subtle invariant, but for our purposes, it is sufficient and much easier to use the radical $N_0(D)$.

The point of D is that it occurs as a discriminant. The trend of thoughts in the direction we are discussing was started by Frey [Fr 87], who associated with each solution of $a + b = c$ the polynomial

$$x(x - a)(x + b),$$

which we call the **Frey polynomial**. (Actually Frey associated the curve defined by the equation $y^2 = x(x - a)(x + b)$, for much deeper reasons, but only the polynomial on the right-hand side will be needed here.) The discriminant of the polynomial is the product of the differences of the roots squared, and so

$$D = (abc)^2.$$

We make a translation

$$\xi = x + \frac{b - a}{3}$$

to get rid of the x^2-term, so that our polynomial can be rewritten

$$\xi^3 - \gamma_2\xi - \gamma_3,$$

where γ_2, γ_3 are homogeneous in a, b of appropriate weight. The discriminant does not change because the roots of the polynomial in ξ are

translations of the roots of the polynomial in x. Then

$$D = 4\gamma_2^3 - 27\gamma_3^2.$$

The translation with $(b - a)/3$ introduces a small denominator. One may avoid this denominator by using the polynomial $x(x - 3a)(x - 3b)$, so that γ_2, γ_3 then come out to be integers, and one can apply the generalized Szpiro conjecture to the discriminant, which then has an extra factor $D = 3^6(abc)^2$.

It is immediately seen that the generalized Szpiro conjecture implies asymptotic Fermat. Conversely:

Generalized Szpiro implies the abc conjecture.

Indeed, the correspondence $(a, b) \leftrightarrow (\gamma_2, \gamma_3)$ is invertible, and has the "right" weight. A simple algebraic manipulation shows that the generalized Szpiro estimates on γ_2, γ_3 imply the desired estimates on $|a|$, $|b|$. (Do Exercise 14.) From the equivalence between *abc* and generalized Szpiro, one can use the examples given earlier to show that the epsilon is needed in the Szpiro conjecture.

Finally, note that the polynomial case of the Mason-Stothers theorem and the case of integers are not independent, or specifically the Davenport theorem and Hall's conjecture are related. Examples in the polynomial case parametrize cases with integers when we substitute integers for the variables. Such examples are given in [BCHS 65], one of them (due to Birch) being

$$f(t) = t^6 + 4t^4 + 10t^2 + 6 \qquad \text{and} \qquad g(t) = t^9 + 6t^7 + 21t^5 + 35t^3 + \tfrac{63}{2}t,$$

whence

$$\deg(f(t)^3 - g(t)^2) = \tfrac{1}{2}\deg f + 1.$$

This example shows that Davenport's inequality is best possible, because the degree attains the lowest possible value permissible under the theorem. Substituting large integral values of $t \equiv 2 \bmod 4$ gives examples of similarly low values for $x^3 - y^2$. For other connections of all these matters, cf. [La 90].

Bibliography

[BCHS 65] B. Birch, S. Chowla, M. Hall, and A. Schinzel, On the difference $x^3 - y^2$, *Norske Vid. Selsk. Forrh.* **38** (1965) pp. 65–69

[BLSTW 83] J. Brillhart, D. H. Lehmer, J. L. Selfridge, B. Tuckerman, and S. Wagstaff Jr., Factorization of $b^n \pm 1$, $b = 2, 3, 5, 6, 7, 10, 11$ up to high powers, *Contemporary Mathematics* Vol. **22**, AMS, Providence, RI, 1983

[Dav 65] H. Davenport, On $f^3(t) - g^2(t)$, *Norske Vid. Selsk. Forrh.* **38** (1965) pp. 86–87

[Fr 87] G. Frey, Links between solutions of $A - B = C$ and elliptic curves, *Number Theory, Lecture Notes* **1380**, Springer-Verlag, New York, 1989 pp. 31–62

[Ha 71] M. HALL, The diophantine equation $x^3 - y^2 = k$, *Computers and Number Theory*, ed. by A. O. L. Atkin and B. Birch, Academic Press, London 1971 pp. 173–198

[La 90] S. LANG, Old and new conjectured diophantine inequalities, *Bull. AMS* Vol. **23** No. 1 (1990) pp. 37–75

[Ma 84a] R. C. MASON, Equations over function fields, Springer Lecture Notes **1068** (1984), pp. 149–157; in *Number Theory, Proceedings of the Noordwijkerhout*, 1983

[Ma 84b] R. C. MASON, Diophantine equations over function fields, *London Math. Soc. Lecture Note Series* Vol. **96**, Cambridge University Press, Cambridge, 1984

[Ma 84c] R. C. MASON, The hyperelliptic equation over function fields, *Math. Proc. Cambridge Philos. Soc.* **93** (1983) pp. 219–230

[Si 88] J. SILVERMAN, Wieferich's criterion and the *abc* conjecture, *Journal of Number Theory* **30** (1988) pp. 226–237

[ST 86] C. L. STEWART and R. TIJDEMAN, On the Oesterle–Masser Conjecture, *Mon. Math.* **102** (1986) pp. 251–257

See additional references at the end of the chapter.

§8. THE RESULTANT

In this section, we assume that the reader is familiar with determinants. The theory of determinants will be covered later. The section can be viewed as giving further examples of symmetric functions.

Let A be a commutative ring and let $v_0, \ldots, v_n$, $w_0, \ldots, w_m$ be algebraically independent over A. We form two polynomials:

$$f_v(X) = v_0 X^n + \cdots + v_n,$$

$$g_w(X) = w_0 X^m + \cdots + w_m.$$

We define the **resultant** of (v, w), or of f_v, g_w, to be the determinant

$$\begin{vmatrix}
v_0 & v_1 & \cdots & v_n & & \\
 & v_0 & v_1 & \cdots & v_n & \\
 & & \cdots & \cdots & \cdots & \cdots \\
 & & & v_0 & v_1 & \cdots & v_n \\
w_0 & w_1 & \cdots & w_m & & \\
 & w_0 & w_1 & \cdots & w_m & \\
 & & \cdots & \cdots & \cdots & \cdots \\
 & & & w_0 & w_1 & \cdots & w_m
\end{vmatrix}$$

(m rows of the v-coefficients, n rows of the w-coefficients; $m + n$ columns)

The blank spaces are supposed to be filled with zeros.

If we substitute elements $(a) = (a_0, \ldots, a_n)$ and $(b) = (b_0, \ldots, b_m)$ in A for (v) and (w) respectively in the coefficients of f_v and g_w, then we obtain polynomials f_a and g_b with coefficients in A, and *we define their* **resultant** *to be the determinant obtained by substituting* (a) *for* (v) *and* (b) *for* (w) *in the determinant.* We shall write the resultant of f_v, g_w in the form

$$\operatorname{Res}(f_v, g_w) \quad \text{or} \quad R(v, w).$$

The resultant $\operatorname{Res}(f_a, g_b)$ is then obtained by substitution of (a), (b) for (v), (w) respectively.

We observe that $R(v, w)$ is a polynomial with integer coefficients, i.e. we may take $A = \mathbf{Z}$. If z is a variable, then

$$R(zv, w) = z^m R(v, w) \qquad \text{and} \qquad R(v, zw) = z^n R(v, w)$$

as one sees immediately by factoring out z from the first m rows (resp. the last n rows) in the determinant. Thus R is homogeneous of degree m in its first set of variables, and homogeneous of degree n in its second set of variables. Furthermore, $R(v, w)$ contains the monomial

$$v_0^m w_m^n$$

with coefficient 1, when expressed as a sum of monomials.

