7

Formal Power Series

In the last chapter, we introduced techniques for computing generating functions $G_S(x) = \sum_{z \in S} x^{\mathrm{wt}(z)}$ where S is a *finite* set of weighted objects. These generating functions are polynomials in the variable x. Now suppose that S is an *infinite* set of weighted objects. By analogy with the finite case, we would like to define a generating function $G_S(x) = \sum_{z \in S} x^{\mathrm{wt}(z)} = \sum_{n \geq 0} a_n x^n$, where a_n is the number of objects in S of weight n. But the resulting expression $G_S(x)$ is no longer a polynomial in x, since a polynomial can have only finitely many terms.

For example, if S is the set of all words over the alphabet $\{0, 1\}$ weighted by length, we have $a_n = 2^n$ for all $n \geq 0$, and so

$$G_S(x) = 1 + 2x + 4x^2 + 8x^3 + \cdots + 2^n x^n + \cdots = \sum_{n=0}^{\infty} 2^n x^n.$$

If we think of x as a real number, then $G_S(x)$ is a *function of a real variable* that is defined for all x sufficiently close to zero. In fact, using the geometric series formula from calculus, one sees that $G_S(x) = \sum_{n \geq 0} (2x)^n = \frac{1}{1-2x}$ for $-1/2 < x < 1/2$. For values of x outside this interval, $G_S(x)$ is undefined. More generally, the power series $\sum_{n=0}^{\infty} a_n x^n$ can be regarded as a function of a real (or complex) variable x that is defined within a certain interval of convergence centered at $x = 0$. However, difficulties can emerge if the coefficients a_n grow too rapidly. For instance, given $a_n = n!$ for all n, one can show using the ratio test that the power series $H(x) = \sum_{n=0}^{\infty} n! x^n$ only converges at $x = 0$. Thus we cannot recover the coefficients $a_n = n!$ from knowledge of the function H, which is only defined at $x = 0$.

As this example shows, using real-valued functions to model combinatorial generating functions can be problematic because one must constantly worry about questions of convergence. We would prefer a purely *formal* notion of a power series in which convergence issues do not arise. The idea is to view a generating function $\sum_{n=0}^{\infty} a_n x^n$ as merely a shorthand for an infinite sequence of integers $(a_0, a_1, a_2, \ldots, a_n, \ldots)$. The letter x is now only a symbol, not a variable; we are *not* allowed to substitute specific real numbers for x.

This chapter gives a rigorous development of the algebraic properties of formal power series. Our goal is to extend the familiar operations on polynomial functions (like addition, multiplication, composition, and differentiation) to the setting of formal power series. In certain situations, we will even be able to define infinite sums and products of formal power series. These algebraic operations will be used to help develop the combinatorics of infinite weighted sets, which is the topic of the next chapter.

In combinatorics, it usually suffices to consider formal power series whose coefficients are integers, rational numbers, or complex numbers. In this chapter, we will consider the slightly more general situation where the coefficients come from any field of characteristic zero (cf. 7.1 below). In fact, much of the algebraic theory is valid for power series with coefficients coming from an arbitrary ring (see 2.2). We shall indicate, as we proceed, which proofs require the stronger assumptions we are imposing on the coefficient ring.

7.1 The Ring of Formal Power Series

7.1. Notational Convention. Throughout this chapter, the letter K will stand for a field (see 2.3) that contains the field $\mathbb{Q}$ of rational numbers.

For example, K might be $\mathbb{Q}$ itself, or $\mathbb{R}$ (the field of real numbers), or $\mathbb{C}$ (the field of complex numbers). K might also be a field $\mathbb{Q}(x)$ of formal rational functions, discussed in 7.46 below.

7.2. Definition: Formal Power Series. A *formal power series in one variable with coefficients in* K is a function $F : \mathbb{N} \to K$. We write $F(n)$ or F_n for the value of the function F on the input $n \in \mathbb{N}$. The set of all such functions will be denoted $K[[x]]$, where x is a symbol called an *indeterminate*.

A formal power series $F \in K[[x]]$ is exactly the same as a *sequence*

$$F = (F_0, F_1, F_2, \ldots, F_n, \ldots) = (F(0), F(1), F(2), \ldots, F(n), \ldots)$$

indexed by nonnegative integers, where each $F_n \in K$. We often display this sequence using *power series notation*, writing

$$F = \sum_{n=0}^{\infty} F_n x^n$$

and calling F_n *the coefficient of* x^n *in* F. For the time being, the symbol x appearing in this notation has no independent meaning, and there are no addition, multiplication, or exponentiation operations being performed on the right side. This notation is used to help motivate the algebraic operations on power series to be introduced below, which are suggested by corresponding operations on one-variable polynomials.

7.3. Remark: Equality of Formal Power Series. Two formal power series $F, G \in K[[x]]$ are *equal* iff $F_n = G_n$ for all $n \in \mathbb{N}$. This follows from the definition of equality of two functions with domain $\mathbb{N}$.

7.4. Example. Consider the functions $G, H : \mathbb{N} \to \mathbb{Q}$ defined by $G(n) = 2^n$ and $H(n) = n!$ for all $n \in \mathbb{N}$. These are two elements of $\mathbb{Q}[[x]]$ which were discussed in the introduction to this chapter. In sequence notation and power series notation, we would write

$$G_S = (1, 2, 4, 8, \ldots, 2^n, \ldots) = \sum_{n \geq 0} 2^n x^n;$$

$$H = (1, 1, 2, 6, 24, 120, 720, \ldots, n!, \ldots) = \sum_{n \geq 0} n! x^n.$$

The function $Z : \mathbb{N} \to K$ such that $Z(n) = 0_K$ for all $n \in \mathbb{N}$ defines a *zero power series* $Z = \sum_{n \geq 0} 0x^n$. We often denote Z (as well as the additive identity of K, the integer zero, etc.) by the symbol 0.

7.5. Example: X_i**.** For each $i \in \mathbb{N}$, define a power series $X_i : \mathbb{N} \to K$ by setting $X_i(i) = 1$ and $X_i(j) = 0$ for all $j \neq i$. Thus X_i is the sequence $(0, 0, \ldots, 1, 0, \ldots)$ where the 1 is preceded by i zeroes. We have

$$X_i = \sum_{n=0}^{\infty} \chi(n = i) x^n.$$

If we omit zero coefficients, it is tempting to write $X_0 = x^0 = 1$, $X_1 = x^1 = x$, and $X_i = x^i$. Strictly speaking, these abbreviations of the official power series notation are not allowed, but soon we will find a way to justify them.

We can now define addition and multiplication of formal power series.

7.6. Definition: Sum and Product of Formal Power Series. Given $F, G \in K[[x]]$, define the *sum* $F + G : \mathbb{N} \to K$ by $(F + G)(n) = F(n) + G(n)$ for all $n \in \mathbb{N}$. Define the *product* $FG : \mathbb{N} \to K$ by

$$(FG)(n) = \sum_{\substack{(i,j)\in\mathbb{N}^2 \\ i+j=n}} F(i)G(j) = \sum_{k=0}^{n} F(k)G(n-k).$$

FG is sometimes called the *convolution* of the functions F and G.

In sequence notation, this definition says

$$(F_n : n \geq 0) + (G_n : n \geq 0) = (F_n + G_n : n \geq 0);$$

$$(F_n : n \geq 0) \times (G_n : n \geq 0) = \left(\sum_{k=0}^{n} F_k G_{n-k} : n \geq 0 \right).$$

Using formal power series notation, these operations can also be written

$$\sum_{n\geq 0} F_n x^n + \sum_{n\geq 0}^{\infty} G_n x^n = \sum_{n\geq 0} (F_n + G_n) x^n;$$

$$\left(\sum_{n\geq 0} F_n x^n \right) \times \left(\sum_{n\geq 0} G_n x^n \right) = \sum_{n\geq 0} \left(\sum_{i+j=n} F_i G_j \right) x^n.$$

These formulas are exactly what we would expect (using the generalized distributive law) if x and every F_n and G_n were elements in some ring, and the sums appearing were finite.

7.7. Example. In $\mathbb{Q}[[x]]$, we have

$$(1, 2, 3, 4, 5, 6, \ldots) + (1, 0, 1, 0, 1, 0, 1, 0, 1, 0, \ldots) = (2, 2, 4, 4, 6, 6, \ldots);$$

$$(1, 2, 3, 4, 5, 6, \ldots) \times (1, 0, 1, 0, 1, 0, 1, 0, 1, 0, \ldots) = (1, 2, 4, 6, 9, 12, 16, 20, \ldots).$$

Given $A = (3, 0, 2, 1, 7, 0, 0, 0, \ldots)$ and $B = (1, 4, 5, 0, 0, \ldots)$ in $K[[x]]$, we have

$$A + B = (4, 4, 7, 1, 7, 0, 0, \ldots);$$

$$AB = (3 \cdot 1, 3 \cdot 4 + 0 \cdot 1, 3 \cdot 5 + 0 \cdot 4 + 2 \cdot 1, \ldots) = (3, 12, 17, 9, 21, 33, 35, 0, 0, \ldots).$$

Compare these formal operations to the ordinary product of the two polynomial functions $p(z) = 3 + 2z^2 + z^3 + 7z^4$ and $q(z) = 1 + 4z + 5z^2$:

$$p(z) + q(z) = 4 + 4z + 7z^2 + 1z^3 + 7z^4 \qquad (z \in \mathbb{R});$$

$$p(z)q(z) = 3 + 12z + 17z^2 + 9z^3 + 21z^4 + 33z^5 + 35z^6 \qquad (z \in \mathbb{R}).$$

Now suppose $F = (F_n : n \in \mathbb{N}) \in K[[x]]$ and $C = (1, 1, 1, \ldots) \in K[[x]]$. Then

$$FC = CF = (F_0, F_0 + F_1, F_0 + F_1 + F_2, \ldots, F_0 + F_1 + \cdots + F_n, \ldots).$$

Thus, multiplication by C replaces a sequence of scalars by the sequence of partial sums of those scalars.

7.8. Theorem: Algebraic Structure of $K[[x]]$. (a) With the sum and product operations defined above, $K[[x]]$ is a commutative ring. (b) $K[[x]]$ contains the field K, provided we identify each $a \in K$ with the sequence $(a, 0, 0, 0, \ldots) \in K[[x]]$. (c) $K[[x]]$ is a vector space over K.

Proof. (a) We verify some of the ring axioms for $K[[x]]$ (see 2.2), leaving the others as exercises. If F and G are functions from $\mathbb{N}$ to K, $F+G$ and FG are also functions that map $\mathbb{N}$ *into* K (using closure of K under addition and multiplication). In other words, $F \in K[[x]]$ and $G \in K[[x]]$ imply $F+G \in K[[x]]$ and $FG \in K[[x]]$, so the closure axioms for $K[[x]]$ are true. To see that addition in $K[[x]]$ is associative, fix $F, G, H \in K[[x]]$. Using associativity of addition in the field K, we see that

$$\begin{aligned}[(F+G)+H]_n &= (F+G)_n + H_n = (F_n + G_n) + H_n \\ &= F_n + (G_n + H_n) = F_n + (G+H)_n = [F + (G+H)]_n\end{aligned}$$

for every $n \in \mathbb{N}$. Thus (by 7.3) $(F+G)+H = F+(G+H)$.

The verification that $(FG)H = F(GH)$ is somewhat more elaborate. On one hand, for a fixed $n \in \mathbb{N}$,

$$[(FG)H]_n = \sum_{\substack{(i,c)\in\mathbb{N}^2:\\ i+c=n}} (FG)_i H_c = \sum_{\substack{(i,c)\in\mathbb{N}^2\\ i+c=n}} \left(\sum_{\substack{(a,b)\in\mathbb{N}^2\\ a+b=i}} (F_a G_b) \right) H_c = \sum_{\substack{(a,b,c)\in\mathbb{N}^3\\ (a+b)+c=n}} (F_a G_b) H_c.$$

The last step used the distributive law in K and a reindexing of the summations (which is permissible since addition in K is commutative). On the other hand,

$$[F(GH)]_n = \sum_{\substack{(a,k)\in\mathbb{N}^2:\\ a+k=n}} F_a (GH)_k = \sum_{\substack{(a,k)\in\mathbb{N}^2\\ a+k=n}} F_a \left(\sum_{\substack{(b,c)\in\mathbb{N}^2\\ b+c=k}} (G_b H_c) \right) = \sum_{\substack{(a,b,c)\in\mathbb{N}^3\\ a+(b+c)=n}} F_a (G_b H_c).$$

Using associativity of addition in $\mathbb{N}$ and associativity of multiplication in K, we see that $[(FG)H]_n = [F(GH)]_n$ for all $n \in \mathbb{N}$, hence $(FG)H = F(GH)$ as desired.

Next we claim that $X_0 = (1, 0, 0, \ldots) = \sum_{n\geq 0} \chi(n=0)x^n$ is the multiplicative identity element in $K[[x]]$. For, given $F \in K[[x]]$ and $n \in \mathbb{N}$, we compute

$$(X_0 F)_n = \sum_{i+j=n} X_0(i) F(j) = 1F(n) + 0F(n-1) + \cdots + 0F(0) = F_n.$$

Thus $X_0 F = F$, and similarly $FX_0 = F$. We let the reader verify the remaining ring axioms, namely: the zero sequence is the additive identity in $K[[x]]$; the additive inverse of $(F_n : n \geq 0)$ is $(-F_n : n \geq 0)$; addition in $K[[x]]$ is commutative; multiplication in $K[[x]]$ is commutative (this uses commutativity of K); and the distributive law holds.

For (b), observe first that the map $a \mapsto (a, 0, 0, \ldots)$ is a bijection of K onto the subset of $K[[x]]$ consisting of sequences that are zero after position zero. The definitions immediately show that

$$\begin{aligned}(a,0,0,\ldots) + (b,0,0,\ldots) &= (a+b,0,0,\ldots); \\ -(a,0,0,\ldots) &= (-a,0,0,\ldots); \\ (a,0,0,\ldots) \times (b,0,0,\ldots) &= (ab,0,0,\ldots);\end{aligned}$$

and furthermore, $0_K \mapsto (0,0,0,\ldots) = 0_{K[[x]]}$ and $1_K \mapsto (1,0,0,\ldots) = 1_{K[[x]]}$. This shows

that operations in $K[[x]]$ on sequences of this form agree with the corresponding field operations in K. So we can view K as embedded in $K[[x]]$ by means of this bijection. (More formally, we have found an "isomorphic copy" of K inside $K[[x]]$.)

(c) Define scalar multiplication in $K[[x]]$ by setting $cF = (cF_n : n \in \mathbb{N})$ for $c \in K$ and $F \in K[[x]]$. One checks that cF is the same as the product of $(c, 0, 0, \ldots)$ and $F = (F_0, F_1, F_2, \ldots)$ in the ring $K[[x]]$. Using this observation, one sees immediately that $K[[x]]$ satisfies all the axioms for a vector space over K, because the required identities are special cases of the ring axioms that have just been verified. □

7.2 Finite Products and Powers of Formal Series

Now that we know $K[[x]]$ is a ring, we can iterate the binary operations of addition and multiplication to define finite sums and products of formal power series. Similarly, for any integer $n \geq 0$ and any $G \in K[[x]]$, the power G^n is defined recursively by setting $G^0 = 1$ and, for $n \geq 0$, setting $G^{n+1} = G^n \cdot G$. Intuitively, G^n is the product of n factors all equal to G. Later, we will see that infinite sums and infinite products of formal power series can be defined in certain situations. We will also obtain a criterion for when the multiplicative inverse G^{-1} (and other negative powers of G) can be formed.

7.9. Example: Powers of x. For $i \in \mathbb{N}$, define $X_i = \sum_{n \geq 0} \chi(n = i)x^n$ as in 7.5. We claim that $X_1^i = X_i$ for all $i \geq 0$. The claim holds when $i = 0$ since $X_1^0 = 1_{K[[x]]}$ by definition, and we saw in 7.8 that $1_{K[[x]]} = X_0$. Fix $i \geq 0$ and assume by induction that $X_1^i = X_i$. Now

$$X_1^{i+1}(n) = (X_1^i \cdot X_1)(n) = (X_i \cdot X_1)(n) = \sum_{a+b=n} X_i(a)X_1(b) \qquad (n \in \mathbb{N}).$$

The only choice of (a, b) that produces a nonzero summand is $a = i$ and $b = 1$, which can occur only for $n = i + 1$. So $X_1^{i+1}(n) = 0 = X_{i+1}(n)$ if $n \neq i + 1$, and $X_1^{i+1}(i + 1) = 1 = X_{i+1}(i + 1)$. Thus $X_1^{i+1} = X_{i+1}$, verifying the claim for $i + 1$.

If we *define* x to be the particular formal power series $X_1 \in K[[x]]$, the claim shows that $x^i = X_i$ for all $i \geq 0$. We have now justified our earlier "notational abuse" in 7.5. Furthermore, for any *finite* sequence of scalars $c_0, c_1, \ldots, c_N \in K$, define $C \in K[[x]]$ by letting $C(n) = c_n$ for $0 \leq n \leq N$ and $C(n) = 0$ for $n > N$. Then the definition of addition and scalar multiplication shows that

$$c_0 + c_1x + c_2x^2 + \cdots + c_Nx^N = (c_0, c_1, \ldots, c_N, 0, 0, \ldots) = C = \sum_{n \geq 0} C_n x^n,$$

where the leftmost expression is built up from $x = X_1$ and the c_i's by algebraic operations in $K[[x]]$, and the rightmost expression is our atomic notation for the series C. Later, after we give a meaning to infinite summations of formal power series, we will see that the analogous identity

$$C_0 + C_1x + C_2x^2 + \cdots + C_nx^n + \cdots = C = \sum_{n \geq 0} C_n x^n$$

is also valid for any $C \in K[[x]]$.

7.10. Theorem: Products of k Series. Suppose $G_1, G_2, \ldots, G_k \in K[[x]]$. For all $n \in \mathbb{N}$,

$$(G_1G_2 \cdots G_k)(n) = \sum_{\substack{(j_1, j_2, \ldots, j_k) \in \mathbb{N}^k: \\ j_1 + j_2 + \cdots + j_k = n}} G_1(j_1)G_2(j_2) \cdots G_k(j_k).$$

Proof. We use induction on k. The case $k = 1$ is immediate, and the case $k = 2$ holds by the definition of the product of two formal power series. Assume $k > 2$ and the result is already known for products of $k-1$ series. Letting $F = G_1G_2\cdots G_{k-1}$, we calculate

$$\begin{aligned}(G_1G_2\cdots G_k)(n) &= (FG_k)(n) = \sum_{\substack{(r,s)\in\mathbb{N}^2\\ r+s=n}} F(r)G_k(s)\\ &= \sum_{\substack{(r,s)\in\mathbb{N}^2\\ r+s=n}} \left(\sum_{\substack{(j_1,j_2,\ldots,j_{k-1})\in\mathbb{N}^{k-1}:\\ j_1+j_2+\cdots+j_{k-1}=r}} G_1(j_1)G_2(j_2)\cdots G_{k-1}(j_{k-1}) \right) G_k(s)\\ &= \sum_{\substack{(j_1,j_2,\ldots,j_k)\in\mathbb{N}^k:\\ j_1+j_2+\cdots+j_k=n}} G_1(j_1)G_2(j_2)\cdots G_{k-1}(j_{k-1})G_k(j_k).\end{aligned}$$

The last step follows by the generalized distributive law and a change in the names of the summation indices. □

Now we can prove a result, similar to the multinomial theorem 2.12, that lets us compute a power of a formal series.

7.11. Theorem: Powers of Formal Series. For all $G \in K[[x]]$ and all $m, n \in \mathbb{N}$,

$$G^m(n) = \sum_{\substack{(k_0,k_1,\ldots,k_n)\in\mathbb{N}^{n+1}:\\ \sum_i k_i = m, \sum_i ik_i = n}} \binom{m}{k_0, k_1, \ldots, k_n} G(0)^{k_0}G(1)^{k_1}\cdots G(n)^{k_n}.$$

Proof. Applying 7.10 with $k = m$ and all G_i's equal to G, we see that

$$G^m(n) = \sum_{\substack{(j_1,\ldots,j_m)\in\mathbb{N}^m:\\ j_1+\cdots+j_m=n}} G(j_1)G(j_2)\cdots G(j_m).$$

Given $(k_0, k_1, \ldots, k_n) \in \mathbb{N}^{n+1}$ satisfying $\sum_i k_i = m$, $\sum_i ik_i = n$, let us group together all the summands indexed by sequences $(j_1, \ldots, j_m) \in \mathcal{R}(0^{k_0}1^{k_1}\cdots n^{k_n})$. The number of such summands is the multinomial coefficient $\binom{m}{k_0,k_1,\ldots,k_n}$ by 1.46, and (by commutativity of multiplication in K) every such summand is equal to $G(0)^{k_0}G(1)^{k_1}\cdots G(n)^{k_n}$. Summing over all possible choices of $(k_0, \ldots, k_n)$ gives the stated formula for $G^m(n)$. (Compare to the proof of 2.12.) □

7.3 Formal Polynomials

Polynomials, like the power series studied in calculus, are often regarded as functions of a real variable x. For a function $p : \mathbb{R} \to \mathbb{R}$ to be a polynomial, there must exist constants $a_i \in \mathbb{R}$ and $n \in \mathbb{N}$ such that $p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ for all $x \in \mathbb{R}$. One can prove that the coefficients a_i are uniquely determined by p. This functional view of polynomials is not really necessary for many algebraic and combinatorial purposes. We now give a rigorous discussion of "formal" polynomials and their algebraic properties. Some of these properties will follow quickly from results already proved for formal power series.

7.12. Definition: Formal Polynomials. A formal power series $F \in K[[x]]$ is a *polynomial* iff $\{n \in \mathbb{N} : F(n) \neq 0\}$ is a finite set. Let $K[x]$ be the set of all polynomials in $K[[x]]$.

Intuitively, a polynomial is a formal power series with only finitely many nonzero coefficients.

7.13. Definition: Degree of a Polynomial. Given a nonzero polynomial $f \in K[x]$, the *degree* of f, denoted $\deg(f)$, is the largest $n \in \mathbb{N}$ with $f(n) \neq 0$. The element $f(n) \in K$ is the *leading coefficient* of f. A polynomial f is called *monic* iff $f(n) = 1$. The degree of the zero polynomial is undefined.

7.14. Theorem: Properties of Degree. For all $f, g \in K[x]$, (a) $f+g$ is a polynomial, and $\deg(f+g) \leq \max(\deg(f), \deg(g))$ whenever both sides are defined. (b) fg is a polynomial. If f and g are nonzero, then $fg \neq 0$, and $\deg(fg) = \deg(f) + \deg(g)$.

Proof. (a) Certainly $f+g$ is a polynomial if $f = 0$ or $g = 0$. Otherwise, let $n = \deg(f)$, $m = \deg(g)$, and $k = \max(m, n)$. For all $i > k$, $(f+g)(i) = f(i) + g(i) = 0 + 0 = 0$. On one hand, this shows that $\{i \in \mathbb{N} : (f+g)(i) \neq 0\} \subseteq \{0, 1, \ldots, k\}$, so that $f+g$ is a polynomial. On the other hand, this also shows that $\deg(f+g) \leq k$ if $f + g \neq 0$.

(b) Certainly fg is a polynomial if $f = 0$ or $g = 0$. Now assume f is nonzero of degree n, and g is nonzero of degree m. Thus $f(i) = 0$ for all $i > n$ and $g(j) = 0$ for all $j > m$. Suppose $k > n + m$ and $(i, j) \in \mathbb{N}^2$ satisfy $i + j = k$. Then we must have either $i > n$ or $j > m$. Thus, every summand in the expression $(fg)(k) = \sum_{i+j=k} f(i)g(j)$ is zero, so $(fg)(k) = 0$ for all $k > n + m$. This shows that fg is a polynomial. We also have $(fg)(n+m) = \sum_{i+j=n+m} f(i)g(j) = f(n)g(m)$ since the only nonzero summand occurs when $i = n$ and $j = m$. Now $f(n)g(m) \neq 0$ since $f(n) \neq 0$ and $g(m) \neq 0$ *and K is a field* (cf. 7.27 below). Thus, $(fg)(n+m) \neq 0$, whereas all higher coefficients of fg are zero. We now see that $fg \neq 0$ and $\deg(fg) = n + m = \deg(f) + \deg(g)$. □

7.15. Theorem: Algebraic Structure of $K[x]$. (a) $K[x]$ is a commutative ring containing the field K. (b) $K[x]$ is a vector space over K with basis $\{X_i : i \geq 0\} = \{x^i : i \geq 0\}$ (see 7.9).

Proof. (a) We have just seen that $K[x]$ is closed under addition and multiplication. All the other ring axioms in 2.2 follow automatically from the corresponding ring axioms for $K[[x]]$, once we notice that $-f$ is a polynomial whenever f is, and $0_{K[[x]]}$ and $1_{K[[x]]}$ are polynomials. More generally, for any $c \in K$, every power series of the form $(c, 0, 0, \ldots)$ is a polynomial. So $K[x]$ contains the field K (or, more precisely, the copy of K inside $K[[x]]$).

(b) Given a nonzero polynomial $f \in K[x]$, let $n = \deg(f)$. We see that

$$f(0)X_0 + f(1)X_1 + \cdots + f(n)X_n = f,$$

since both series take the same value at every $k \in \mathbb{N}$ (cf. 7.9). Thus $\{X_i : i \in \mathbb{N}\}$ is a spanning set for the vector space $K[x]$. To see that this set is linearly independent, consider a finite linear combination

$$c_{i_1}X_{i_1} + c_{i_2}X_{i_2} + \cdots + c_{i_k}X_{i_k} = 0,$$

where the i_j's are distinct indices and each $c_{i_j} \in K$. Evaluating the left side at i_j, we see that $c_{i_j} = 0$ for all j. Thus, $\{X_i : i \geq 0\}$ is a basis of $K[x]$. □

Note that $B = \{x^i : i \geq 0\}$ is a basis for $K[x]$ but not a basis for $K[[x]]$. The set B does not span $K[[x]]$, because we are not allowed to form "infinite linear combinations" $\sum_{i=0}^{\infty} c_i x^i$ when determining the span (in the linear-algebraic sense) of the set B. It is true that $K[[x]]$ has some basis (as does every vector space) — but this basis will be much larger than the collection B and cannot be specified explicitly.

7.16. Example. The sequences $f = (2, 0, 1, 3, 0, 0, \ldots)$ and $g = (1, -1, 0, -3, 0, 0, \ldots)$ are polynomials of degree 3, which can be written in terms of the basis B as $f = 2 + x^2 + 3x^3$ and $g = 1 - x - 3x^3$. We calculate

$$f + g = 3 - x + x^2, \qquad fg = 2 - 2x + x^2 - 4x^3 - 3x^4 - 3x^5 - 9x^6.$$

We have $\deg(f) = 3 = \deg(g)$, $\deg(fg) = 6 = \deg(f) + \deg(g)$, and $\deg(f + g) = 2 \leq \max(\deg(f), \deg(g))$. We see that strict inequality occurs in the last formula, since the leading coefficients of f and g cancel in $f + g$.

We stress once more that formal polynomials are not "functions of x." Two formal polynomials $f = \sum_{n \geq 0} f_n x^n$ and $g = \sum_{n \geq 0} g_n x^n$ are equal iff $f_n = g_n$ for all $n \in \mathbb{N}$. This is true by the definition of formal polynomials as sequences (functions with domain $\mathbb{N}$) and is equivalent to the linear independence of the basis B. Nevertheless, we can use a formal polynomial to define an associated polynomial function, as follows.

7.17. Definition: Polynomial Functions. Given a nonzero polynomial $f \in K[x]$ and a commutative ring R containing the field K, the *polynomial function associated to f with domain R* is the function $P_f : R \to R$ defined by

$$P_f(z) = \sum_{n=0}^{\deg(f)} f_n z^n \qquad (z \in R).$$

If $f = 0$, we let $P_f(z) = 0$ for all $z \in R$.

One can show that, *because R contains the infinite field $\mathbb{Q}$*, $f = g$ (equality of formal polynomials) iff $P_f = P_g$ (equality of functions). However, this statement fails if one considers polynomial functions defined on *finite* rings and fields (see the exercises).

7.18. Example. Let $f = 2 + x^2 + 3x^3 \in \mathbb{Q}[x]$, and let $R = \mathbb{C}$ (the complex numbers). Then

$$P_f(\sqrt{2}) = 2 + (\sqrt{2})^2 + 3(\sqrt{2})^3 = 4 + 6\sqrt{2}; \qquad P_f(i) = 2 + i^2 + 3i^3 = 1 - 3i.$$

7.19. Example. Let $h = 1 - x + x^2 \in \mathbb{Q}[x]$, and let $R = \mathbb{Q}[x]$. Then $P_h(2x^3) = 1 - (2x^3) + (2x^3)^2 = 1 - 2x^3 + 4x^6$ and $P_h(h) = 1 - (1 - x + x^2) + (1 - x + x^2)^2 = 1 - x + 2x^2 - 2x^3 + x^4$. Note also that $P_h(x) = 1 - x + x^2 = h$; more generally, $P_f(x) = f$ for any $f \in K[x]$. Next suppose $R = \mathbb{Q}[[x]]$ and $z = \sum_{n \geq 0} x^n \in R$. Then

$$P_h(z) = (1, 0, 0, 0, 0, \ldots) - (1, 1, 1, 1, 1, \ldots) + (1, 2, 3, 4, 5, \ldots) = (1, 1, 2, 3, 4, \ldots).$$

Intuitively, the next result confirms that the algebraic operations on formal polynomials agree with the familiar algebraic operations on polynomial functions.

7.20. Theorem: Comparison of Algebraic Operations on Formal Polynomials and Polynomial Functions. Let $f, g \in K[x]$, let $c \in K \subseteq K[x]$, let R be a commutative ring containing the field K, and let $z \in R$. (a) $P_{f+g}(z) = P_f(z) + P_g(z)$. (b) $P_{fg}(z) = P_f(z) P_g(z)$. (c) $P_c(z) = c$.

Proof. We prove (b), leaving (a) and (c) as exercises. Both sides in (b) are zero if $f = 0$ or $g = 0$, so assume $f \neq 0$ and $g \neq 0$. Write $n = \deg(f)$ and $m = \deg(g)$, so $\deg(fg) = n + m$.