If we substitute 0 for v_0 and w_0 in the resultant, we obtain 0, because the first column of the determinant vanishes.

Let us work over the integers $\mathbf{Z}$. We consider the linear equations

$$\begin{aligned}
X^{m-1}f_v(X) &= v_0X^{n+m-1} + v_1X^{n+m-2} + \cdots + v_nX^{m-1} \\
X^{m-2}f_v(X) &= \qquad v_0X^{n+m-2} + \cdots + v_nX^{m-2} \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
f_v(X) &= \qquad\qquad v_0X^n + \cdots + v_n \\
X^{n-1}g_w(X) &= w_0X^{n+m-1} + w_1X^{n+m-2} + \cdots + w_mX^{n-1} \\
X^{n-2}g_w(X) &= \qquad w_0X^{n+m-2} + \cdots + w_mX^{n-2} \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
g_w(X) &= \qquad\qquad w_0X^m + \cdots + w_m.
\end{aligned}$$

Let C be the column vector on the left-hand side, and let

$$C_0, \ldots, C_{m+n}$$

be the column vectors of coefficients. Our equations can be written

$$C = X^{n+m-1}C_0 + \cdots + 1 \cdot C_{m+n}.$$

By Cramer's rule, applied to the last coefficient which is $= 1$,

$$R(v, w) = \det(C_0, \ldots, C_{m+n}) = \det(C_0, \ldots, C_{m+n-1}, C).$$

From this we see that there exist polynomials $\varphi_{v,w}$ *and* $\psi_{v,w}$ *in* $\mathbf{Z}[v, w][X]$ *such that*

$$\varphi_{,v,w}f_v + \psi_{v,w}g_w = R(v,w) = \operatorname{Res}(f_v, f_w).$$

Note that $R(v, w) \in \mathbf{Z}[v, w]$ but that the polynomials on the left-hand side involve the variable X.

If $\lambda\colon \mathbf{Z}[v, w] \to A$ is a homomorphism into a commutative ring A and we let $\lambda(v) = (a)$, $\lambda(w) = (b)$, then

$$\varphi_{a,b}f_a + \psi_{a,b}g_b = R(a, b) = \operatorname{Res}(f_a, f_b).$$

Thus from the universal relation of the resultant over $\mathbf{Z}$ we obtain a similar relation for every pair of polynomials, in any commutative ring A.

Proposition 8.1. *Let* K *be a subfield of a field* L, *and let* f_a, g_b *be polynomials in* $K[X]$ *having a common root* ξ *in* L. *Then* $R(a, b) = 0$.

Proof. If $f_a(\xi) = g_b(\xi) = 0$, then we substitute ξ for X in the expression obtained for $R(a, b)$ and find $R(a, b) = 0$.

Next, we shall investigate the relationship between the resultant and the roots of our polynomials f_v, g_w. We need a lemma.

Lemma 8.2. *Let* $h(X_1, \ldots, X_n)$ *be a polynomial in* n *variables over the integers* $\mathbf{Z}$. *If* h *has the value* 0 *when we substitute* X_1 *for* X_2 *and leave the other* X_i *fixed* $(i \neq 2)$, *then* $h(X_1, \ldots, X_n)$ *is divisible by* $X_1 - X_2$ *in* $\mathbf{Z}[X_1, \ldots, X_n]$.

Proof. Exercise for the reader.

Let $v_0, t_1, \ldots, t_n, w_0, u_1, \ldots, u_m$ be algebraically independent over $\mathbf{Z}$ and form the polynomials

$$f_v = v_0(X - t_1) \cdots (X - t_n) = v_0X^n + \cdots + v_n,$$
$$g_w = w_0(X - u_1) \cdots (X - u_m) = w_0X^m + \cdots + w_m.$$

Thus we let

$$v_i = (-1)^i v_0 s_i(t) \qquad \text{and} \qquad w_j = (-1)^j w_0 s_j(u).$$

We leave to the reader the easy verification that

$$v_0, v_1, \ldots, v_n, w_0, w_1, \ldots, w_m$$

are algebraically independent over $\mathbf{Z}$.

Proposition 8.3. *Notation being as above, we have*

$$\operatorname{Res}(f_v, g_w) = v_0^m w_0^n \prod_{i=1}^{n} \prod_{j=1}^{m} (t_i - u_j).$$

Proof. Let S be the expression on the right-hand side of the equality in the statement of the proposition.

Since $R(v, w)$ is homogeneous of degree m in its first variables, and homogeneous of degree n in its second variables, it follows that

$$R = v_0^m w_0^n h(t, u)$$

where $h(t, u) \in \mathbf{Z}[t, u]$. By Proposition 8.1, the resultant vanishes when we substitute t_i for u_j ($i = 1, \ldots, n$ and $j = 1, \ldots, m$), whence by the lemma, viewing R as an element of $\mathbf{Z}[v_0, w_0, t, u]$ it follows that R is divisible by $t_i - u_j$ for each pair (i, j). Hence S divides R in $\mathbf{Z}[v_0, w_0, t, u]$, because $t_i - u_j$ is obviously a prime in that ring, and different pairs (i, j) give rise to different primes.

From the product expression for S, namely

$$S = v_0^m w_0^n \prod_{i=1}^{n} \prod_{j=1}^{m} (t_i - u_j), \tag{1}$$

we obtain

$$\prod_{i=1}^{n} g(t_i) = w_0^n \prod_{i=1}^{n} \prod_{j=1}^{m} (t_i - u_j),$$

whence

$$S = v_0^m \prod_{i=1}^{n} g(t_i). \tag{2}$$

Similarly,

$$S = (-1)^{nm} w_0^n \prod_{j=1}^{m} f(u_j). \tag{3}$$

From (2) we see that S is homogeneous and of degree n in (w), and from (3) we see that S is homogeneous and of degree m in (v). Since R has exactly the same homogeneity properties, and is divisible by S, it follows that $R = cS$ for some integer c. Since both R and S have a monomial $v_0^m w_m^n$ occurring in them with coefficient 1, it follows that $c = 1$, and our proposition is proved.

We also note that the three expressions found for S above now give us a factorization of R. We also get a converse for Proposition 8.1.

Corollary 8.4. *Let f_a, g_b be polynomials with coefficients in a field K, such that $a_0 b_0 \neq 0$, and such that f_a, g_b split in factors of degree 1 in $K[X]$. Then* $\operatorname{Res}(f_a, g_b) = 0$ *if and only if f_a and g_b have a root in common.*

Proof. Assume that the resultant is 0. If

$$f_a = a_0(X - \alpha_1) \cdots (X - \alpha_n),$$

$$g_b = b_0(X - \beta_1) \cdots (X - \beta_n),$$

is the factorization of f_a, g_b, then we have a homomorphism

$$\mathbf{Z}[v_0, t, w_0, u] \to K$$

such that $v_0 \mapsto a_0$, $w_0 \mapsto b_0$, $t_i \mapsto \alpha_i$, and $u_j \mapsto \beta_j$ for all i, j. Then

$$0 = \operatorname{Res}(f_a, g_b) = a_0^m b_0^n \prod_i \prod_j (\alpha_i - \beta_j),$$

whence f_a, f_b have a root in common. The converse has already been proved.

We deduce one more relation for the resultant in a special case. Let f_v be as above,

$$f_v(X) = v_0 X^n + \cdots + v_n = v_0(X - t_1) \cdots (X - t_n).$$

From (2) we know that if f_v' is the derivative of f_v, then

$$\operatorname{Res}(f_v, f_v') = v_0^{n-1} \prod_i f'(t_i). \tag{4}$$

Using the product rule for differentiation, we find:

$$f_v'(X) = \sum_i v_0(X - t_1) \cdots \widehat{(X - t_i)} \cdots (X - t_n),$$

$$f_v'(t_i) = v_0(t_i - t_1) \cdots \widehat{(t_i - t_i)} \cdots (t_i - t_n),$$

where a roof over a term means that this term is to be omitted.

We define the **discriminant** of f_v to be

$$D(f_v) = D(v) = (-1)^{n(n-1)/2} v_0^{2n-2} \prod_{i \neq j} (t_i - t_j).$$

Proposition 8.5. *Let f_v be as above and have algebraically independent coefficients over $\mathbf{Z}$. Then*

$$\operatorname{Res}(f_v, f_v') = v_0^{2n-1} \prod_{i \neq j} (t_i - t_j) = (-1)^{n(n-1)/2} v_0 D(f_v). \tag{5}$$

Proof. One substitutes the expression obtained for $f_v'(t_i)$ into the product (4). The result follows at once.

When we substitute 1 for v_0, we find that the discriminant as we defined it in the preceding section coincides with the present definition. In particular, we find an explicit formula for the discriminant. The formulas in the special case of polynomials of degree 2 and 3 will be given as exercises.

Note that the discriminant can also be written as the product

$$D(f_v) = v_0^{2n-2} \prod_{i<j} (t_i - t_j)^2.$$

Serre once pointed out to me that the sign $(-1)^{n(n-1)/2}$ was missing in the first edition of this book, and that this sign error is quite common in the literature, occurring as it does in van der Waerden, Samuel, and Hilbert (but not in his collected works, corrected by Olga Taussky); on the other hand the sign is correctly given in Weber's *Algebra*, Vol. I, 50.

For a continuation of this section, see Chapter IX, §3 and §4.

§9. POWER SERIES

Let X be a letter, and let G be the monoid of functions from the set $\{X\}$ to the natural numbers. If $v \in \mathbf{N}$, we denote by X^v the function whose value at X is v. Then G is a multiplicative monoid, already encountered when we discussed polynomials. Its elements are $X^0, X^1, X^2, \ldots, X^v, \ldots$.

Let A be a commutative ring, and let $A[[X]]$ be the set of functions from G into A, without any restriction. Then an element of $A[[X]]$ may be viewed as assigning to each monomial X^v a coefficient $a_v \in A$. We denote this element by

$$\sum_{v=0}^{\infty} a_v X^v.$$

The summation symbol is not a sum, of course, but we shall write the above expression also in the form

$$a_0 X^0 + a_1 X^1 + \cdots$$

and we call it a **formal power series** with coefficients in A, in one variable. We call $a_0, a_1, \ldots$ its coefficients.

Given two elements of $A[[X]]$, say

$$\sum_{v=0}^{\infty} a_v X^v \quad \text{and} \quad \sum_{\mu=0}^{\infty} b_\mu X^\mu,$$

we define their product to be

$$\sum_{i=0}^{\infty} c_i X^i$$

where

$$c_i = \sum_{v+\mu=i} a_v b_\mu.$$

Just as with polynomials, one defines their sum to be

$$\sum_{v=0}^{\infty} (a_v + b_v) X^v.$$

Then we see that the power series form a ring, the proof being the same as for polynomials.

One can also construct the power series ring in several variables $A[[X_1, \ldots, X_n]]$ in which every element can be expressed in the form

$$\sum_{(v)} a_{(v)} X_1^{v_1} \cdots X_n^{v_n} = \sum a_{(v)} M_{(v)}(X_1, \ldots, X_n)$$

with unrestricted coefficients $a_{(v)}$ in bijection with the n-tuples of integers $(v_1, \ldots, v_n)$ such that $v_i \geqq 0$ for all i. It is then easy to show that there is an isomorphism between $A[[X_1, \ldots, X_n]]$ and the repeated power series ring $A[[X_1]] \cdots [[X_n]]$. We leave this as an exercise for the reader.

The next theorem will give an analogue of the Euclidean algorithm for power series. However, instead of dealing with power series over a field, it is important to have somewhat more general coefficients for certain applications, so we have to introduce a little more terminology.

Let A be a ring and I an ideal. We assume that

$$\bigcap_{v=1}^{\infty} I^v = \{0\}.$$

We can view the powers I^v as defining neighborhoods of 0 in A, and we can transpose the usual definition of Cauchy sequence in analysis to this situation, namely: we define a sequence $\{a_n\}$ in A to be **Cauchy** if given some power I^v there exists an integer N such that for all $m, n \geqq N$ we have

$$a_m - a_n \in I^v.$$

Thus I^v corresponds to the given ϵ of analysis. Then we have the usual notion of **convergence** of a sequence to an element of A. One says that A is **complete in the I-adic topology** if every Cauchy sequence converges.

Perhaps the most important example of this situation is when A is a local ring and $I = \mathfrak{m}$ is its maximal ideal. By a **complete local ring**, one always means a local ring which is complete in the $\mathfrak{m}$-adic topology.

Let k be a field. Then the power series ring

$$R = k[[X_1, \ldots, X_n]]$$

in n variables is such a complete local ring. Indeed, let $\mathfrak{m}$ be the ideal generated by the variables $X_1, \ldots, X_n$. Then $R/\mathfrak{m}$ is naturally isomorphic to the field k itself, so $\mathfrak{m}$ is a maximal ideal. Furthermore, any power series of the form

$$f(X) = c_0 - f_1(X)$$

with $c_0 \in k$, $c_0 \neq 0$ and $f_1(X) \in \mathfrak{m}$ is invertible. To prove this, one may first assume without loss of generality that $c_0 = 1$. Then

$$(1 - f_1(X))^{-1} = 1 + f_1(X) + f_1(X)^2 + f_1(X)^3 + \cdots$$

gives the inverse. Thus we see that $\mathfrak{m}$ is the unique maximal ideal and R is local. It is immediately verified that R is complete in the sense we have just defined. The same argument shows that if k is not a field but c_0 is invertible in k, then again $f(X)$ is invertible.

Again let A be a ring. We may view the power series ring in n variables ($n > 1$) as the ring of power series in one variable X_n over the ring of power series in $n - 1$ variables, that is we have a natural identification

$$A[[X_1, \ldots, X_n]] = A[[X_1, \ldots, X_{n-1}]][[X_n]].$$

If $A = k$ is a field, the ring $k[[X_1, \ldots, X_{n-1}]]$ is then a complete local ring. More generally, if $\mathfrak{o}$ is a complete local ring, then the power series ring $\mathfrak{o}[[X]]$ is a complete local ring, whose maximal ideal is $(\mathfrak{m}, X)$ where $\mathfrak{m}$ is the maximal ideal of $\mathfrak{o}$. Indeed, if a power series $\sum a_v X^v$ has unit constant

term $a_0 \in \mathfrak{o}^*$, then the power series is a unit in $\mathfrak{o}[[X]]$, because first, without loss of generality, we may assume that $a_0 = 1$, and then we may invert $1 + h$ with $h \in (\mathfrak{m}, X)$ by the geometric series $1 - h + h^2 - h^3 + \cdots$.