Now, compute

$$\begin{aligned}
P_{fg}(z) &= \sum_{k=0}^{n+m} (fg)_k z^k \text{ (definition of } P_{fg}) \\
&= \sum_{k=0}^{n+m} \left(\sum_{i+j=k} f_i g_j \right) z^k \text{ (definition of } fg) \\
&= \sum_{k=0}^{n+m} \sum_{i+j=k} (f_i z^i g_j z^j) \text{ (distributive law in } R \text{ and } \textit{commutativity} \text{ of } R) \\
&= \left(\sum_{i=0}^{n} f_i z^i \right) \left(\sum_{j=0}^{m} g_j z^j \right) \text{ (generalized distributive law in } R) \\
&= P_f(z) P_g(z) \text{ (definition of } P_f \text{ and } P_g). \quad \square
\end{aligned}$$

We can rephrase this result in a somewhat more sophisticated way, using the following definition.

7.21. Definition: Ring Homomorphisms. Let R and S be rings. A map $f : R \to S$ is a *ring homomorphism* iff $f(1_R) = 1_S$ and for all $x, y \in R$, $f(x + y) = f(x) + f(y)$ and $f(xy) = f(x)f(y)$.

7.22. Theorem: Evaluation Homomorphisms on $K[x]$. Suppose R is a commutative ring containing the field K. For each $z \in R$, there exists a unique ring homomorphism $\mathrm{ev}_z : K[x] \to R$ such that $\mathrm{ev}_z(x) = z$ and $\mathrm{ev}_z(c) = c$ for all $c \in K$. Furthermore, $\mathrm{ev}_z(f) = P_f(z)$ for all $f \in K[x]$. We call ev_z the *evaluation homomorphism on $K[x]$ determined by evaluating x at z.*

Proof. The particular map $\mathrm{ev}_z(f) = P_f(z)$ *is* a ring homomorphism sending x to z and fixing K, since (by the previous theorem) for all $f, g \in K[x]$ and $c \in K$,

$$\begin{aligned}
\mathrm{ev}_z(f+g) &= P_{f+g}(z) = P_f(z) + P_g(z) = \mathrm{ev}_z(f) + \mathrm{ev}_z(g); \\
\mathrm{ev}_z(fg) &= P_{fg}(z) = P_f(z) P_g(z) = \mathrm{ev}_z(f)\,\mathrm{ev}_z(g); \\
\mathrm{ev}_z(c) &= P_c(z) = c \qquad (\text{so } \mathrm{ev}_z(1_{K[x]}) = 1_R);
\end{aligned}$$

and $\mathrm{ev}_z(x) = P_x(z) = z$. To prove uniqueness, let $E : K[x] \to R$ be any ring homomorphism with $E(x) = z$ and $E(c) = c$ for $c \in K$. For a nonzero $f \in K[x]$ of degree n, we have

$$E(f) = E\left(\sum_{k=0}^{n} f_k x^k \right) = \sum_{k=0}^{n} E(f_k x^k) = \sum_{k=0}^{n} E(f_k) E(x)^k = \sum_{k=0}^{n} f_k z^k = P_f(z) = \mathrm{ev}_z(f).$$

As $E(0) = 0_R = \mathrm{ev}_z(0)$, we conclude that $E = \mathrm{ev}_z$. $\square$

7.4 Order of Formal Power Series

We now discuss the order of a formal power series, which is analogous to the degree of a formal polynomial. The degree of a polynomial F is the largest n with $F(n) \neq 0$. Such an n will not exist if F is a formal power series that is not a polynomial. So we instead proceed as follows.

7.23. Definition: Order of a Formal Power Series. Let $F = \sum_{n \geq 0} F_n x^n \in K[[x]]$ be a nonzero series. The *order* of F, denoted $\text{ord}(F)$, is the least $n \geq 0$ such that $F(n) \neq 0$. Since $\{k \in \mathbb{N} : F(k) \neq 0\}$ is a nonempty subset of $\mathbb{N}$, the order of F is well defined. The order of the zero series is undefined.

7.24. Example. The order of $(0, 0, 0, 4, 2, 1, 4, 2, 1, 4, 2, 1, \ldots)$ is 3. Nonzero polynomials have both a degree and an order; for instance, $x^2 + x^5 + 3x^7$ has degree 7 and order 2.

The properties of order are analogous to the properties of degree.

7.25. Theorem: Properties of Order. Let F and G be nonzero formal series in $K[[x]]$. (a) If $F + G \neq 0$, then $\text{ord}(F + G) \geq \min(\text{ord}(F), \text{ord}(G))$. (b) $FG \neq 0$, and $\text{ord}(FG) = \text{ord}(F) + \text{ord}(G)$.

Proof. Let $n = \text{ord}(F)$ and $m = \text{ord}(G)$. (a) Let $k = \min(n, m)$. For any $i < k$, $(F + G)_i = F_i + G_i = 0 + 0 = 0$, so $\text{ord}(F + G) \geq k$.

(b) For any $p < n + m$, $(FG)_p = \sum_{i+j=p} F_i G_j$. For any pair $(i, j) \in \mathbb{N}^2$ with sum p, either $i < n$ or $j < m$. Thus $F_i = 0$ or $G_j = 0$, so every summand $F_i G_j = 0$. Hence, $(FG)_p = 0$. On the other hand, for $p = n + m$, we only get a nonzero summand when $i = n$ and $j = m$. So $(FG)_{n+m} = F_n G_m \neq 0$, since $F_n \neq 0$ and $G_m \neq 0$ *and K is a field* (cf. 7.27 below). This shows that $FG \neq 0$ and that $n+m$ is the least element p of $\mathbb{N}$ with $(FG)_p \neq 0$, hence $\text{ord}(FG) = n + m = \text{ord}(F) + \text{ord}(G)$. □

7.26. Definition: Integral Domains. A commutative ring R with more than one element is an *integral domain* iff for all nonzero $x, y \in R$, $xy \neq 0$. Equivalently, for all $x, y \in R$, $xy = 0$ implies $x = 0$ or $y = 0$.

7.27. Example. Every field F (see 2.3) is an integral domain. For if $x, y \in K$, $xy = 0$, and $x \neq 0$, then x has a multiplicative inverse in K. Multiplying $xy = 0$ by the inverse of x, we see that $y = 0$. The ring $\mathbb{Z}$ is an integral domain that is not a field. Part (b) of 7.14 shows that $K[x]$ is an integral domain. Part (b) of 7.25 shows that $K[[x]]$ is an integral domain. A key step in both proofs was the deduction that $F_n G_m \neq 0$ since $F_n \neq 0$ and $G_m \neq 0$. Thus, these proofs tacitly used the fact that the field K is an integral domain. This hypothesis on K (which is weaker than assuming that K is a field) is enough to ensure that $K[x]$ and $K[[x]]$ will be integral domains.

7.5 Formal Limits, Infinite Sums, and Infinite Products

Even in the algebraic setting of formal power series, one can imitate the limiting operations that play such a prominent role in calculus. In particular, we can use formal limits of formal power series to define infinite sums and infinite products of formal power series in certain situations.

7.28. Definition: Limit of a Sequence of Formal Power Series. Suppose $(F_k : k \in \mathbb{N})$ is a sequence of elements of $K[[x]]$ (so $F_k : \mathbb{N} \to K$ for each $k \in \mathbb{N}$), and $G \in K[[x]]$. Write

$$\lim_{k \to \infty} F_k = G \qquad (\text{or } F_k \to G)$$

iff for each $n \geq 0$, there exists an index $K(n)$ such that $k \geq K(n)$ implies $F_k(n) = G(n)$.

Informally, the sequence $(F_k : k \in \mathbb{N})$ of formal power series converges to some (necessarily unique) limit series in $K[[x]]$ iff for each $n \in \mathbb{N}$, the coefficient of x^n in F_k eventually becomes constant for large enough values of k (and this constant is the coefficient of x^n in the limit series G).

7.29. Example. We have $\lim_{k\to\infty} x^k = 0$ in $K[[x]]$. To prove this, fix any n and then note that $k > n$ implies $x^k(n) = 0 = 0(n)$.

Now that limits are available, we can define infinite sums (resp. products) as limits of partial finite sums (resp. products).

7.30. Definition: Infinite Sums and Products of Formal Series. Suppose $(F_k : k \in \mathbb{N})$ is a sequence of formal power series in $K[[x]]$. For each $N \geq 0$, let $G_N = F_0 + F_1 + \cdots + F_N \in K[[x]]$ be the Nth partial sum of this sequence. If $H = \lim_{N\to\infty} G_N$ exists in $K[[x]]$, then we write $\sum_{k=0}^{\infty} F_k = H$. Similarly, let $P_N = F_0 F_1 \cdots F_N \in K[[x]]$ be the Nth partial product of the sequence of F_k's. If $Q = \lim_{N\to\infty} P_N$ exists in $K[[x]]$, then we write $\prod_{k=0}^{\infty} F_k = Q$. Analogous definitions are made for sums and products ranging over any countably infinite index set (e.g., for k ranging from 1 to ∞).

7.31. Example. Given $F = \sum_{n=0}^{\infty} F_n x^n \in K[[x]]$, define a formal series $G_k = F_k X_k = F_k x^k$ (see 7.9) for each $k \geq 0$. We have

$$\sum_{k=0}^{n} G_k = \sum_{k=0}^{n} F_k x^k = (F_0, F_1, \ldots, F_n, 0, 0, \ldots).$$

Given $m \in \mathbb{N}$, it follows that the coefficient of x^m in any partial sum $\sum_{k=0}^{n} G_k$ with $n \geq m$ is $F_m = F(m)$. Thus, by definition, $\sum_{k=0}^{\infty} G_k$ has limit F. In other words,

$$\sum_{k=0}^{\infty} F_k X_k = F = \sum_{k=0}^{\infty} F_k x^k,$$

where the left side is an infinite sum of formal power series, and the right side is our notation for the single formal power series F. This equality finally justifies the use of the "power series notation" for elements of $K[[x]]$.

The previous example can be rephrased as follows.

7.32. Theorem: Density of Polynomials in $K[[x]]$. For each $F \in K[[x]]$, there exists a sequence of polynomials $f_n \in K[x]$ such that $\lim_{n\to\infty} f_n = F$. Specifically, we can take $f_n = \sum_{k=0}^{n} F_k x^k$.

Testing the convergence of infinite sums and products of real-valued functions is a delicate and often difficult problem. On the other hand, we can use the notion of order to give simple and convenient criteria ensuring the existence of infinite sums and infinite products of *formal* power series. Recall from calculus that a sequence of integers $(e_n : n \geq 0)$ tends to infinity (in $\mathbb{R}$) iff for every integer $K > 0$, there exists N such that $n \geq N$ implies $e_n > K$.

7.33. Theorem: Existence Criteria for Limits of Formal Series. Suppose $(F_k : k \in \mathbb{N})$ is a sequence of nonzero formal power series in $K[[x]]$.
(a) $\lim_{k\to\infty} F_k = 0$ in $K[[x]]$ iff $\lim_{k\to\infty} \operatorname{ord}(F_k) = \infty$ in $\mathbb{R}$.
(b) $\sum_{k=0}^{\infty} F_k$ exists in $K[[x]]$ iff $\lim_{k\to\infty} \operatorname{ord}(F_k) = \infty$ in $\mathbb{R}$.
(c) If $F_k(0) = 0$ for all k, then $\prod_{k=0}^{\infty}(1 + F_k)$ exists in $K[[x]]$ iff $\lim_{k\to\infty} \operatorname{ord}(F_k) = \infty$ in $\mathbb{R}$.

Proof. (a) Assume $F_k \to 0$ in $K[[x]]$. Choose a fixed integer $M \geq 0$. For each n between 0 and M, there exists an index k_n such that $k \geq k_n$ implies $F_k(n) = 0$. Hence, whenever $k \geq K = \max(k_0, k_1, \ldots, k_M)$, we have $F_k(n) = 0$ for all $n \leq M$. It follows that $\operatorname{ord}(F_k) > M$ whenever $k \geq K$. This proves that the sequence of integers $(\operatorname{ord}(F_k) : k \in \mathbb{N})$ tends to infinity as k goes to infinity. Conversely, suppose $\operatorname{ord}(F_k) \to \infty$ as $k \to \infty$. Fix n, and choose K so that $k \geq K$ implies $\operatorname{ord}(F_k) > n$. It follows immediately that $F_k(n) = 0 = 0(n)$ for all $k \geq K$. Thus, $F_k \to 0$ in $K[[x]]$.

(b) Suppose $\sum_{k \geq 0} F_k$ converges to G in $K[[x]]$. Given an index n, we can therefore choose K so that $k \geq K$ implies

$$(F_0 + F_1 + \cdots + F_k)(n) = G(n).$$

Given $k > K$, note that

$$F_k(n) = (F_0 + \cdots + F_k)(n) - (F_0 + \cdots + F_{k-1})(n) = G(n) - G(n) = 0.$$

This proves that $F_k \to 0$ as $k \to \infty$, and hence $\operatorname{ord}(F_k) \to \infty$ by (a). Conversely, suppose $\operatorname{ord}(F_k) \to \infty$, so that $F_k \to 0$ by (a). For each fixed n, the coefficient of x^n in F_k is eventually zero, and hence the coefficient of x^n in the partial sum $F_0 + F_1 + \cdots + F_k$ eventually stabilizes. Thus, these partial sums have a limit in $K[[x]]$.

(c) Suppose the indicated infinite product exists. We must show that for all n, $\operatorname{ord}(F_k)$ is eventually $\geq n$. We prove this by induction on n. The statement is true for $n = 1$ since $F_k(0) = 0$ for all k. Assume the statement holds for some $n \geq 1$. Choose k_0 so that $\operatorname{ord}(F_k) \geq n$ for all $k \geq k_0$. Next, using the hypothesis that the infinite product exists, choose k_1 so that $j, k \geq k_1$ implies

$$\left[\prod_{i=0}^{j}(1 + F_i)\right](n) = \left[\prod_{i=0}^{k}(1 + F_i)\right](n).$$

Note that

$$\prod_{i=0}^{k}(1 + F_i) - \prod_{i=0}^{k-1}(1 + F_i) = \left[\prod_{i=0}^{k-1}(1 + F_i)\right](1 + F_k - 1) = F_k \prod_{i=0}^{k-1}(1 + F_i). \tag{7.1}$$

For $k > k_1$, the coefficient of x^n on the left side is zero. On the other hand, for $k \geq k_0$, the fact that $\operatorname{ord}(F_k) \geq n$ implies that

$$\left[F_k \prod_{i=0}^{k-1}(1 + F_i)\right](n) = F_k(n),$$

since the partial product has constant term 1. Combining these facts, we see that $F_k(n) = 0$ for $k > \max(k_0, k_1)$. This shows that $\operatorname{ord}(F_k)$ eventually exceeds n, completing the induction.

Conversely, suppose $\operatorname{ord}(F_k) \to \infty$ as $k \to \infty$. Fix n; we must show that the coefficient of x^n in the partial products $\prod_{i=0}^{k}(1 + F_i)$ eventually stabilizes. Choose k_0 so that $k \geq k_0$ implies $\operatorname{ord}(F_k) > n$. It suffices to show that for all $k > k_0$,

$$\left[\prod_{i=0}^{k}(1 + F_i)\right](n) = \left[\prod_{i=0}^{k-1}(1 + F_i)\right](n).$$

Subtracting and using (7.1), we see that the condition we want is equivalent to

$$\left[F_k \prod_{i=0}^{k-1}(1 + F_i)\right](n) = 0 \qquad (k > k_0).$$

This holds because the product appearing on the left side here has order greater than n. □

7.34. Example. The infinite product $\prod_{n=1}^{\infty}(1+x^n)$ is a well-defined element of $K[[x]]$, since $\text{ord}(x^n) = n \to \infty$ as $n \to \infty$.