In a number of problems, it is useful to reduce certain questions about power series in several variables over a field to questions about power series in one variable over the more complicated ring as above. We shall now apply this decomposition to the Euclidean algorithm for power series.

Theorem 9.1. *Let $\mathfrak{o}$ be a complete local ring with maximal ideal $\mathfrak{m}$. Let*

$$f(X) = \sum_{i=0}^{\infty} a_i X^i$$

be a power series in $\mathfrak{o}[[X]]$ (one variable), such that not all a_i lie in $\mathfrak{m}$. Say $a_0, \ldots, a_{n-1} \in \mathfrak{m}$, and $a_n \in \mathfrak{o}^*$ is a unit. Given $g \in \mathfrak{o}[[X]]$ we can solve the equation

$$g = qf + r$$

uniquely with $q \in \mathfrak{o}[[X]]$, $r \in \mathfrak{o}[X]$, and $\deg r \leqq n - 1$.

Proof (Manin). Let α and τ be the projections on the beginning and tail end of the series, given by

$$\alpha: \sum b_i X^i \mapsto \sum_{i=0}^{n-1} b_i X^i = b_0 + b_1 X + \cdots + b_{n-1} X^{n-1},$$

$$\tau: \sum b_i X^i \mapsto \sum_{i=n}^{\infty} b_i X^{i-n} = b_n + b_{n+1} X + b_{n+2} X^2 + \cdots.$$

Note that $\tau(hX^n) = h$ for any $h \in \mathfrak{o}[[X]]$; and h is a polynomial of degree $< n$ if and only if $\tau(h) = 0$.

The existence of q, r is equivalent with the condition that there exists q such that

$$\tau(g) = \tau(qf).$$

Hence our problem is equivalent with solving

$$\tau(g) = \tau\big(q\alpha(f)\big) + \tau\big(q\tau(f)X^n\big) = \tau\big(q\alpha(f)\big) + q\tau(f).$$

Note that $\tau(f)$ is invertible. Put $Z = q\tau(f)$. Then the above equation is equivalent with

$$\tau(g) = \tau\left(Z\frac{\alpha(f)}{\tau(f)}\right) + Z = \left(I + \tau \circ \frac{\alpha(f)}{\tau(f)}\right)Z.$$

Note that

$$\tau \circ \frac{\alpha(f)}{\tau(f)}: \mathfrak{o}[[X]] \to \mathfrak{m}\mathfrak{o}[[X]],$$

because $\alpha(f)/\tau(f) \in \mathfrak{m}\mathfrak{o}[[X]]$. We can therefore invert to find Z, namely

$$Z = \left(I + \tau \circ \frac{\alpha(f)}{\tau(f)}\right)^{-1} \tau(g),$$

which proves both existence and uniqueness and concludes the proof.

Theorem 9.2. (Weierstrass Preparation). *The power series f in the previous theorem can be written uniquely in the form*

$$f(X) = (X^n + b_{n-1}X^{n-1} + \cdots + b_0)u,$$

where $b_i \in \mathfrak{m}$, and u is a unit in $\mathfrak{o}[[X]]$.

Proof. Write uniquely

$$X^n = qf + r,$$

by the Euclidean algorithm. Then q is invertible, because

$$q = c_0 + c_1 X + \cdots,$$

$$f = \cdots + a_n X^n + \cdots,$$

so that

$$1 \equiv c_0 a_n \pmod{\mathfrak{m}},$$

and therefore c_0 is a unit in $\mathfrak{o}$. We obtain $qf = X^n - r$, and

$$f = q^{-1}(X^n - r),$$

with $r \equiv 0 \pmod{\mathfrak{m}}$. This proves the existence. Uniqueness is immediate.

The integer n in Theorems 9.1 and 9.2 is called the **Weierstrass degree** of f, and is denoted by $\deg_W f$. We see that a power series not all of whose coefficients lie in $\mathfrak{m}$ can be expressed as a product of a polynomial having the given Weierstrass degree, times a unit in the power series ring. Furthermore, all the coefficients of the polynomial except the leading one lie in the maximal ideal. Such a polynomial is called **distinguished**, or a **Weierstrass polynomial.**

Remark. I rather like the use of the Euclidean algorithm in the proof of the Weierstrass Preparation theorem. However, one can also give a direct proof exhibiting explicitly the recursion relations which solve for the coefficients of u, as follows. Write $u = \sum c_i X^i$. Then we have to solve the equations

$$b_0 c_0 = a_0,$$

$$b_0 c_1 + b_1 c_0 = a_1,$$

$$\cdots$$

$$b_0 c_{n-1} + \cdots + b_{n-1} c_0 = a_{n-1},$$

$$b_0 c_n + \cdots + c_0 = a_n,$$

$$b_0 c_{n+1} + \cdots + c_1 = a_{n+1},$$

$$\cdots$$

In fact, the system of equations has a unique solution mod $\mathfrak{m}^r$ for each positive integer r, after selecting c_0 to be a unit, say $c_0 = 1$. Indeed, from the first n equations (from 0 to $n-1$) we see that $b_0, \ldots, b_{n-1}$ are uniquely determined to be 0 mod $\mathfrak{m}$. Then $c_n, c_{n+1}, \ldots$ are uniquely determined mod $\mathfrak{m}$ by the subsequent equations. Now inductively, suppose we have shown that the coefficients b_i, c_j are uniquely determined mod $\mathfrak{m}^r$. Then one sees immediately that from the conditions $a_0, \ldots, a_{n-1} \equiv 0 \bmod \mathfrak{m}$ the first n equations define b_i uniquely mod $\mathfrak{m}^{r+1}$ because all $b_i \equiv 0 \bmod \mathfrak{m}$. Then the subsequent equations define $c_j \bmod \mathfrak{m}^{r+1}$ uniquely from the values of $b_i \bmod \mathfrak{m}^{r+1}$ and $c_j \bmod \mathfrak{m}^r$. The unique system of solutions mod $\mathfrak{m}^r$ for each r then defines a solution in the projective limit, which is the complete local ring.

We now have all the tools to deal with unique factorization in one important case.

Theorem 9.3. *Let k be a field. Then $k[[X_1, \ldots, X_n]]$ is factorial.*

Proof. Let $f(x) = f(X_1, \ldots, X_n) \in k[[X]]$ be $\neq 0$. After making a sufficiently general linear change of variables (when k is infinite)

$$x_i = \sum c_{ij} Y_j \quad \text{with} \quad c_{ij} \in k,$$

we may assume without loss of generality that $f(0, \ldots, 0, x_n) \neq 0$. (When k is finite, one has to make a non-linear change, cf. Theorem 2.1 of Chapter VIII.) Indeed, if we write $f(X) = f_d(X) +$ higher terms, where $f_d(X)$ is a homogeneous polynomial of degree $d \geqq 0$, then changing the variables as above preserves the degree of each homogeneous component of f, and since k is assumed infinite, the coefficients c_{ij} can be taken so that in fact each power Y_i^d $(i = 1, \ldots, n)$ occurs with non-zero coefficient.