7.35. Theorem: Limit Rules for Sums and Products. Suppose $F_n, G_n, P, Q \in K[[x]]$ are formal series such that $F_n \to P$ and $G_n \to Q$. Then $F_n + G_n \to P+Q$ and $F_n G_n \to PQ$.

Proof. We prove the second statement, leaving the first as an exercise. Fix $m \in \mathbb{N}$. We must find $N \in \mathbb{N}$ so that $n \geq N$ implies $(F_n G_n)(m) = (PQ)(m)$. For each $j \leq m$, there is an $N_j \in \mathbb{N}$ such that $n \geq N_j$ implies $F_n(j) = P(j)$. Similarly, for each $k \leq m$, there is an $M_k \in \mathbb{N}$ such that $n \geq M_k$ implies $G_n(k) = Q(k)$. Let $N = \max(N_0, \ldots, N_m, M_0, \ldots, M_m) \in \mathbb{N}$. For any $n \geq N$,

$$(PQ)(m) = \sum_{j+k=m} P(j)Q(k) = \sum_{j+k=m} F_n(j)G_n(k) = (F_n G_n)(m). \qquad \square$$

7.6 Multiplicative Inverses in $K[x]$ and $K[[x]]$

In any ring S, it is of interest to know which elements of S have multiplicative inverses in S.

7.36. Definition: Units of a Ring. An element x in a ring S is called a *unit of* S iff there exists $y \in S$ with $xy = yx = 1_S$.

Suppose $y, z \in S$ satisfy $xy = yx = 1_S$ and $xz = zx = 1_S$. Then $y = y1 = y(xz) = (yx)z = 1z = z$, so $y = z$. Thus, if x has a multiplicative inverse in S, this inverse is unique. We write x^{-1} or $1/x$ to denote this inverse.

7.37. Example. If $|S| > 1$, then $1_S \neq 0_S$, and zero is not a unit of S. By definition, every nonzero element of a field F is a unit of F. In particular, the units of $\mathbb{Q}$ are the nonzero rational numbers. On the other hand, the only units of the ring $\mathbb{Z}$ are 1 and -1. So 2 is a unit of $\mathbb{Q}$ but not a unit of $\mathbb{Z}$.

Next we characterize the units of the polynomial ring $K[x]$ and the formal power series ring $K[[x]]$. Our first result says that the only units in the polynomial ring $K[x]$ are the nonzero scalars.

7.38. Theorem: Units of $K[x]$. A polynomial $f \in K[x]$ is a unit in $K[x]$ iff $\deg(f) = 0$.

Proof. The zero polynomial is not a unit of $K[x]$, so assume $f \neq 0$ henceforth. First, suppose $f = a_0 x^0$ is a degree zero polynomial, so $a_0 \in K$ is nonzero. Since K is a field, a_0^{-1} exists in K. In $K[x]$, we have $a_0 a_0^{-1} = 1_{K[x]} = a_0^{-1} a_0$, so a_0^{-1} is also a multiplicative inverse for f in the ring $K[x]$. Thus, f is a unit of $K[x]$.

Conversely, suppose $\deg(f) > 0$. For any nonzero $g \in K[x]$, we know $\deg(fg) = \deg(f) + \deg(g) > 0$ (this result uses the fact that K is a field). Thus $fg \neq 1_{K[x]}$, since $\deg(1_{K[x]}) = 0$. So f does not have a multiplicative inverse g in $K[x]$. $\square$

Intuitively, a non-constant polynomial cannot be a unit since there is no way to get rid of the positive powers of x. Perhaps surprisingly, when we pass to the larger ring of formal power series, almost every element in the ring becomes a unit. More precisely, every series with nonzero constant term has an inverse in $K[[x]]$. Before proving this, we consider an example that occurs frequently.

7.39. Example: Formal Geometric Series. Consider the series $F = (1, -1, 0, 0, 0, 0, \ldots)$ and $G = (1, 1, 1, 1, 1, \ldots)$. Multiplying these series, we discover that

$$FG = GF = (1, 0, 0, 0, \ldots) = 1_{K[[x]]}.$$

Thus the polynomial $1 - x$ is invertible in $K[[x]]$, and

$$(1 - x)^{-1} = 1 + x + x^2 + \cdots + x^n + \cdots = \sum_{n \geq 0} x^n.$$

This is a formal version of the "geometric series formula" learned in calculus. In calculus, one requires $|x| < 1$ to ensure convergence. There is no such restriction here, since the letter x in our formula does not denote a real number!

7.40. Theorem: Units in $K[[x]]$. A formal power series $F \in K[[x]]$ is a unit in $K[[x]]$ iff $F(0) \neq 0$.

Proof. Assume $F(0) = 0$. For any $G \in K[[x]]$, $(FG)(0) = F(0)G(0) = 0 \neq 1 = 1_{K[[x]]}(0)$. So $FG \neq 1$ in $K[[x]]$, and F is not a unit of $K[[x]]$.

Conversely, assume $F(0) \neq 0$. Our goal is to find a series $G = \sum_{n \geq 0} G_n x^n$ such that $FG = GF = 1_{K[[x]]}$. The desired equation $FG = 1$ holds in $K[[x]]$ iff the following infinite system of equations holds in the field K:

$$\begin{array}{rcl} F_0 G_0 & = & 1 \\ F_0 G_1 + F_1 G_0 & = & 0 \\ F_0 G_2 + F_1 G_1 + F_2 G_0 & = & 0 \\ \cdots & & \cdots \\ \sum_{k=0}^{n} F_k G_{n-k} & = & 0 \\ \cdots & & \cdots \end{array} \tag{7.2}$$

We claim that there exist unique scalars $G_0, G_1, \ldots, G_n, \ldots$ solving this system. To prove existence, we "solve the preceding system for the unknowns G_n." More precisely, we recursively define $G_0 = F_0^{-1} \in K$, $G_1 = F_0^{-1}(-F_1 G_0)$, $G_2 = F_0^{-1}(-F_1 G_1 - F_2 G_0)$, and in general

$$G_n = -F_0^{-1} \sum_{k=1}^{n} F_k G_{n-k}. \tag{7.3}$$

By construction, the scalars $G_n \in K$ defined in this way satisfy (7.2), and therefore $G = \sum_{n \geq 0} G_n x^n$ satisfies $FG = GF = 1$. Since $G = F^{-1}$ in $K[[x]]$, G (and hence the G_n) are uniquely determined by F. □

The preceding proof gives a recursive algorithm for calculating any given coefficient of $1/F$ in terms of the coefficients of F and lower coefficients of $1/F$. In some situations, the following theorem (which generalizes 7.39) gives a more convenient way to calculate the multiplicative inverse of a formal series.

7.41. Theorem: Formal Geometric Series. If $G \in K[[x]]$ satisfies $G(0) = 0$, then

$$(1 - G)^{-1} = \frac{1}{1 - G} = 1 + G + G^2 + G^3 + \cdots = \sum_{k=0}^{\infty} G^k \in K[[x]].$$

Proof. The theorem holds if $G = 0$, so assume G is nonzero. Suppose $\operatorname{ord}(G) = d > 0$; then $\operatorname{ord}(G^k) = kd$, which goes to infinity as $k \to \infty$. It follows that $H = \sum_{k=0}^{\infty} G^k$ exists.

Consider the coefficient of x^n in $(1-G)H$. By order considerations, this coefficient is the same as the coefficient of x^n in

$$(1-G)(1+G+G^2+\cdots+G^n) = 1-G^{n+1}.$$

Since $G^{n+1}(n) = 0$, we see that the coefficient in question is 1 if $n = 0$ and is 0 if $n > 0$. Thus, $(1-G)H = H(1-G) = 1$, so $H = (1-G)^{-1}$. □

7.42. Example. Taking $G = x^i$ in the theorem (where $i \geq 1$ is a fixed integer), we find that

$$\frac{1}{1-x^i} = 1 + x^i + x^{2i} + x^{3i} + \cdots \in K[[x]].$$

This series has the form $1 + F$ where $\text{ord}(F) = i$. It follows that the infinite product $\prod_{i=1}^{\infty} \frac{1}{1-x^i}$ exists. Similarly, $\prod_{i=1}^{\infty}(1-x^i)$ exists because $\text{ord}(-x^i) = i$, which goes to infinity as i goes to infinity. Next, we claim that

$$\prod_{i=1}^{\infty}(1-x^i)\prod_{i=1}^{\infty}\frac{1}{1-x^i} = 1.$$

One might at first believe that this identity is automatically true (due to "cancellation"), but care is required since we are dealing with formal infinite products. To justify this identity carefully, let $P_n = \prod_{i=1}^{n}(1-x^i)$ and $Q_n = \prod_{i=1}^{n}(1-x^i)^{-1}$ for each $n \in \mathbb{N}$. Using 7.35, we see that

$$\prod_{i=1}^{\infty}(1-x^i)\prod_{i=1}^{\infty}\frac{1}{1-x^i} = \left(\lim_{n\to\infty} P_n\right)\cdot\left(\lim_{n\to\infty} Q_n\right) = \lim_{n\to\infty} P_nQ_n = \lim_{n\to\infty} 1 = 1.$$

Note that we can rearrange the factors in each *finite* product P_nQ_n (since multiplication is commutative) and cancel to obtain 1.

7.43. Example. The geometric series formula can be used to invert any invertible formal power series, not just series with constant term 1. For, suppose $F = \sum_{n\geq 0} F_n x^n$ where $F_0 \neq 0$. We can write $F = F_0(1-G)$ where $G = \sum_{n\geq 1}(-F_0^{-1}F_n)x^n$. Then

$$F^{-1} = F_0^{-1}(1-G)^{-1} = F_0^{-1}[1+G+G^2+\cdots+G^k+\cdots].$$

7.7 Formal Laurent Series

We saw in the last section that $G \in K[[x]]$ is a unit in $K[[x]]$ iff $G(0) \neq 0$. Sometimes we want to divide by elements of $K[[x]]$ that are not units. To do this, we need to operate in the *field of fractions* of the integral domain $K[[x]]$. We summarize this construction now, omitting many routine details; more thorough treatments may be found in texts on abstract algebra.

7.44. Construction: Field of Fractions of an Integral Domain. Let D be an integral domain, and let D^* be the set of nonzero elements of D. Let $X = D \times D^*$ be the set of pairs (a,b) where $a, b \in D$ and b is nonzero. Define a relation $\sim$ on X by setting $(a,b) \sim (c,d)$ iff $ad = bc$. One may verify that $\sim$ is an equivalence relation (checking transitivity requires the assumption that D is an integral domain). Write $F = \text{Frac}(D)$ to denote the set of

equivalence classes of this equivalence relation; also, write a/b to denote the equivalence class of (a, b).

Given two elements a/b and c/d in F, define addition and multiplication as follows:

$$(a/b) + (c/d) = (ad + bc)/(bd); \qquad (a/b) \times (c/d) = (ac)/(bd).$$

It must be checked that these operations are independent of the representatives chosen for the equivalence classes. For example, one must show that $a/b = a'/b'$ and $c/d = c'/d'$ imply $(ad + bc)/(bd) = (a'd' + b'c')/(b'd')$. Also define zero and one in F by $0_F = 0_D/1_D$ and $1_F = 1_D/1_D$. One may now check that $(F, +, \times, 0_F, 1_F)$ is a commutative ring with $1_F \neq 0_F$. The map $i : D \to F$ such that $i(a) = a/1_D$ for $a \in D$ is an injective ring homomorphism that embeds D as a subring of F and allows us to regard D as a subset of F. Finally (and this is the point of the whole construction), every nonzero element $a/b \in F$ has a multiplicative inverse in F, namely b/a. This follows since $(a/b) \times (b/a) = (ab)/(ba)$, and this equals $1_D/1_D = 1_F$ because $(ab)1 = (ba)1$ in D. Therefore, F is a field.

The field F has the following *universal mapping property*: for any ring homomorphism $g : D \to L$ into a *field* L, there exists a unique ring homomorphism $g' : F \to L$ extending g (more precisely, such that $g = g' \circ i$). This homomorphism must be given by $g'(a/b) = g(a)g(b)^{-1} \in L$ (proving uniqueness); for existence, one checks that the formula just written does give a well-defined ring homomorphism extending g.

7.45. Example: $\mathbb{Z}$ and $\mathbb{Q}$. Let $\mathbb{Z}$ be the ring of integers, which is an integral domain. The field $\mathbb{Q}$ of rational numbers is, by definition, the field of fractions of $\mathbb{Z}$.

7.46. Definition: Formal Rational Functions. The symbol $K(x)$ denotes the fraction field of the integral domain $K[x]$. Elements of $K(x)$ are *formal rational functions* p/q, where p, q are polynomials with q nonzero; we have $p/q = s/t$ in $K(x)$ iff $pt = qs$ in $K[x]$.

7.47. Definition: Formal Laurent Series. The symbol $K((x))$ denotes the fraction field of the integral domain $K[[x]]$. Elements of $K((x))$ are *formal Laurent series* in one indeterminate. A formal Laurent series is a quotient G/H, where G, H are formal power series with H nonzero; we have $G/H = P/Q$ in $K((x))$ iff $GQ = HP$ in $K[[x]]$.

We can use our characterization of units in $K[[x]]$ to find a canonical description of the elements of $K((x))$.

7.48. Theorem: Representation of Laurent Series. For every nonzero $S \in K((x))$, there exists a unique integer N and a unique series $F \in K[[x]]$ such that $S = x^N F$ and $F(0) \neq 0$.

7.49. Remark. We call N the *order* of S. When $N \geq 0$, so that $S \in K[[x]]$, this is consistent with the previous definition of $\mathrm{ord}(S)$. When $N = -m$ is negative, $S = x^N F$ is the fraction F/x^m. In this case, we often use the "Laurent series notation"

$$S = F_0 x^{-m} + F_1 x^{-m+1} + F_2 x^{-m+2} + \cdots + F_m x^0 + F_{m+1} x^1 + \cdots = \sum_{n=-m}^{\infty} F_{n+m} x^n.$$

Proof. Given a nonzero $S \in K((x))$, there exist nonzero series $G, H \in K[[x]]$ with $S = G/H$; G and H are not unique. Write $\mathrm{ord}(G) = i$, $\mathrm{ord}(H) = j$, $G = \sum_{n \geq i} G_n x^n$, $H = \sum_{n \geq j} H_n x^n$, where G_i and H_j are nonzero. Let $H^* = \sum_{n \geq 0} H_{n+j} x^n$, which is a unit in $K[[x]]$ since $H^*(0) = H_j \neq 0$. Let $Q \in K[[x]]$ be the inverse of H^*, which satisfies $Q(0) \neq 0$. Similarly, write $G^* = \sum_{n \geq 0} G_{n+i} x^n$ and note that $G^*(0) \neq 0$. Now,

$$S = \frac{G}{H} = \frac{x^i G^*}{x^j H^*} = \frac{x^i G^* Q}{x^j H^* Q} = x^N F$$

if we set $N = i - j \in \mathbb{Z}$ and $F = G^*Q \in K[[x]]$. This proves existence of N and F.