We now proceed by induction on n. Let $R_n = k[[X_1, \ldots, X_n]]$ be the power series in n variables, and assume by induction that R_{n-1} is factorial. By Theorem 9.2, write $f = gu$ where u is a unit and g is a Weierstrass polynomial in $R_{n-1}[X_n]$. By Theorem 2.3, $R_{n-1}[X_n]$ is factorial, and so we can write g as a product of irreducible elements $g_1, \ldots, g_r \in R_{n-1}[X_n]$, so $f = g_1 \cdots g_r u$, where the factors g_i are uniquely determined up to multiplication by units. This proves the existence of a factorization. As to uniqueness, suppose f is expressed as a product of irreducible elements in R_n, $f = f_1 \cdots f_s$. Then $f_q(0, \ldots, 0, x_n) \neq 0$ for each $q = 1, \ldots, s$, so we can write $f_q = h_q u_q'$ where u_q' is a unit and h_q is a Weierstrass polynomial, necessarily irreducible in $R_{n-1}[X_n]$. Then $f = gu = \prod h_q \prod u_q'$ with g and all h_q Weierstrass polynomials. By Theorem 9.2, we must have $g = \prod h_q$, and since $R_{n-1}[X_n]$ is factorial, it follows that the polynomials h_q are the same as the polynomials g_i, up to units. This proves uniqueness.

Remark. As was pointed out to me by Dan Anderson, I incorrectly stated in a previous printing that if $\mathfrak{O}$ is a factorial complete local ring, then $\mathfrak{O}[[X]]$ is also factorial. This assertion is false, as shown by the example

$$k(t)[[X_1, X_2, X_3]]/(X_1^2 + X_2^2 + X_3^2)$$

due to P. Salmon, *Su un problema posto da P. Samuel*, Atti Acad. Naz. Lincei Rend. Cl. Sc. Fis. Matem. **40(8)** (1966) pp. 801–803. It is true that if $\mathfrak{O}$ is a regular local ring *in addition* to being complete, then $\mathfrak{O}[[X]]$ is factorial, but this is a deeper theorem. The simple proof I gave for the power series over a field is classical. I chose the exposition in [GrH 78].

Theorem 9.4. *If A is Noetherian, then $A[[X]]$ is also Noetherian.*

Proof. Our argument will be a modification of the argument used in the proof of Hilbert's theorem for polynomials. We shall consider elements of lowest degree instead of elements of highest degree.

Let $\mathfrak{A}$ be an ideal of $A[[X]]$. We let $\mathfrak{a}_i$ be the set of elements $a \in A$ such that a is the coefficient of X^i in a power series

$$aX^i + \text{terms of higher degree}$$

lying in $\mathfrak{A}$. Then $\mathfrak{a}_i$ is an ideal of A, and $\mathfrak{a}_i \subset \mathfrak{a}_{i+1}$ (the proof of this assertion being the same as for polynomials). The ascending chain of ideals stops:

$$\mathfrak{a}_0 \subset \mathfrak{a}_1 \subset \mathfrak{a}_2 \subset \cdots \subset \mathfrak{a}_r = \mathfrak{a}_{r+1} = \cdots$$

As before, let a_{ij} ($i = 0, \ldots, r$ and $j = 1, \ldots, n_i$) be generators for the ideals $\mathfrak{a}_i$, and let f_{ij} be power series in A having a_{ij} as beginning coefficient. Given $f \in \mathfrak{A}$, starting with a term of degree d, say $d \leqq r$, we can find elements $c_1, \ldots, c_{n_d} \in A$ such that

$$f - c_1 f_{d1} - \cdots - c_{n_d} f_{dn_d}$$

starts with a term of degree $\geqq d + 1$. Proceeding inductively, we may assume that $d > r$. We then use a linear combination

$$f - c_1^{(d)} X^{d-r} f_{r1} - \cdots - c_{n_r}^{(d)} X^{d-r} f_{rn_r}$$

to get a power series starting with a term of degree $\geqq d + 1$. In this way, if we start with a power series of degree $d > r$, then it can be expressed as a linear combination of $f_{r1}, \ldots, f_{rn_r}$ by means of the coefficients

$$g_1(X) = \sum_{\nu=d}^{\infty} c_1^{(\nu)} X^{\nu-r}, \ldots, g_{n_r}(X) = \sum_{\nu=d}^{\infty} c_{n_r}^{(\nu)} X^{\nu-r},$$

and we see that the f_{ij} generate our ideal $\mathfrak{A}$, as was to be shown.

Corollary 9.5. *If A is a Noetherian commutative ring, or a field, then $A[[X_1, \ldots, X_n]]$ is Noetherian.*

Examples. Power series in one variable are at the core of the theory of functions of one complex variable, and similarly for power series in several variables in the higher-dimensional case. See for instance [Gu 90].

Weierstrass polynomials occur in several contexts. First, they can be used to reduce questions about power series to questions about polynomials, in studying analytic sets. See for instance [GrH 78], Chapter 0. In a number-

theoretic context, such polynomials occur as characteristic polynomials in the Iwasawa theory of cyclotomic fields. Cf. [La 90], starting with Chapter 5.

Power series can also be used as generating functions. Suppose that to each positive integer n we associate a number $a(n)$. Then the **generating function** is the power series $\sum a(n)t^n$. In significant cases, it turns out that this function represents a rational function, and it may be a major result to prove that this is so.

For instance in Chapter X, §6 we shall consider a Poincaré series, associated with the length of modules. Similarly, in topology, consider a topological space X such that its homology groups (say) are finite dimensional over a field k of coefficients. Let $h_n = \dim H_n(X, k)$, where H_n is the n-th homology group. The **Poincaré series** is defined to be the generating series

$$P_X(t) = \sum h_n t^n.$$

Examples arise in the theory of dynamical systems. One considers a mapping $T: X \to X$ from a space X into itself, and we let N_n be the number of fixed points of the n-th iterate $T^n = T \circ T \circ \cdots \circ T$ (n times). The generating function is $\sum N_n t^n$. Because of the number of references I give here, I list them systematically at the end of the section. See first Artin–Mazur [ArM 65]; a proof by Manning of a conjecture of Smale [Ma 71]; and Shub's book [Sh 87], especially Chapter 10, Corollary 10.42 (Manning's theorem).

For an example in algebraic geometry, let V be an algebraic variety defined over a finite field k. Let K_n be the extension of k of degree n (in a given algebraic closure). Let N_n be the number of points of V in K_n. One defines the **zeta function** $Z(t)$ as the power series such that $Z(0) = 1$ and

$$Z'/Z(t) = \sum_{n=1}^{\infty} N_n t^{n-1}.$$

Then $Z(t)$ is a rational function (F. K. Schmidt when the dimension of V is 1, and Dwork in higher dimensions). For a discussion and references to the literature, see Appendix C of Hartshorne [Ha 77].

Finally we mention the **partition function** $p(n)$, which is the number of ways a positive integer can be expressed as a sum of positive integers. The generating function was determined by Euler to be

$$1 + \sum_{n=1}^{\infty} p(n)t^n = \prod_{n=1}^{\infty} (1 - t^n)^{-1}.$$

See for instance Hardy and Wright [HardW 71], Chapter XIX. The generating series for the partition function is related to the power series usually expressed in terms of a variable q, namely

$$\Delta = q \prod_{n=1}^{\infty} (1 - q^n)^{24} = \sum_{n=1}^{\infty} \tau(n) q^n,$$

which is the generating series for the **Ramanujan function** $\tau(n)$. The power series for Δ is also the expansion of a function in the theory of modular functions. For an introduction, see Serre's book [Se 73], last chapter, and books on elliptic functions, e.g. mine. We shall mention one application of the power series for Δ in the Galois theory chapter.

Generating power series also occur in K-theory, topological and algebraic geometric, as in Hirzebruch's formalism for the Riemann–Roch theorem and its extension by Grothendieck. See Atiyah [At 67], Hirzebruch [Hi 66], and [FuL 86]. I have extracted some formal elementary aspects having directly to do with power series in Exercises 21–27, which can be viewed as basic examples. See also Exercises 31–34 of the next chapter.