For uniqueness, assume $x^N F = x^M P$ for some $M \in \mathbb{Z}$ and some $P \in K[[x]]$ with $P(0) \neq 0$. Choose $k > \max(|N|, |M|)$; then $x^{k+N}F = x^{k+M}P$. Both sides are nonzero series in $K[[x]]$ and hence have an order. Since F and P have nonzero constant term, comparison of orders gives $k + N = k + M$ and hence $N = M$. Dividing by x^N (which is a unit in $K((x))$!), we see that $F = P$. □

7.50. Remark. The proof shows that, for *any* representation of a nonzero $S \in K((x))$ as a quotient F/G with $F, G \in K[[x]]$, we have $\operatorname{ord}(S) = \operatorname{ord}(F) - \operatorname{ord}(G)$.

7.8 Formal Derivatives

We now define a formal version of the derivative operation studied in calculus.

7.51. Definition: Formal Derivatives. Given a formal series $F = \sum_{n\geq 0} F_n x^n \in K[[x]]$, the *formal derivative* of F is

$$F' = \frac{dF}{dx} = DF = \sum_{n\geq 0} (n+1)F_{n+1}x^n.$$

Higher-order formal derivatives are defined recursively by setting $F^{(k+1)} = (F^{(k)})'$ for $k \geq 1$. It follows that

$$F^{(k)} = \sum_{n\geq 0} (n+1)(n+2)\cdots(n+k)F_{n+k}x^n.$$

The integer coefficients appearing in these formulas make sense, since we have assumed that K is a field containing $\mathbb{Q}$.

To give examples of formal differentiation, we introduce formal versions of some familiar functions from calculus.

7.52. Definition: Formal Versions of Exponential, Logarithmic, and Trigonometric Functions. Define the following elements in $K[[x]]$ (recall K contains $\mathbb{Q}$):

$$e^x = (1, 1, 1/2, 1/6, 1/24, 1/120, \ldots) = \sum_{n\geq 0} \frac{1}{n!}x^n;$$

$$\sin x = (0, 1, 0, -1/6, 0, 1/120, 0, \ldots) = \sum_{n\geq 0} \chi(n \text{ is odd})\frac{(-1)^{(n-1)/2}}{n!}x^n;$$

$$\cos x = (1, 0, -1/2, 0, 1/24, 0, -1/720, \ldots) = \sum_{n\geq 0} \chi(n \text{ is even})\frac{(-1)^{n/2}}{n!}x^n;$$

$$\log(1+x) = (0, 1, -1/2, 1/3, -1/4, 1/5, \ldots) = \sum_{n\geq 1} \frac{(-1)^{n+1}}{n}x^n.$$

7.53. Example. Let $F = e^x$, $G = \sin x$, $H = \cos x$, and $P = \log(1+x)$ in $K[[x]]$. Using the definition of formal derivatives, we find that

$$F' = (1\cdot 1, 2\cdot(1/2), 3\cdot(1/3!), \ldots, (n+1)\cdot(1/(n+1)!), \ldots) = (1, 1, 1/2, 1/6, \ldots, 1/n!, \ldots) = F.$$

Thus the formal power series e^x equals its own derivative, in accordance with the situation in calculus. Iterating, we see that $f^{(k)} = f$ for all $k \geq 1$. Similar calculations show that $G' = H$, $H' = -G$, and $P' \cdot (1, 1, 0, 0, \ldots) = 1_{K[[x]]}$. We can express the last fact by writing

$$\frac{d}{dx} \log(1+x) = (1+x)^{-1}.$$

Formal derivatives obey many of the same differentiation rules that ordinary derivatives satisfy. However, each of these rules must be reproved in the formal setting. Some of these rules are stated in the next theorem.

7.54. Theorem: Formal Differentiation Rules. Let $F, G, H_n \in K[[x]]$ and $c, c_n \in K$.
(a) $(F+G)' = F' + G'$ (sum rule).
(b) $(cF)' = c(F')$ (scalar rule).
(c) For $N \in \mathbb{N}$, $\left(\sum_{n=1}^{N} c_n H_n\right)' = \sum_{n=1}^{N} c_n H_n'$ (linear combination rule).
(d) $\frac{d}{dx}(x^k) = kx^{k-1}$ for all $k \geq 0$ (power rule).
(e) $(FG)' = F(G') + (F')G$ (product rule).
(f) If $H_n \to F$, then $H_n' \to F'$ (derivative of a limit).
(g) If $S = \sum_{n=1}^{\infty} H_n$ exists, then $S' = \sum_{n=1}^{\infty} H_n'$ (derivative of an infinite sum).

Proof. We prove (d), (e), and (f), leaving the others as exercises.

(d) Recall that $x^k = \sum_{n \geq 0} \chi(n=k) x^n$. The definition of formal derivative gives

$$\frac{d}{dx}(x^k) = \sum_{n \geq 0} (n+1)\chi(n+1=k)x^n = \sum_{n \geq 0} k\chi(n=k-1)x^n = kx^{k-1}.$$

(e) Note on the one hand that

$$(FG)'_m = (m+1) \cdot (FG)_{m+1} = (m+1) \sum_{k=0}^{m+1} F_k G_{m+1-k}.$$

On the other hand,

$$\begin{aligned}
(FG' + F'G)_m &= (FG')_m + (F'G)_m \\
&= \sum_{k=0}^{m} F_k G'_{m-k} + \sum_{j=0}^{m} F'_j G_{m-j} \\
&= \sum_{k=0}^{m} F_k (m+1-k) G_{m+1-k} + \sum_{j=0}^{m} (j+1) F_{j+1} G_{m-j}.
\end{aligned}$$

In the first summation, we can let k go from 0 to $m+1$ (which adds a zero term). In the second summation, change the summation variable to $k = j+1$ and add a zero term corresponding to $k = 0$. We get

$$\begin{aligned}
(FG' + F'G)_m &= \sum_{k=0}^{m+1} (m+1-k) F_k G_{m+1-k} + \sum_{k=0}^{m+1} k F_k G_{m+1-k} \\
&= \sum_{k=0}^{m+1} (m+1) F_k G_{m+1-k} = (FG)'_m.
\end{aligned}$$

This completes the proof of the product rule.

(f) Assume $\lim_{i \to \infty} H_i = F$. To prove $\lim_{i \to \infty} H_i' = F'$, fix $m \in \mathbb{N}$. Choose N so that $n \geq N$ implies $H_n(m+1) = F(m+1)$. By definition of formal derivatives, $n \geq N$ implies

$$H_n'(m) = (m+1)H_n(m+1) = (m+1)F(m+1) = F'(m). \qquad \square$$

A formal version of the chain rule will be given shortly, once we define formal composition of two formal power series. We turn next to a formal analogue of the Maclaurin series of a real-valued function.

7.55. Theorem: Formal Maclaurin Series. For all $F \in K[[x]]$, $F = \sum_{k\geq 0} \frac{F^{(k)}(0)}{k!} x^k$.

Proof. We have $F^{(k)} = \sum_{n\geq 0}(n+1)(n+2)\cdots(n+k)F_{n+k}x^n$, so $F^{(k)}(0) = k!F_k$. Since K contains $\mathbb{Q}$, we can divide both sides by $k!$ in K. Thus, $F_k = F^{(k)}(0)/k!$ for all $k \in \mathbb{N}$. □

7.9 Composition of Polynomials

Given functions $f, g : \mathbb{R} \to \mathbb{R}$, we can form the composite function $g \circ f : \mathbb{R} \to \mathbb{R}$ by setting $(g \circ f)(z) = g(f(z))$ for each $z \in \mathbb{R}$. We would like to introduce a version of this composition operation for formal power series. There is an immediate difficulty, since the formal series $F = \sum_{n\geq 0} F_n x^n \in K[[x]]$ is not a function of x and need not correspond to a convergent power series in the variable x. On the other hand, we saw in 7.17 that a polynomial $f \in K[x]$ can be viewed as a function $P_f : R \to R$, where R is any commutative ring containing K. So we can define the formal composition of two *polynomials* as follows.

7.56. Definition: Composition of Polynomials. Given $f, g \in K[x]$, the *formal composition of f and g* is $f \bullet g = P_f(g)$. More explicitly, if $f = \sum_{k=0}^{n} f_k x^k$ and $g = \sum_{j=0}^{m} g_j x^j$, then

$$f \bullet g = \sum_{k=0}^{n} f_k \left(\sum_{j=0}^{m} g_j x^j \right)^k \in K[x].$$

Note that the filled circle $\bullet$ denotes formal composition, whereas an open circle $\circ$ denotes ordinary composition of functions. We can also write $f \bullet g = \mathrm{ev}_g(f)$, where $\mathrm{ev}_g : K[x] \to K[x]$ is the evaluation homomorphism that sets x equal to g (see 7.22).

The following theorem shows that formal composition of polynomials satisfies properties analogous to those satisfied by ordinary composition of functions. The proof makes heavy use of the properties of evaluation homomorphisms.

7.57. Theorem: Properties of Polynomial Composition. Suppose $f, g, h \in K[x]$ are polynomials, $c \in K$, and R is a commutative ring containing K.
(a) $P_{f \bullet g} = P_f \circ P_g : R \to R$ (comparison of formal and ordinary composition).
(b) $(f \bullet g) \bullet h = f \bullet (g \bullet h)$ (associativity).
(c) $(f+g) \bullet h = (f \bullet h) + (g \bullet h)$, $(f \cdot g) \bullet h = (f \bullet h) \cdot (g \bullet h)$, $c \bullet h = c$, and $(cf) \bullet h = c(f \bullet h)$ (homomorphism properties).
(d) $(f \bullet g)' = (f' \bullet g) \cdot g'$ (formal chain rule).

Proof. (a) Given $z \in R$, we must show that $P_{f \bullet g}(z) = P_f(P_g(z))$. Let us rewrite each side in terms of evaluation homomorphisms. The left side is

$$P_{f \bullet g}(z) = \mathrm{ev}_z(f \bullet g) = \mathrm{ev}_z(\mathrm{ev}_g(f)) = (\mathrm{ev}_z \circ \mathrm{ev}_g)(f).$$

The right side is

$$P_f(P_g(z)) = P_f(\mathrm{ev}_z(g)) = \mathrm{ev}_{\mathrm{ev}_z(g)}(f).$$

The equality in (a) will follow if we can show that the two ring homomorphisms $\mathrm{ev}_z \circ \mathrm{ev}_g$

and $\mathrm{ev}_{\mathrm{ev}_z(g)}$ from $K[x]$ into R are equal. By the uniqueness part of 7.22, we need only verify that the two homomorphisms have the same effect on the polynomial x. This holds because

$$(\mathrm{ev}_z \circ \mathrm{ev}_g)(x) = \mathrm{ev}_z(\mathrm{ev}_g(x)) = \mathrm{ev}_z(g) = \mathrm{ev}_{\mathrm{ev}_z(g)}(x).$$

(b) In part (a), let R be the ring $K[x]$ and apply both sides to $z = h$. We obtain

$$(f \bullet g) \bullet h = P_{f \bullet g}(h) = P_f(P_g(h)) = P_f(g \bullet h) = f \bullet (g \bullet h).$$

(c) This is a transcription of the statement that ev_h is a ring homomorphism fixing all elements of K; for example,

$$(fg) \bullet h = \mathrm{ev}_h(fg) = \mathrm{ev}_h(f)\,\mathrm{ev}_h(g) = (f \bullet h)(g \bullet h).$$

(d) Let us first prove the special case where $f = x^n$ for some $n \geq 0$. Since $f' = nx^{n-1}$, we must show that

$$(g^n)' = ng^{n-1}g' \qquad (g \in K[x]).$$

We proceed by induction on $n \geq 0$. The base case $n = 0$ holds because both sides are zero. Assuming the formula holds for some $n \geq 0$, we use the formal product rule to calculate

$$(g^{n+1})' = (g^n g)' = (g^n)'g + g^n g' = ng^{n-1}g'g + g^n g' = (n+1)g^n g'.$$

The general case of (d) now follows because the formula is "K-linear in f." More precisely, if (d) holds for polynomials f_1 and f_2, it also holds for $f_1 + f_2$ because (by (c))

$$\begin{aligned} ((f_1 + f_2) \bullet g)' &= ((f_1 \bullet g) + (f_2 \bullet g))' = (f_1 \bullet g)' + (f_2 \bullet g)' \\ &= (f_1' \bullet g)g' + (f_2' \bullet g)g' = (f_1' \bullet g + f_2' \bullet g)g' \\ &= ((f_1' + f_2') \bullet g)g' = ((f_1 + f_2)' \bullet g)g'. \end{aligned}$$

Similarly, if (d) holds for some $f \in K[x]$, then (d) holds for cf whenever $c \in K$. Since every polynomial is a finite K-linear combination of powers of x, it follows that (d) is true for all polynomials f, as desired. □

7.10 Composition of Formal Power Series

We can extend the definition of the formal composition $f \bullet g$ to the case where $f \in K[x]$ is a polynomial and $G \in K[[x]]$ is a formal series by setting $f \bullet G = P_f(G) = \mathrm{ev}_G(f)$, as in 7.56. A more challenging problem is to define the composition $F \bullet G$ when F and G are both formal series. To see what can go wrong, suppose $F = \sum_{n \geq 0} F_n x^n$ and $G = \sum_{m \geq 0} G_m x^m$ are formal series. By analogy with the preceding definitions, we would like to define

$$F \bullet G = \sum_{n \geq 0} F_n G^n = \sum_{n \geq 0} F_n \left(\sum_{k \geq 0} G_k x^k \right)^n \in K[[x]].$$

The trouble is that the infinite sum of formal series $\sum_{n \geq 0} F_n G^n$ may not be well-defined. Indeed, if $F_n \neq 0$ for infinitely many values of n and $G_0 = 1$, consideration of the constant term shows that $\sum_{n \geq 0} F_n G^n$ does not exist. However, we can escape this difficulty by requiring that the right-hand factor in a formal composition $F \bullet G$ have zero constant term. This leads to the following definition.

7.58. Definition: Composition of Formal Power Series. Given $F, G \in K[[x]]$ with $G(0) = 0$, the *formal composition of F and G* is

$$F \bullet G = \sum_{n \geq 0} F_n G^n \in K[[x]].$$

This infinite sum converges by 7.33, because $\text{ord}(F_n G^n) \geq n$ whenever $F_n G^n \neq 0$ and so the orders of the nonzero summands go to ∞ as $n \to \infty$.

7.59. Theorem: Identity for Formal Composition. For all $F \in K[[x]]$, $F \bullet x = F$. For all $G \in K[[x]]$ with $G(0) = 0$, $x \bullet G = G$.

Proof. For any $F \in K[[x]]$, $F \bullet x = \sum_{n \geq 0} F_n x^n = F$. Next, recall that $x = X_1 = \sum_{n \geq 0} \chi(n = 1) x^n$. So, for $G \in K[[x]]$ with $G(0) = 0$,

$$x \bullet G = \sum_{n \geq 0} \chi(n = 1) G^n.$$

One may check that this infinite sum of formal series converges to $G^1 = G$. □

The next technical result will aid us in proving further facts about formal composition.

7.60. Theorem: Coefficient in a Composition. For all $F, G \in K[[x]]$ with $G(0) = 0$ and all $m \in \mathbb{N}$,

$$(F \bullet G)_m = \left(\sum_{n=0}^{m} F_n G^n \right)_m.$$

Proof. Since $(F \bullet G)_m = \left(\sum_{n=0}^{\infty} F_n G^n \right)_m$, it suffices to show that

$$\left(\sum_{n=0}^{p} F_n G^n \right)_m = \left(\sum_{n=0}^{p+1} F_n G^n \right)_m$$

for all $p \geq m$. This holds since

$$\left(\sum_{n=0}^{p+1} F_n G^n \right)_m = \left(\sum_{n=0}^{p} F_n G^n \right)_m + (F_{p+1} G^{p+1})_m,$$

and $F_{p+1} G^{p+1}$ is either zero or has order at least $p + 1 > m$ (since $\text{ord}(G) \geq 1$). □

7.61. Theorem: Joint Continuity of Formal Composition. Suppose $F_n, G_n, P, Q \in K[[x]]$ are formal series such that $G_n(0) = 0$ for all $n \in \mathbb{N}$, $F_n \to P$, and $G_n \to Q$ (forcing $Q(0) = 0$). Then $F_n \bullet G_n \to P \bullet Q$.

Proof. Fix $m \in \mathbb{N}$; we must show that $(P \bullet Q)(m) = (F_n \bullet G_n)(m)$ for all sufficiently large n. By 7.60,

$$(P \bullet Q)(m) = \left(\sum_{i=0}^{m} P(i) Q^i \right)_m; \tag{7.4}$$

$$(F_n \bullet G_n)(m) = \left(\sum_{i=0}^{m} F_n(i) G_n^i \right)_m. \tag{7.5}$$

Now, for each fixed $i \leq m$, iteration of 7.35 shows that $G_n^i \to Q^i$ as $n \to \infty$. For each

fixed i, the sequence of (order zero) power series $F_n(i)$ converges to $P(i)$ as $n \to \infty$, since $F_n \to P$. Using 7.35 again, we see that $F_n(i)G_n^i \to P(i)Q^i$ for each fixed $i \leq m$, and hence (by 7.35)

$$\lim_{n\to\infty} \left(\sum_{i=0}^{m} F_n(i) G_n^i \right) = \sum_{i=0}^{m} P(i) Q^i.$$

It follows that the right sides of (7.4) and (7.5) do agree for large enough n, which is what we needed to show. □

The previous result combined with 7.32 allows us to use "continuity arguments" to deduce properties of composition of formal power series from corresponding properties of composition of polynomials. The next few theorems illustrate this technique.

7.62. Theorem: Homomorphism Properties of Formal Composition.
Let $G \in K[[x]]$ satisfy $G(0) = 0$.
(a) For all $F, H \in K[[x]]$, $(F + H) \bullet G = (F \bullet G) + (H \bullet G)$.
(b) For all $F, H \in K[[x]]$, $(FH) \bullet G = (F \bullet G)(H \bullet G)$.
(c) For all $c \in K$, $c \bullet G = c$.
So, the *evaluation map* $\mathrm{ev}_G : K[[x]] \to K[[x]]$ given by $\mathrm{ev}_G(F) = F \bullet G$ is a ring homomorphism fixing K and sending x to G.

Proof. We prove (a), leaving (b) and (c) as exercises. Fix $F, G, H \in K[[x]]$ with $G(0) = 0$. Use 7.32 to choose polynomials $f_n, g_n, h_n \in K[x]$ with $g_n(0) = 0$ for all n, $f_n \to F$, $g_n \to G$, and $h_n \to H$. For each $n \in \mathbb{N}$, we know from 7.57(c) that

$$(f_n + h_n) \bullet g_n = (f_n \bullet g_n) + (h_n \bullet g_n).$$

Take the limit of both sides as $n \to \infty$. Using 7.35 and 7.61, we get $(F + H) \bullet G = (F \bullet G) + (H \bullet G)$ as desired. □

7.63. Theorem: Associativity of Formal Composition. Suppose $F, G, H \in K[[x]]$ satisfy $G(0) = 0 = H(0)$. Then

$$(F \bullet G) \bullet H = F \bullet (G \bullet H).$$

Proof. First note that all compositions in the theorem statement are defined; in particular, $F \bullet (G \bullet H)$ is defined because $G \bullet H$ has zero constant term. Use 7.32 to choose polynomials $f_n, g_n, h_n \in K[x]$ with $g_n(0) = 0 = h_n(0)$ for all n, $f_n \to F$, $g_n \to G$, and $h_n \to H$. For each $n \in \mathbb{N}$, we know from 7.57(b) that $(f_n \bullet g_n) \bullet h_n = f_n \bullet (g_n \bullet h_n)$. Taking limits and using 7.61 repeatedly gives the desired result. □

A similar continuity argument (left as an exercise) establishes the following differentiation rule.

7.64. Theorem: Formal Chain Rule. For all $F, G \in K[[x]]$ with $G(0) = 0$,

$$(F \bullet G)' = (F' \bullet G)G'.$$

7.65. Theorem: Inverses for Formal Composition. Let $S = \{F \in K[[x]] : F(0) = 0 \text{ and } F(1) \neq 0\}$.
(a) If $F, G \in S$, then $F \bullet G \in S$ (closure of S).
(b) If $F \in S$, there exists a unique $G \in S$ with $F \bullet G = x = G \bullet F$ (inverses).

(Together with 7.59 and 7.63, this proves that $(S, \bullet)$ is a *group* as defined in 9.1. The proof will show that if $F, G \in S$ and $G \bullet F = x$, then $F \bullet G = x$ automatically follows.)

Proof. (a) Suppose F and G belong to S. On one hand, since $F(0) = 0 = G(0)$, $F \bullet G$ is defined and also has zero constant term. On the other hand, 7.60 gives $(F \bullet G)_1 = F_1 G_1^1 \neq 0$. So $F \circ G \in S$.

(b) First we prove that for each $F \in S$, there exists a unique $G \in S$ with $G \bullet F = x$ (we call G a "left inverse" of F). By 7.60,

$$(G \bullet F)_n = \left(\sum_{m=0}^{n} G_m F^m \right)_n \qquad (n \in \mathbb{N}).$$

We can use this equation to give a recursive prescription for the coefficients G_n. At the start, we must set $G_0 = 0$ and $G_1 = 1/F_1 \neq 0$. (Note $1/F_1$ exists because F_1 is a nonzero element of the *field* K.) Assume $n > 1$ and $G_0, G_1, \ldots, G_{n-1}$ have already been determined. Since $(G_n F^n)_n = G_n (F_1)^n$, we need to choose G_n so that

$$0 = \left(\sum_{m=0}^{n} G_m F^m \right)_n = G_n F_1^n + \left(\sum_{m=0}^{n-1} G_m F^m \right)_n .$$

Evidently there is a unique $G_n \in K$ that will work, namely

$$G_n = -\frac{1}{F_1^n} \left(\sum_{m=0}^{n-1} G_m F^m \right)_n . \tag{7.6}$$

Since $G \in S$, we have shown that F has a unique left inverse in S.

To finish the proof, fix $F \in S$. Let G be the left inverse of F, and let H be the left inverse of G. Then

$$H = H \bullet x = H \bullet (G \bullet F) = (H \bullet G) \bullet F = x \bullet F = F.$$

Since $H = F$, we see that both $G \bullet F$ and $F \bullet G = H \bullet G$ equal the identity element x. Thus, G is the two-sided inverse of F. □

7.66. Remark. Lagrange's inversion formula (8.15) provides an alternate way to determine the coefficients of the compositional inverse of a formal series F, which is sometimes easier to use than the recursive formula for G_n in the preceding proof.

7.67. Example. Consider the series $E = e^x - 1 = \sum_{n \geq 1} x^n/n!$ and $L = \log(1 + x) = \sum_{n \geq 1} (-1)^{n-1} x^n/n$. Let us show that L is the two-sided inverse of E relative to formal composition. Set $H = L \bullet E$; since $H(0) = 0$, it will suffice to prove that $H' = 1$ (cf. 7.132). First, a routine formal differentiation shows that $E' = e^x = 1 + E$ and $L' = 1 - x + x^2 - x^3 + \cdots = (1 + x)^{-1}$. We also have $L' \bullet E = 1 - E + E^2 - E^3 + \cdots = (1 + E)^{-1}$ by 7.41. The formal chain rule now gives

$$H' = (L \bullet E)' = (L' \bullet E) E' = (1 + E)^{-1} (1 + E) = 1.$$

We conclude that $L \bullet E = x$, hence also $E \bullet L = x$.

7.11 Generalized Binomial Expansion

Given a formal power series $F \in K[[x]]$, we would like to give meaning to expressions like $F^{1/2} = \sqrt{F}$. To prepare for this, we will first define, for each $r \in K$, a power series $\mathrm{Pow}_r \in K[[x]]$ that is a formal analogue of the function $x \mapsto (1+x)^r$. The following example from calculus motivates the definition of Pow_r.

7.68. Example. Consider the real-valued function $f(x) = (1+x)^r$, where r is a fixed real constant, and x ranges over real numbers > -1. *If* f has a Taylor series expansion about $x = 0$, then the coefficients of the power series must be given by Taylor's formula

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n$$

(cf. 7.55). Computing the successive derivatives of f, we find $f'(x) = r(1+x)^{r-1}$, $f''(x) = r(r-1)(1+x)^{r-2}$, and in general

$$f^{(n)}(x) = r(r-1)(r-2)\cdots(r-n+1)(1+x)^{r-n} = (r)\downarrow_n (1+x)^{r-n}$$

(here we use the falling factorial notation from 2.76). Evaluating the derivatives at $x = 0$, we conclude that

$$(1+x)^r = \sum_{n=0}^{\infty} \frac{(r)\downarrow_n}{n!} x^n$$

for x close enough to zero, *provided* $f(x)$ converges to its Taylor series expansion. One can prove convergence by bounding the remainder term in Taylor's theorem. We will omit the details since they are not needed in the formal setting considered here.

Motivated by the formula in the previous example, we define the following series to model the function $(1+x)^r$.

7.69. Definition: Falling Factorials and Pow_r. For every $r \in K$ and every integer $n \geq 1$, define the *falling factorial*

$$(r)\downarrow_n = r(r-1)(r-2)\cdots(r-n+1) \in K.$$

Let $(r)\downarrow_0 = 1$. Define the formal power series

$$\mathrm{Pow}_r = \sum_{n=0}^{\infty} \frac{(r)\downarrow_n}{n!} x^n \in K[[x]].$$

(This definition uses the assumption that K is a field containing $\mathbb{Q}$.)

7.70. Example: Pow_r for Integral r. Suppose $r \in \mathbb{N} \subseteq K$. Let us show that $\mathrm{Pow}_r = (1+x)^r$ (the product of r copies of the series $1+x$) in this case. First note that $(r)\downarrow_n$ reduces to the binomial coefficient $\binom{r}{n} = \frac{r!}{n!(r-n)!}$ when r is a nonnegative integer (and this coefficient is zero for $n > r$). Next, invoking the binomial theorem 2.14 in the commutative ring $K[[x]]$, we see that

$$(1+x)^r = \sum_{n=0}^{r} \binom{r}{n} x^n = \sum_{n=0}^{\infty} \frac{(r)\downarrow_n}{n!} x^n = \mathrm{Pow}_r\,.$$

Similarly, if $r = -1$, the definition of falling factorials shows that $(r)\downarrow_n / n! = (-1)^n$ for all $n \geq 0$. On the other hand, we have seen that the multiplicative inverse of $1+x$ in $K[[x]]$ is

$$(1+x)^{-1} = 1 - x + x^2 - x^3 + \cdots = \sum_{n=0}^{\infty} (-1)^n x^n = \sum_{n=0}^{\infty} \frac{(r)\downarrow_n}{n!} x^n = \mathrm{Pow}_{-1}\,.$$

So $\mathrm{Pow}_r = (1+x)^r$ is also true when $r = -1$. We will see in a moment that the same result holds for all negative integers r.

If x, r, s are *real numbers*, we have the familiar law of exponents: $(1+x)^{r+s} = (1+x)^r \cdot (1+x)^s$. We now prove the formal analogue of this result.

7.71. Theorem: Formal Exponent Law for Pow_r. For all $r, s \in K$, $\mathrm{Pow}_{r+s} = \mathrm{Pow}_r \mathrm{Pow}_s$.

Proof. We show $\mathrm{Pow}_{r+s}(n) = (\mathrm{Pow}_r \mathrm{Pow}_s)(n)$ for each $n \geq 0$. On one hand, $\mathrm{Pow}_{r+s}(n) = (r+s)\downarrow_n / n!$. On the other hand,

$$(\mathrm{Pow}_r \mathrm{Pow}_s)(n) = \sum_{k=0}^{n} \mathrm{Pow}_r(k)\,\mathrm{Pow}_s(n-k) = \sum_{k=0}^{n} \frac{(r)\downarrow_k}{k!}\frac{(s)\downarrow_{n-k}}{(n-k)!}.$$

Comparing these expressions, we see that we must prove the identities

$$(r+s)\downarrow_n = \sum_{k=0}^{n} \binom{n}{k} (r)\downarrow_k \, (s)\downarrow_{n-k}$$

for all $n \geq 0$. We use induction on n. When $n = 0$, the identity reads $1 = \binom{0}{0} \cdot 1 \cdot 1$, which is true. Assume the identity holds for some $n \geq 0$. Using the recursion $\binom{n+1}{j} = \binom{n}{j-1} + \binom{n}{j}$, we compute

$$\begin{aligned}
(r+s)\downarrow_{n+1} &= (r+s)\downarrow_n \, (r+s-n) \\
&= \sum_{k=0}^{n} \binom{n}{k} (r)\downarrow_k \, (s)\downarrow_{n-k} \, ((r-k) + (s-(n-k))) \\
&= \sum_{k=0}^{n} \binom{n}{k} (r)\downarrow_{k+1} \, (s)\downarrow_{n-k} + \sum_{k=0}^{n} \binom{n}{k} (r)\downarrow_k \, (s)\downarrow_{n-k+1} \\
&= \sum_{j=1}^{n+1} \binom{n}{j-1} (r)\downarrow_j \, (s)\downarrow_{n+1-j} + \sum_{j=0}^{n} \binom{n}{j} (r)\downarrow_j \, (s)\downarrow_{n+1-j} \\
&= \sum_{j=0}^{n+1} \binom{n+1}{j} (r)\downarrow_j \, (s)\downarrow_{n+1-j} \, .
\end{aligned}$$

In the next-to-last step, we changed summation variables to $j = k+1$ in the first sum, and $j = k$ in the second sum. The reader should check that the last equality is valid even for the extreme terms $j = 0$ and $j = n+1$. This completes the induction argument. □

7.72. Theorem: Negative Binomial Formula. For every integer $r > 0$,

$$(1+x)^{-r} = \mathrm{Pow}_{-r} = \sum_{n=0}^{\infty} \binom{r+n-1}{n, r-1} (-1)^n x^n \in K[[x]].$$

Proof. The first equality follows by iterating 7.71 r times, recalling that $(1+x)^{-1} = \mathrm{Pow}_{-1}$. The second equality follows from 7.69 and the identity

$$\begin{aligned}
\frac{(-r)\downarrow_n}{n!} &= \frac{(-r)(-r-1)(-r-2)\cdots(-r-(n-1))}{n!} \\
&= (-1)^n \frac{(r+n-1)\downarrow_n}{n!} = \binom{r+n-1}{n} (-1)^n. \quad \square
\end{aligned}$$

We have now shown that $\mathrm{Pow}_r = (1+x)^r$ holds for *all* integers r. So we can introduce the following notation for Pow_r without risk of ambiguity.

7.73. Definition: $(1+x)^r$. For any $r \in K$, let $(1+x)^r$ denote the series $\mathrm{Pow}_r \in K[[x]]$.

7.12 Generalized Powers of Formal Series

We now have the necessary tools to define operations such as $\sqrt{F}$, provided the formal power series F satisfies suitable hypotheses.