Bibliography

[ArM 65] M. ARTIN and B. MAZUR, On periodic points, *Ann. Math.* (2) **81** (1965) pp. 89–99

[At 67] M. ATIYAH, *K-Theory*, Addison-Wesley 1991 (reprinted from the Benjamin Lecture Notes, 1967)

[FuL 85] W. FULTON and S. LANG, *Riemann–Roch Algebra*, Springer-Verlag, New York, 1985

[GrH 78] P. GRIFFITHS and J. HARRIS, *Principles of Algebraic Geometry*, Wiley–Interscience, New York, 1978

[Gu 90] R. GUNNING, *Introduction to Holomorphic Functions of Several Variables*, Vol. II: *Local Theory*, Wadsworth and Brooks/Cole, 1990

[HardW 71] G. H. HARDY and E. M. WRIGHT, *An Introduction to the Theory of Numbers*, Oxford University Press, Oxford, UK, 1938–1971 (several editions)

[Hart 77] R. HARTSHORNE, *Algebraic Geometry*, Springer-Verlag, New York, 1977

[Hi 66] F. HIRZEBRUCH, *Topological Methods in Algebraic Geometry*, Springer-Verlag, New York, 1966 (translated and expanded from the original German, 1956)

[La 90] S. LANG, *Cyclotomic Fields*, I and II, Springer-Verlag, New York, 1990, combined edition of the original editions, 1978, 1980

[Ma 71] A. MANNING, Axiom A diffeomorphisms have rational zeta functions, *Bull. Lond. Math. Soc.* **3** (1971) pp. 215–220

[Se 73] J. P. SERRE, *A Course in Arithmetic*, Springer-Verlag, New York, 1973

[Sh 87] M. SHUB, *Global Stability of Dynamical Systems*, Springer-Verlag, New York, 1987

EXERCISES

1. Let k be a field and $f(X) \in k[X]$ a non-zero polynomial. Show that the following conditions are equivalent:
   (a) The ideal $(f(X))$ is prime.
   (b) The ideal $(f(X))$ is maximal.
   (c) $f(X)$ is irreducible.

2. (a) State and prove the analogue of Theorem 5.2 for the rational numbers.
   (b) State and prove the analogue of Theorem 5.3 for positive integers.

3. Let f be a polynomial in one variable over a field k. Let X, Y be two variables. Show that in $k[X, Y]$ we have a "Taylor series" expansion

$$f(X + Y) = f(X) + \sum_{i=1}^{n} \varphi_i(X) Y^i,$$

   where $\varphi_i(X)$ is a polynomial in X with coefficients in k. If k has characteristic 0, show that

$$\varphi_i(X) = \frac{D^i f(X)}{i!}.$$

4. Generalize the preceding exercise to polynomials in several variables (introduce partial derivatives and show that a finite Taylor expansion exists for a polynomial in several variables).

5. (a) Show that the polynomials $X^4 + 1$ and $X^6 + X^3 + 1$ are irreducible over the rational numbers.
   (b) Show that a polynomial of degree 3 over a field is either irreducible or has a root in the field. Is $X^3 - 5X^2 + 1$ irreducible over the rational numbers?
   (c) Show that the polynomial in two variables $X^2 + Y^2 - 1$ is irreducible over the rational numbers. Is it irreducible over the complex numbers?

6. Prove the integral root test of §3.

7. (a) Let k be a finite field with $q = p^m$ elements. Let $f(X_1, \ldots, X_n)$ be a polynomial in $k[X]$ of degree d and assume $f(0, \ldots, 0) = 0$. An element $(a_1, \ldots, a_n) \in k^{(n)}$ such that $f(a) = 0$ is called a zero of f. If $n > d$, show that f has at least one other zero in $k^{(n)}$. [*Hint*: Assume the contrary, and compare the degrees of the reduced polynomial belonging to

$$1 - f(X)^{q-1}$$

   and $(1 - X_1^{q-1}) \cdots (1 - X_n^{q-1})$. The theorem is due to Chevalley.]
   (b) Refine the above results by proving that the number N of zeros of f in $k^{(n)}$ is $\equiv 0 \pmod{p}$, arguing as follows. Let i be an integer $\geqq 1$. Show that

$$\sum_{x \in k} x^i = \begin{cases} q - 1 = -1 & \text{if } q - 1 \text{ divides } i, \\ 0 & \text{otherwise.} \end{cases}$$

   Denote the preceding function of i by $\psi(i)$. Show that

$$N \equiv \sum_{x \in k^{(n)}} (1 - f(x)^{q-1})$$

and for each n-tuple $(i_1, \ldots, i_n)$ of integers $\geqq 0$ that

$$\sum_{x \in k^{(n)}} x_1^{i_1} \cdots x_n^{i_n} = \psi(i_1) \cdots \psi(i_n).$$

Show that both terms in the sum for N above yield $0 \bmod p$. (The above argument is due to Warning.)

(c) Extend Chevalley's theorem to r polynomials $f_1, \ldots, f_r$ of degrees $d_1, \ldots, d_r$ respectively, in n variables. If they have no constant term and $n > \sum d_i$, show that they have a non-trivial common zero.

(d) Show that an arbitrary function $f: k^{(n)} \to k$ can be represented by a polynomial. (As before, k is a finite field.)

8. Let A be a commutative entire ring and X a variable over A. Let $a, b \in A$ and assume that a is a unit in A. Show that the map $X \mapsto aX + b$ extends to a unique automorphism of $A[X]$ inducing the identity on A. What is the inverse automorphism?

9. Show that every automorphism of $A[X]$ inducing the identity on A is of the type described in Exercise 8.

10. Let K be a field, and $K(X)$ the quotient field of $K[X]$. Show that every automorphism of $K(X)$ which induces the identity on K is of type

$$X \mapsto \frac{aX + b}{cX + d}$$

with $a, b, c, d \in K$ such that $(aX + b)/(cX + d)$ is not an element of K, or equivalently, $ad - bc \neq 0$.

11. Let A be a commutative entire ring and let K be its quotient field. We show here that some formulas from calculus have a purely algebraic setting. Let $D: A \to A$ be a **derivation**, that is an additive homomorphism satisfying the rule for the derivative of a product, namely

$$D(xy) = xDy + yDx \qquad \text{for} \quad x, y \in A.$$

(a) Prove that D has a unique extension to a derivation of K into itself, and that this extension satisfies the rule

$$D(x/y) = \frac{yDx - xDy}{y^2}$$

for $x, y \in A$ and $y \neq 0$. [Define the extension by this formula, prove that it is independent of the choice of x, y to write the fraction x/y, and show that it is a derivation having the original value on elements of A.]

(b) Let $L(x) = Dx/x$ for $x \in K^*$. Show that $L(xy) = L(x) + L(y)$. The homomorphism L is called the **logarithmic derivative**.

(c) Let D be the standard derivative in the polynomial ring $k[X]$ over a field k. Let $R(X) = c \prod (X - \alpha_i)^{m_i}$ with $\alpha_i \in k$, $c \in k$, and $m_i \in \mathbf{Z}$, so $R(X)$ is a rational

function. Show that

$$R'/R = \sum \frac{m_i}{X - \alpha_i}.$$

12. (a) If $f(X) = aX^2 + bX + c$, show that the discriminant of f is $b^2 - 4ac$.
    (b) If $f(X) = a_0X^3 + a_1X^2 + a_2X + a_3$, show that the discriminant of f is

$$a_1^2a_2^2 - 4a_0a_2^3 - 4a_1^3a_3 - 27a_0^2a_3^2 + 18a_0a_1a_2a_3.$$

    (c) Let $f(X) = (X - t_1) \cdots (X - t_n)$. Show that

$$D_f = (-1)^{n(n-1)/2} \prod_{i=1}^{n} f'(t_i).$$

13. Polynomials will be taken over an algebraically closed field of characteristic 0.
    (a) Prove

Davenport's theorem. *Let $f(t)$, $g(t)$ be polynomials such that $f^3 - g^2 \neq 0$. Then*

$$\deg(f^3 - g^2) \geqq \tfrac{1}{2} \deg f + 1.$$

Or put another way, let $h = f^3 - g^2$ and assume $h \neq 0$. Then

$$\deg f \leqq 2 \deg h - 2.$$

To do this, first assume f, g relatively prime and apply Mason's theorem. In general, proceed as follows.

    (b) *Let A, B, f, g be polynomials such that Af, Bg are relatively prime $\neq 0$. Let $h = Af^3 + Bg^2$. Then*

$$\deg f \leqq \deg A + \deg B + 2 \deg h - 2.$$

This follows directly from Mason's theorem. Then starting with f, g not necessarily relatively prime, start factoring out common factors until no longer possible, to effect the desired reduction. When I did it, I needed to do this step three times, so don't stop until you get it.

    (c) Generalize (b) to the case of $f^m - g^n$ for arbitrary positive integer exponents m and n.

14. Prove that the generalized Szpiro conjecture implies the *abc* conjecture.

15. Prove that the *abc* conjecture implies the following conjecture: There are infinitely many primes p such that $2^{p-1} \not\equiv 1 \bmod p^2$. [Cf. the reference [Sil 88] and [La 90] at the end of §7.]

16. Let w be a complex number, and let $c = \max(1, |w|)$. Let F, G be non-zero polynomials in one variable with complex coefficients, of degrees d and d' respectively, such that $|F|, |G| \geqq 1$. Let R be their resultant. Then

$$|R| \leqq c^{d+d'}[|F(w)| + |G(w)|]|F|^{d'}|G|^{d}(d + d')^{d+d'}.$$

(We denote by $|F|$ the maximum of the absolute values of the coefficients of F.)

17. Let d be an integer $\geqq 3$. Prove the existence of an irreducible polynomial of degree d over $\mathbf{Q}$, having precisely $d - 2$ real roots, and a pair of complex conjugate roots. Use the following construction. Let $b_1, \ldots, b_{d-2}$ be distinct

integers, and let a be an integer > 0. Let

$$g(X) = (X^2 + a)(X - b_1) \cdots (X - b_{d-2}) = X^d + c_{d-1}X^{d-1} + \cdots + c_0.$$

Observe that $c_i \in \mathbf{Z}$ for all i. Let p be a prime number, and let

$$g_n(X) = g(X) + \frac{p}{p^{dn}}$$

so that g_n converges to g (i.e. the coefficients of g_n converge to the coefficients of g).

(a) Prove that g_n has precisely $d - 2$ real roots for n sufficiently large. (You may use a bit of calculus, or use whatever method you want.)

(b) Prove that g_n is irreducible over $\mathbf{Q}$.

Integral-valued polynomials

18. Let $P(X) \in \mathbf{Q}[X]$ be a polynomial in one variable with rational coefficients. It may happen that $P(n) \in \mathbf{Z}$ for all sufficiently large integers n without necessarily P having integer coefficients.

(a) Give an example of this.

(b) Assume that P has the above property. Prove that there are integers $c_0, c_1, \ldots, c_r$ such that

$$P(X) = c_0\binom{X}{r} + c_1\binom{X}{r-1} + \cdots + c_r,$$

where

$$\binom{X}{r} = \frac{1}{r!}X(X-1)\cdots(X-r+1)$$

is the binomial coefficient function. In particular, $P(n) \in \mathbf{Z}$ for all n. Thus we may call P **integral valued**.

(c) Let $f: \mathbf{Z} \to \mathbf{Z}$ be a function. Assume that there exists an integral valued polynomial Q such that the difference function Δf defined by

$$(\Delta f)(n) = f(n) - f(n-1)$$

is equal to $Q(n)$ for all n sufficiently large positive. Show that there exists an integral-valued polynomial P such that $f(n) = P(n)$ for all n sufficiently large.

Exercises on symmetric functions

19. (a) Let $X_1, \ldots, X_n$ be variables. Show that any homogeneous polynomial in $\mathbf{Z}[X_1, \ldots, X_n]$ of degree $> n(n-1)$ lies in the ideal generated by the elementary symmetric functions $s_1, \ldots, s_n$.

(b) With the same notation show that $\mathbf{Z}[X_1, \ldots, X_n]$ is a free $\mathbf{Z}[s_1, \ldots, s_n]$ module with basis the monomials

$$X^{(r)} = X_1^{r_1} \cdots X_n^{r_n}$$

with $0 \leqq r_i \leqq n - i$.

(c) Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ be two independent sets of variables. Let $s_1, \ldots, s_n$ be the elementary symmetric functions of X and $s'_1, \ldots, s'_m$ the elementary symmetric functions of Y (using vector vector notation). Show that $\mathbf{Z}[X, Y]$ is free over $\mathbf{Z}[s, s']$ with basis $X^{(r)}Y^{(q)}$, and the exponents (r), (q) satisfying inequalities as in (b).

(d) Let I be an ideal in $\mathbf{Z}[s, s']$. Let J be the ideal generated by I in $\mathbf{Z}[X, Y]$. Show that

$$J \cap \mathbf{Z}[s, s'] = I.$$

20. Let A be a commutative ring. Let t be a variable. Let

$$f(t) = \sum_{i=0}^{m} a_i t^i \qquad \text{and} \qquad g(t) = \sum_{i=0}^{n} b_i t^i$$

be polynomials whose constant terms are $a_0 = b_0 = 1$. If

$$f(t)g(t) = 1,$$

show that there exists an integer N $\big(= (m + n)(m + n - 1)\big)$ such that any monomial

$$a_1^{r_1} \cdots a_n^{r_n}$$

with $\sum jr_j > N$ is equal to 0. [*Hint*: Replace the a's and b's by variables. Use Exercise 19(b) to show that any monomial $M(a)$ of weight $> N$ lies in the ideal I generated by the elements

$$c_k = \sum_{i=0}^{k} a_i b_{k-i}$$

(letting $a_0 = b_0 = 1$). Note that c_k is the k-th elementary symmetric function of the $m + n$ variables (X, Y).]

[*Note*: For some interesting contexts involving symmetric functions, see Cartier's talk at the Bourbaki Seminar, 1982–1983.]

λ-rings

The following exercises start a train of thought which will be pursued in Exercise 33 of Chapter V; Exercises 22–24 of Chapter XVIII; and Chapter XX, §3. These originated to a large extent in Hirzebruch's Riemann–Roch theorem and its extension by Grothendieck who defined λ-rings in general.