7.74. Definition: Generalized Powers. Suppose $F \in K[[x]]$ has $F(0) = 1$, and $r \in K$. Let F^r be the composition $\mathrm{Pow}_r \bullet (F-1)$, which is defined since $F-1$ has zero constant term.

Informally, $\mathrm{Pow}_r \bullet (F-1) = (1+(F-1))^r$, so this definition is reasonable. Observe that F^r always has constant term 1. When $r = 1/n$ for n a positive integer, we also write $\sqrt[n]{F}$ to denote $F^{1/n}$.

Many familiar rules for manipulating powers remain true in the formal setting, but they must be reproved formally before they can be used.

7.75. Theorem: Properties of Formal Powers. Suppose $F \in K[[x]]$ has $F(0) = 1$. For any $r, s \in K$, $F^{r+s} = F^r \cdot F^s$. Furthermore, when r is an integer, F^r (as defined in 7.74) coincides with the customary algebraic definition of F^r (namely the product of r copies of F for $r \geq 0$, or $|r|$ copies of $1/F$ for $r < 0$).

Proof. Recall from 7.71 that $\mathrm{Pow}_{r+s} = \mathrm{Pow}_r \mathrm{Pow}_s$. Using the fact that composing on the right by $F-1$ is a ring homomorphism (see 7.62(b)), we obtain

$$\begin{aligned} F^{r+s} &= \mathrm{Pow}_{r+s} \bullet (F-1) = (\mathrm{Pow}_r \mathrm{Pow}_s) \bullet (F-1) \\ &= (\mathrm{Pow}_r \bullet (F-1))(\mathrm{Pow}_s \bullet (F-1)) = F^r F^s. \end{aligned}$$

For the rest of this proof, n will denote an integer and F^n will have the usual algebraic meaning (repeated multiplication). We must prove $\mathrm{Pow}_n \bullet (F-1) = F^n$ for all $n \in \mathbb{Z}$. When $n = 0$, $\mathrm{Pow}_0 \bullet (F-1) = 1 \bullet (F-1) = 1 = F^0$. When $n = 1$, $\mathrm{Pow}_1 \bullet (F-1) = (1+x) \bullet (F-1) = 1 + (F-1) = F = F^1$. By induction, assuming $n \geq 1$ and $\mathrm{Pow}_n \bullet (F-1) = F^n$, the result just proved shows that

$$\mathrm{Pow}_{n+1} \bullet (F-1) = (\mathrm{Pow}_n \bullet (F-1))(\mathrm{Pow}_1 \bullet (F-1)) = F^n F^1 = F^{n+1}.$$

(The last step uses the algebraic definition of the power F^{n+1}.) Similarly, the known identity

$$(\mathrm{Pow}_{-n} \bullet (F-1))(\mathrm{Pow}_n \bullet (F-1)) = \mathrm{Pow}_0 \bullet (F-1) = 1$$

shows that $\mathrm{Pow}_{-n} \bullet (F-1)$ is the multiplicative inverse of $\mathrm{Pow}_n \bullet (F-1) = F^n$. In other words, $\mathrm{Pow}_{-n} \bullet (F-1) = (F^n)^{-1} = F^{-n}$. □

7.76. Example: Negative Binomial Expansion. Suppose $F = 1 - cx$ where $c \in K$ is a constant. Let r be a positive integer. Using 7.72 and the definition of composition, we find that

$$(1-cx)^{-r} = \left(\frac{1}{1-cx}\right)^r = \sum_{n=0}^{\infty} \binom{n+r-1}{n, r-1} c^n x^n. \tag{7.7}$$

This identity is used often when computing with generating functions.

Next we prove a partial version of another familiar law of exponents.

7.77. Theorem: Iterated Exponents. For all $F \in K[[x]]$, all $r \in K$ and all *integers* n, $(F^r)^n = F^{rn}$.

Proof. The idea is to iterate the known identity $F^{r+s} = F^r F^s$, which holds for all $r, s \in K$. First, we prove the result for integers $n \geq 0$ by induction. When $n = 0$, $(F^r)^0 = 1 = F^{r0}$. When $n = 1$, $(F^r)^1 = F^r = F^{r1}$. Assuming $n \geq 1$ and $(F^r)^n = F^{rn}$ is already known, we calculate

$$(F^r)^{n+1} = (F^r)^n (F^r)^1 = F^{rn} F^r = F^{rn+r} = F^{r(n+1)}.$$

Next, $(F^r)^{-n}(F^r)^n = (F^r)^{-n+n} = 1$, and hence $(F^r)^{-n} = ((F^r)^n)^{-1} = (F^{rn})^{-1}$. Similarly, $F^{-rn}F^{rn} = F^{-rn+rn} = 1$, so that $(F^{rn})^{-1} = F^{-rn}$. So finally $(F^r)^{-n} = F^{r(-n)}$, which establishes the result for negative integers. □

The next result is the analogue of the fact that every positive real number has a unique positive nth root, for each $n \geq 0$.

7.78. Theorem: Existence and Uniqueness of nth Roots. Suppose $F \in K[[x]]$ satisfies $F(0) = 1$. For every integer $n \geq 1$, there exists a unique $G \in K[[x]]$ with $G(0) = 1$ such that $G^n = F$, namely $G = F^{1/n} = \sqrt[n]{F}$.

Proof. Existence of G follows from the previous result, since $(F^{1/n})^n = F^{(1/n)\cdot n} = F^1 = F$. To prove uniqueness, suppose $G, H \in K[[x]]$ satisfy $G^n = F = H^n$ and $G(0) = 1 = H(0)$. By the distributive law, we have the factorization

$$0 = G^n - H^n = (G - H)(G^{n-1} + G^{n-2}H^1 + G^{n-3}H^2 + \cdots + H^{n-1}).$$

Since $K[[x]]$ is an integral domain, this implies that either $G - H = 0$ or $\sum_{i=0}^{n-1} G^{n-1-i}H^i = 0$. The first alternative gives $G = H$, as desired. The second alternative is impossible, since the left side has constant term $\sum_{i=0}^{n-1} 1^{n-1-i}1^i = n > 0$ while the right side has constant term zero. (This proof uses the assumption that K is a field containing $\mathbb{Q}$.) □

7.79. Theorem: Formal Power Rule. Suppose $F \in K[[x]]$ satisfies $F(0) = 1$. For $r \in K$,

$$(F^r)' = rF^{r-1}F'.$$

Proof. Let us first show that $\mathrm{Pow}_r' = r\,\mathrm{Pow}_{r-1}$. The formal derivative of $\mathrm{Pow}_r = \sum_{n\geq 0}((r)\downarrow_n / n!)x^n$ is

$$\mathrm{Pow}_r' = \sum_{n\geq 0} \frac{(n+1)(r)\downarrow_{n+1}}{(n+1)!}x^n = \sum_{n\geq 0} \frac{r(r-1)\downarrow_n}{n!}x^n = r\,\mathrm{Pow}_{r-1}\,.$$

It follows that

$$(F^r)' = [\mathrm{Pow}_r \bullet (F-1)]' = [\mathrm{Pow}_r' \bullet (F-1)](F-1)' = [r\,\mathrm{Pow}_{r-1}' \bullet (F-1)]F' = rF^{r-1}F'. \quad \square$$

We now have the necessary machinery to solve certain quadratic equations involving formal power series.

7.80. Example. Suppose $F \in \mathbb{Q}[[x]]$ is a formal series such that $xF^2 - F + 1 = 0$. Let us "solve for F," determining F_n for all $n \in \mathbb{N}$. The given equation immediately implies $F_0 = 1$. Multiplying the quadratic equation by $4x$ gives $4x^2F^2 - 4xF + 4x = 0$. Completing the square leads to $(1 - 2xF)^2 = 1 - 4x$. Since $\sqrt{1-4x}$ is the *unique* power series with constant term 1 that squares to $1 - 4x$, we conclude that

$$1 - 2xF = \sqrt{1-4x}.$$

Rearranging, $xF = \frac{1}{2}(1 - \sqrt{1-4x})$. Since the power series on the right has zero constant term, we may safely write

$$F = \frac{1 - \sqrt{1-4x}}{2x},$$

where the notation $\frac{1}{x}\sum_{n\geq 1} a_n x^n$ can be regarded as shorthand for the series $\sum_{n\geq 0} a_{n+1}x^n$. (Note x is not a unit in $\mathbb{Q}[[x]]$, so algebraic division by x is not permitted in this ring, although we could allow it by passage to the field of fractions (§7.7). What we are really doing is cancelling the nonzero element x in the integral domain $\mathbb{Q}[[x]]$; cf. 7.135.)

We remark that our formula for F is exactly what we would have obtained by formally applying the quadratic formula for solving $AF^2 + BF + C = 0$, which gives $F = \frac{-B\pm\sqrt{B^2-4AC}}{2A}$. However, one must take care in blindly applying this formula since we cannot divide by arbitrary formal power series, and the sign ambiguity in the square root must be resolved somehow. In our example, we are forced to choose the minus sign.

Now we are ready to find F_n for each $n \geq 0$. By 7.74,

$$\sqrt{1-4x} = \sum_{m=0}^{\infty} \frac{(1/2)\downarrow_m}{m!}(-4x)^m.$$

For $m \geq 1$, the coefficient of x^m here is

$$\begin{aligned}(-1)^m 4^m \frac{(1/2)(-1/2)(-3/2)\cdots(-(2m-3)/2)}{m!} &= -\frac{2^m \cdot 1\cdot 1\cdot 3\cdot 5\cdot \cdots \cdot (2m-3)}{m!} \\ &= -\frac{2^m(2m-2)!}{m!\cdot 2\cdot 4\cdot 6\cdot \cdots\cdot (2m-2)} \\ &= -2\frac{(2m-2)!}{m!(m-1)!},\end{aligned}$$

where the last step follows by using $m-1$ powers of 2 in the numerator to divide each of the even numbers in the denominator by 2. It now follows that

$$\left(\frac{1-\sqrt{1-4x}}{2}\right)_m = \frac{1}{2m-1}\binom{2m-1}{m, m-1} \qquad (m \geq 1).$$

Finally,

$$F_n = \left(\frac{1-\sqrt{1-4x}}{2}\right)_{n+1} = \frac{1}{2n+1}\binom{2n+1}{n+1, n} \qquad (n \geq 0).$$

Thus, $F = \sum_{n\geq 0} C_n x^n$ is the generating function for the Catalan numbers (see 1.55).

Calculations such as those in the preceding example occur frequently in the application of formal power series to combinatorial problems.

7.13 Partial Fraction Expansions

Suppose g is a polynomial in $\mathbb{C}[x]$ with nonzero constant term, and f is any polynomial in $\mathbb{C}[x]$. We have seen (§7.6) that g is a unit in $\mathbb{C}[[x]]$, so that we can write $f/g = \sum_{n=0}^{\infty} b_n x^n$ for suitable complex numbers b_n. This section presents a technique for finding explicit expressions for the coefficients b_n, which is a formal version of the "method of partial fractions" from calculus. We will see that this technique can be used to find explicit closed formulas for certain recursively defined sequences. Our starting point is the famous fundamental theorem of algebra, which we state here without proof.

7.81. Fundamental Theorem of Algebra. Let $p \in \mathbb{C}[x]$ be a monic polynomial of degree $n \geq 1$. There exist pairwise distinct complex numbers $r_1, \ldots, r_k$ (unique up to reordering) and unique positive integers $n_1, \ldots, n_k$ such that

$$p = (x - r_1)^{n_1}(x - r_2)^{n_2} \cdots (x - r_k)^{n_k} \in \mathbb{C}[x].$$

The number r_i is called a *root* of p of *multiplicity* n_i.

The following variant of the fundamental theorem is needed in partial fraction problems because of the form of the negative binomial expansion (see 7.76).

7.82. Theorem: Factorization of Polynomials in $\mathbb{C}[x]$. Let $p \in \mathbb{C}[x]$ be a polynomial of degree $n \geq 1$ with $p(0) = 1$. There exist pairwise distinct, nonzero complex numbers $r_1, \ldots, r_k$ and positive integers $n_1, \ldots, n_k$ such that

$$p(x) = (1 - r_1 x)^{n_1}(1 - r_2 x)^{n_2} \cdots (1 - r_k x)^{n_k} \in \mathbb{C}[x].$$

Proof. Consider the polynomial $q = x^n P_p(1/x)$. We have $p = \sum_{i=0}^{n} p_i x^i$ and $q = \sum_{i=0}^{n} p_{n-i} x^i$, so that q is obtained from p by "reversing the coefficient sequence." Since $p_0 = 1$, q is a monic polynomial of degree n. Using the fundamental theorem of algebra, we write

$$x^n P_p(1/x) = q = \prod_{i=1}^{k} (x - r_i)^{n_i},$$

where $\sum n_i = n$. Since the constant term of q is nonzero, no r_i is equal to zero. Reversing the coefficient sequence again, it follows that

$$p = x^n P_q(1/x) = x^n \prod_{i=1}^{k} ((1/x) - r_i)^{n_i} = \prod_{i=1}^{k} x^{n_i} \left(\frac{1 - r_i x}{x} \right)^{n_i} = \prod_{i=1}^{k} (1 - r_i x)^{n_i}. \qquad \square$$

The next step is to rewrite a general fraction f/g as a sum of fractions whose denominators have the form $(1 - rx)^m$. Note that, as long as $g(0) \neq 0$, we can always arrange $g(0) = 1$ by multiplying numerator and denominator by a suitable scalar in K.

7.83. Theorem: Splitting a Denominator. Suppose $f, g \in \mathbb{C}[x]$ are polynomials such that $g(0) = 1$, and let g have factorization $g(x) = \prod_{i=1}^{k} (1 - r_i x)^{n_i}$, where $r_1, \ldots, r_k \in \mathbb{C}$ are distinct and nonzero. There exist polynomials $p_0, p_1, \ldots, p_k$ with $\deg(p_i) < n_i$ (or $p_i = 0$) for $1 \leq i \leq k$, such that

$$\frac{f}{g} = p_0 + \sum_{i=1}^{k} \frac{p_i}{(1 - r_i x)^{n_i}}.$$

Proof. For $1 \leq i \leq k$, define a polynomial $h_i = g/(1 - r_i x)^{n_i} = \prod_{j:j \neq i} (1 - r_j x)^{n_j}$. Since $r_1, \ldots, r_k$ are distinct, $\gcd(h_1, \ldots, h_k) = 1$. By a well-known result from polynomial algebra, it follows that there exist polynomials $q_1, \ldots, q_k \in \mathbb{C}[x]$ with $q_1 h_1 + \cdots + q_k h_k = 1$. Therefore,

$$\begin{aligned} \frac{f}{g} &= \frac{f \cdot 1}{g} = \frac{f q_1 h_1 + \cdots + f q_k h_k}{g} \\ &= \sum_{i=1}^{k} \frac{f q_i}{(1 - r_i x)^{n_i}}. \end{aligned}$$

This is almost the answer we want, but the degrees of the numerators may be too high. Using

polynomial division (see 5.87), we can write $fq_i = a_i(1 - r_i x)^{n_i} + p_i$ where $a_i, p_i \in \mathbb{C}[x]$, and either $p_i = 0$ or $\deg(p_i) < n_i$. Dividing by $(1 - r_i x)^{n_i}$, we see that

$$\frac{f}{g} = p_0 + \sum_{i=1}^{k} \frac{p_i}{(1 - r_i x)^{n_i}}$$

holds if we take $p_0 = \sum_{i=1}^{k} a_i \in \mathbb{C}[x]$. □

The fractions $p_i/(1 - r_i x)^{n_i}$ (with $\deg(p_i) < n_i$ or $p_i = 0$) can be further reduced into sums of fractions where the numerators are complex constants.