Let K be a commutative ring. By **λ-operations** we mean a family of mappings

$$\lambda^i : K \to K$$

for each integer $i \geqq 0$ satisfying the relations for all $x \in K$:

$$\lambda^0(x) = 1, \qquad \lambda^1(x) = x,$$

and for all integers $n \geqq 0$, and $x, y \in K$,

$$\lambda^n(x + y) = \sum_{i=0}^{n} \lambda^i(x)\lambda^{n-i}(y).$$

The reader will meet examples of such operations in the chapter on the alternating and symmetric products, but the formalism of such operations depends only on the above relations, and so can be developed here in the context of formal power series. Given a λ-operation, in which case we also say that K is a **λ-ring**, we define the power series

$$\lambda_t(x) = \sum_{i=0}^{\infty} \lambda^i(x)t^i.$$

Prove the following statements.

21. The map $x \mapsto \lambda_t(x)$ is a homomorphism from the additive group of K into the multiplicative group of power series $1 + tK[[t]]$ whose constant term is equal to 1. Conversely, any such homomorphism such that $\lambda_t(x) = 1 + xt +$ higher terms gives rise to λ-operations.

22. Let $s = at +$ higher terms be a power series in $K[[t]]$ such that a is a unit in K. Show that there is a power series

$$t = g(s) = \sum b_i s^i \qquad \text{with} \quad b_i \in K.$$

Show that any power series $f(t) \in K[[t]]$ can be written in the form $h(s)$ for some other power series with coefficients in K.

Given a λ-operation on K, define the corresponding **Grothendieck power series**

$$\gamma_t(x) = \lambda_{t/(1-t)}(x) = \lambda_s(x)$$

where $s = t/(1-t)$. Then the map

$$x \mapsto \gamma_t(x)$$

is a homomorphism as before. We define $\gamma^i(x)$ by the relation

$$\gamma_t(x) = \sum \gamma^i(x)t^i.$$

Show that γ satisfies the following properties.

23. (a) For every integer $n \geqq 0$ we have

$$\gamma^n(x + y) = \sum_{i=0}^{n} \gamma^i(x)\gamma^{n-i}(y).$$

(b) $\gamma_t(1) = 1/(1-t)$.

(c) $\gamma_t(-1) = 1 - t$.

24. Assume that $\lambda^i u = 0$ for $i > 1$. Show:

(a) $\gamma_t(u - 1) = 1 + (u-1)t$.

(b) $\gamma_t(1 - u) = \sum_{i=0}^{\infty} (1-u)^i t^i$.

25. **Bernoulli numbers.** Define the Bernoulli numbers B_k as the coefficients in the power series

$$F(t) = \frac{t}{e^t - 1} = \sum_{k=0}^{\infty} B_k \frac{t^k}{k!}.$$

Of course, $e^t = \sum t^n/n!$ is the standard power series with rational coefficients $1/n!$. Prove:

(a) $B_0 = 1$, $B_1 = -\frac{1}{2}$, $B_2 = \frac{1}{6}$.

(b) $F(-t) = t + F(t)$, and $B_k = 0$ if k is odd $\neq 1$.

26. **Bernoulli polynomials.** Define the Bernoulli polynomials $\mathbf{B}_k(X)$ by the power series expansion

$$F(t, X) = \frac{te^{tX}}{e^t - 1} = \sum_{k=0}^{\infty} \mathbf{B}_k(X)\frac{t^k}{k!}.$$

It is clear that $B_k = \mathbf{B}_k(0)$, so the Bernoulli numbers are the constant terms of the Bernoulli polynomials. Prove:

(a) $\mathbf{B}_0(X) = 1$, $\mathbf{B}_1(X) = X - \frac{1}{2}$, $\mathbf{B}_2(X) = X^2 - X + \frac{1}{6}$.

(b) For each positive integer N,

$$\mathbf{B}_k(X) = N^{k-1} \sum_{a=0}^{N-1} \mathbf{B}_k\left(\frac{X + a}{N}\right).$$

(c) $\mathbf{B}_k(X) = X^k - \frac{1}{2}kX^{k-1} +$ lower terms.

(d) $F(t, X + 1) - F(t, X) = te^{Xt} = t\sum X^k \dfrac{t^k}{k!}$.

(e) $\mathbf{B}_k(X + 1) - \mathbf{B}_k(X) = kX^{k-1}$ for $k \geqq 1$.

27. Let N be a positive integer and let f be a function on $\mathbf{Z}/N\mathbf{Z}$. Form the power series

$$F_f(t, X) = \sum_{a=0}^{N-1} f(a)\frac{te^{(a+X)t}}{e^{Nt} - 1}.$$

Following Leopoldt, define the **generalized Bernoulli polynomials** relative to the function f by

$$F_f(t, X) = \sum_{k=0}^{\infty} \mathbf{B}_{k,f}(X)\frac{t^k}{k!}.$$

In particular, the constant term of $\mathbf{B}_{k,f}(X)$ is defined to be the **generalized Bernoulli number** $B_{k,f} = \mathbf{B}_{k,f}(0)$ introduced by Leopoldt in cyclotomic fields. Prove:

(a) $F_f(t, X + k) = e^{kt}F_f(t, X)$.

(b) $F_f(t, X + N) - F_f(t, X) = (e^{Nt} - 1)F_f(t, X)$.

(c) $\dfrac{1}{k}[\mathbf{B}_{k,f}(X + N) - \mathbf{B}_{k,f}(X)] = \sum_{a=0}^{N-1} f(a)(a + X)^{k-1}$.

(d) $\mathbf{B}_{k,f}(X) = \sum_{i=0}^{k} \binom{k}{i} B_{i,f}X^{n-i}$

$$= B_{k,f} + kB_{k-1,f}X + \cdots + kB_{1,f}X^{k-1} + B_{0,f}X^k.$$

Note. The exercises on Bernoulli numbers and polynomials are designed not only to give examples for the material in the text, but to show how this material leads into major areas of mathematics: in topology and algebraic geometry centering

around Riemann–Roch theorems; analytic and algebraic number theory, as in the theory of the zeta functions and the theory of modular forms, cf. my *Introduction to Modular Forms*, Springer-Verlag, New York, 1976, Chapters XIV and XV; my *Cyclotomic Fields*, I and II, Springer-Verlag, New York, 1990, Chapter 2, §2; Kubert–Lang's *Modular Units*, Springer-Verlag, New York, 1981; etc.

Further Comments, 1996–2001. I was informed by Umberto Zannier that what has been called Mason's theorem was proved three years earlier by Stothers [Sto 81], Theorem 1.1. Zannier himself has published some results on Davenport's theorem [Za 95], without knowing of the paper by Stothers, using a method similar to that of Stothers, and rediscovering some of Stothers' results, but also going beyond. Indeed, Stothers uses the "Belyi method" belonging to algebraic geometry, and increasingly appearing as a fundamental tool. Mason gave a very elementary proof, accessible at the basic level of algebra. An even shorter and very elegant proof of the Mason-Stothers theorem was given by Noah Snyder [Sny 00]. I am much indebted to Snyder for showing me that proof before publication, and I reproduced it in [La 99b]. But I recommend looking at Snyder's version.

[La 99b] S. LANG, *Math Talks for Undergraduates*, Springer Verlag 1999

[Sny 00] N. SNYDER, An alternate proof of Mason's theorem, *Elemente der Math.* **55** (2000) pp. 93–94

[Sto 81] W. STOTHERS, Polynomial identities and hauptmoduln, *Quart. J. Math. Oxford* (2) **32** (1981) pp. 349–370

[Za 95] U. ZANNIER, On Davenport's bound for the degree of $f^3 - g^2$ and Riemann's existence theorem, *Acta Arithm.* **LXXI.2** (1995) pp. 107–137