7.84. Theorem: Division by $(1 - rx)^n$. Given a fraction $p/(1 - rx)^n$ where $p \in \mathbb{C}[x]$, $\deg(p) < n$ (or $p = 0$), and $0 \neq r \in \mathbb{C}$, there exist complex numbers $a_1, \ldots, a_n$ such that

$$\frac{p}{(1 - rx)^n} = \sum_{i=1}^{n} \frac{a_i}{(1 - rx)^i}.$$

Proof. Consider the evaluation homomorphism $E : \mathbb{C}[x] \to \mathbb{C}[x]$ such that $E(x) = 1 - rx$ (see 7.22). The evaluation homomorphism $E' : \mathbb{C}[x] \to \mathbb{C}[x]$ such that $E'(x) = (1-x)/r$ is a two-sided inverse to E (since $E(E'(x)) = x = E'(E(x))$), so E is a bijection. In particular, E is surjective, so $p = E(q)$ for some $q \in \mathbb{C}[x]$. Now, one may check that E and E' each map polynomials of degree $< n$ to polynomials of degree $< n$, and it follows that $\deg(q) < n$ (or $q = 0$). Write $q = c_0 + c_1 x + c_2 x^2 + \cdots + c_{n-1} x^{n-1}$, with $c_i \in \mathbb{C}$. Then

$$p = E(q) = c_0 + c_1(1 - rx) + c_2(1 - rx)^2 + \cdots + c_{n-1}(1 - rx)^{n-1}.$$

Dividing by $(1 - rx)^n$, we see that we may take $a_1 = c_{n-1}, \ldots, a_{n-1} = c_1, a_n = c_0$. □

The next result summarizes the partial fraction manipulations in the last two theorems. The uniqueness proof given below also provides a convenient algorithm for finding the coefficients in the partial fraction decomposition.

7.85. Theorem: Partial Fraction Decompositions in $\mathbb{C}(x)$. Suppose $f, g \in \mathbb{C}[x]$ are polynomials with $g(0) = 1$; let $g = \prod_{i=1}^{k}(1 - r_i x)^{n_i}$ where the r_i are distinct nonzero complex numbers. There exist a unique polynomial $h \in \mathbb{C}[x]$ and unique complex numbers a_{ij} (where $1 \leq i \leq k, 1 \leq j \leq n_i$) with

$$\frac{f}{g} = h + \sum_{i=1}^{k} \sum_{j=1}^{n_i} \frac{a_{ij}}{(1 - r_i x)^j}. \tag{7.8}$$

Viewing $f/g \in \mathbb{C}[[x]]$, we have (for all $m \in \mathbb{N}$)

$$(f/g)_m = h_m + \sum_{i=1}^{k} \sum_{j=1}^{n_i} a_{ij} \binom{m + j - 1}{m, j - 1} r_i^m.$$

Proof. Existence of the decomposition follows by combining 7.83 and 7.84. The formula for the coefficient of x^m follows from the negative binomial expansion 7.76. We must now prove uniqueness of h and the a_{ij}'s. Note first that the numbers r_i and n_i appearing in the factorization of g are unique (this follows from the uniqueness assertion in the fundamental theorem of algebra). Now consider any expression of the form (7.8). Multiplying both sides by g produces an equation

$$f = gh + \sum_{i=1}^{k} \sum_{j=1}^{n_i} a_{ij}(1 - r_i x)^{n_i - j} \prod_{s \neq i} (1 - r_s x)^{n_s}, \tag{7.9}$$

where both sides are polynomials. Furthermore, the terms in the double sum add up to a polynomial that is either zero or has degree less than $\deg(g)$. Thus h must be the quotient when f is divided by g using the polynomial division algorithm, and this quotient is known to be unique. Next, we show how to recover the "top coefficients" a_{i,n_i} for $1 \le i \le k$. Fix i, and apply the functions associated to the polynomials on each side of (7.9) to $z = 1/r_i \in \mathbb{C}$. Since any positive power of $(1 - r_i x)$ becomes zero for this choice of x, all but one term on the right side becomes zero. We are left with

$$P_f(1/r_i) = a_{i,n_i} \prod_{s \neq i} (1 - r_s/r_i)^{n_s}.$$

Since $r_s \neq r_i$ for $s \neq i$, the product is nonzero. Thus there is a unique $a_{i,n_i} \in \mathbb{C}$ for which this equation holds. We can use the displayed formula to calculate each a_{i,n_i} given f and g.

To find the remaining a_{ij}'s, subtract the recovered summands $a_{i,n_i}/(1-r_i x)^{n_i}$ from both sides of (7.8) (thus replacing f/g by a new fraction f_1/g_1) to obtain a new problem in which all n_i's have been reduced by one. We now repeat the procedure of the previous paragraph to find a_{i,n_i-1} for all i. Continuing similarly, we eventually recover all the a_{ij}. □

7.86. Example. Let us find the partial fraction expansion of

$$\frac{f}{g} = \frac{x^2 - 2}{1 - 2x - x^2 + 2x^3}.$$

To find the required factorization of the denominator, we first reverse the coefficient sequence to obtain $x^3 - 2x^2 - x + 2$. This polynomial factors as $(x-2)(x-1)(x+1)$, so the original denominator can be rewritten as

$$1 - 2x - x^2 + 2x^3 = (1 - 2x)(1 - x)(1 + x)$$

(see the proof of 7.82). We know that

$$\frac{x^2 - 2}{1 - 2x - x^2 + 2x^3} = \frac{A}{1 - 2x} + \frac{B}{1 - x} + \frac{C}{1 + x} \tag{7.10}$$

for suitable complex constants A, B, C. To find A, multiply both sides by $1 - 2x$ to get

$$\frac{x^2 - 2}{(1 - x)(1 + x)} = A + \frac{B(1 - 2x)}{1 - x} + \frac{C(1 - 2x)}{1 + x}.$$

Now set $x = 1/2$ to see that $A = (-7/4)/(3/4) = -7/3$. Similarly,

$$\begin{aligned} B &= \left.\frac{x^2 - 2}{(1 - 2x)(1 + x)}\right|_{x=1} = 1/2; \\ C &= \left.\frac{x^2 - 2}{(1 - 2x)(1 - x)}\right|_{x=-1} = -1/6. \end{aligned}$$

It now follows from (7.10) that

$$\left(\frac{x^2 - 2}{1 - 2x - x^2 + 2x^3}\right)_n = -\frac{7}{3} \cdot 2^n + \frac{1}{2} - \frac{1}{6} \cdot (-1)^n \qquad (n \in \mathbb{N}).$$

7.87. Example. We will find the partial fraction expansion of

$$\frac{f}{g} = \frac{1}{1 - 9x + 30x^2 - 46x^3 + 33x^4 - 9x^5}.$$

Factoring the numerator as in the last example, we find that $g(x) = (1-x)^3(1-3x)^2$. We can therefore write

$$\frac{f}{g} = \frac{A}{(1-x)^3} + \frac{B}{(1-x)^2} + \frac{C}{1-x} + \frac{D}{(1-3x)^2} + \frac{E}{1-3x}. \tag{7.11}$$

To find A, multiply both sides by $(1-x)^3$ and then substitute $x = 1$ to get $A = 1/(-2)^2 = 1/4$. Similarly, multiplication by $(1-3x)^2$ reveals that $D = 1/(2/3)^3 = 27/8$. Having found A and D, we subtract $A/(1-x)^3$ and $D/(1-3x)^2$ from both sides of (7.11). After simplifying, we are left with

$$\frac{(3/8)(3x-7)}{(1-x)^2(1-3x)} = \frac{B}{(1-x)^2} + \frac{C}{1-x} + \frac{E}{1-3x}.$$

Now we repeat the process. Multiplying by $(1-x)^2$ and setting $x = 1$ shows that $B = 3/4$. Similarly, $E = -81/16$. Subtracting these terms from both sides leaves $(27/16)/(1-x)$, so $C = 27/16$. Using (7.11) and (7.76), we conclude that

$$(f/g)_n = \frac{1}{4}\binom{n+2}{2} + \frac{3}{4}\binom{n+1}{1} + \frac{27}{16} + \frac{27}{8}\binom{n+1}{1}3^n - \frac{81}{16}3^n \qquad (n \in \mathbb{N}).$$

7.14 Application to Recursions

In Chapter 2, we saw that many enumeration problems in combinatorics lead naturally to recursion relations. Formal power series and partial fraction expansions provide a powerful method for solving a wide class of recursions. Before stating the general method, we consider some typical examples.

7.88. Example. In 2.22, we found that the number a_n of subsets of an n-element set satisfies the following recursion and initial condition:

$$a_n = 2a_{n-1} \qquad (n \geq 1); \qquad a_0 = 1.$$

It is not hard to guess that the solution to this recursion is $a_n = 2^n$ for all $n \geq 0$, and it is then routine to prove that this guess is correct by induction on n. However, for more complicated recursions, one is unlikely to find the solution by guessing. Thus, let us see how to solve the recursion using formal power series.

We introduce the formal series $F = \sum_{n\geq 0} a_n x^n$ whose coefficients are given by the unknown sequence (a_n). Notice that $xF = \sum_{m\geq 0} a_m x^{m+1} = \sum_{n\geq 1} a_{n-1}x^n$. By the recursive description of the a_n's, we see that $(F - 2xF)_n = 0$ for all $n \geq 1$. On the other hand, $(F-2xF)_0 = F_0 = a_0 = 1$ by the initial condition. It follows that

$$F - 2xF = 1 \in \mathbb{C}[[x]].$$

Solving for F, we find that

$$F = \frac{1}{1-2x} = \sum_{n=0}^{\infty} 2^n x^n.$$

Comparing coefficients of x^n leads to the expected solution $a_n = 2^n$.

Now let us modify the problem by changing the initial condition to $a_0 = 3$. The same reasoning as above leads to $F = 3/(1-2x)$, so that the new solution is $a_n = 3 \cdot 2^n$.

For a more subtle modification, let us change the recursion to $a_n = 2a_{n-1} + 1$, with

initial condition $a_0 = 0$. (This recursion describes the number of moves needed to solve the famous "Tower of Hanoi" puzzle.) Define $F = \sum_{n\geq 0} a_n x^n$ as before. We now have $(F - 2xF)_n = a_n - 2a_{n-1} = 1$ for all $n \geq 1$, and $(F - 2xF)_0 = a_0 = 0$. We conclude that

$$F - 2xF = x + x^2 + x^3 + \cdots + x^n + \cdots = \frac{x}{1-x}.$$

Solving for F and using partial fractions, we get

$$F = \frac{x}{(1-x)(1-2x)} = \frac{1}{1-2x} - \frac{1}{1-x}.$$

Extracting the coefficient of x^n yields the solution $a_n = 2^n - 1^n = 2^n - 1$ for all $n \geq 0$.

7.89. Example: Fibonacci Recursion. The *Fibonacci numbers* are defined by the recursion $f_n = f_{n-1} + f_{n-2}$ for all $n \geq 2$, with initial conditions $f_0 = 0$, $f_1 = 1$. (Sometimes other initial conditions are used, such as $f_0 = f_1 = 1$, which leads to a shift in the indexing of the sequence.) Let us use formal power series to find an explicit closed formula for the numbers f_n. Define $F = \sum_{n\geq 0} f_n x^n$. Since $(xF)_n = f_{n-1}$ and $(x^2F)_n = f_{n-2}$ for all $n \geq 2$, the recursion gives

$$(F - xF - x^2F)_n = 0 \qquad (n \geq 2).$$

On the other hand, the initial conditions show that

$$(F - xF - x^2F)_0 = f_0 = 0; \qquad (F - xF - x^2F)_1 = f_1 - f_0 = 1.$$

It follows that $F - xF - x^2F = 0 + 1x + 0x^2 + \cdots = x$. Solving for F gives

$$F = \frac{x}{1 - x - x^2}.$$

We now apply the method of partial fractions. First, reversing the coefficient sequence in the denominator gives the polynomial $x^2 - x - 1 = (x - r_1)(x - r_2)$, where (by the quadratic formula)

$$r_1 = \frac{1+\sqrt{5}}{2}, \qquad r_2 = \frac{1-\sqrt{5}}{2}.$$

It follows that $1 - x - x^2 = (1 - r_1 x)(1 - r_2 x)$. Next write

$$F = \frac{x}{(1-r_1x)(1-r_2x)} = \frac{A}{1-r_1x} + \frac{B}{1-r_2x}.$$

Multiplying both sides by $(1 - r_1 x)$ and setting $x = 1/r_1$, we find that $A = (1/r_1)/(1 - r_2/r_1) = 1/(r_1 - r_2) = 1/\sqrt{5}$. Similarly we find that $B = -1/\sqrt{5}$, so that

$$F = \frac{x}{(1-r_1x)(1-r_2x)} = \frac{1}{\sqrt{5}}\left(\frac{1}{1-r_1x} - \frac{1}{1-r_2x}\right).$$

Extracting the coefficient of x^n, we conclude that the Fibonacci numbers are given by the following exact formula:

$$f_n = (r_1^n - r_2^n)/\sqrt{5} = \frac{1}{2^n\sqrt{5}}\left[(1+\sqrt{5})^n - (1-\sqrt{5})^n\right] \qquad (n \geq 0).$$

Note that $|r_2| \approx 0.618 < 1$, so that $\lim_{n\to\infty} r_2^n = 0$. It follows that, for very large n,

$$f_n \approx r_1^n/\sqrt{5} \approx (0.447214) \cdot (1.61803)^n.$$

This formula tells us the asymptotic growth rate of the Fibonacci numbers.

The next theorem gives a general method for solving recursion relations with constant coefficients.

7.90. Theorem: Recursions with Constant Coefficients. Suppose we are given the following data: a positive integer k, constants $c_1, c_2, \ldots, c_k, d_0, \ldots, d_{k-1} \in K$, and a function $g : \mathbb{N} \to K$. The recursion

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_k a_{n-k} + g(n) \qquad (n \geq k)$$

with initial conditions $a_i = d_i$ for $0 \leq i < k$ has a unique solution. Setting $d'_i = d_i - c_1 d_{i-1} - c_2 d_{i-2} - \cdots - c_i d_0$, $F = \sum_{n\geq 0} a_n x^n$, $G = \sum_{i=0}^{k-1} d'_i x^i + \sum_{n\geq k} g(n) x^n$, and $p = 1 - c_1 x - c_2 x^2 - \cdots - c_k x^k$, we have $F = G/p$.

Proof. The existence and uniqueness of the sequence $(a_n : n \geq 0)$ satisfying the given recursion and initial conditions is intuitively plausible and can be informally established by an induction argument. (A formal proof requires the *recursion theorem* from set theory; see Section 12 of Halmos [66] for a discussion of this theorem.) It follows that the formal series $F \in K[[x]]$ in the theorem statement is well defined. Consider next the formal series

$$H = (1 - c_1 x - c_2 x^2 - \cdots - c_k x^k)F = pF.$$

For each $n \geq k$, the recursion shows that

$$H_n = a_n - c_1 a_{n-1} - \cdots - c_k a_{n-k} = g(n) = G_n.$$

On the other hand, for $0 \leq n < k$, the initial conditions show that

$$H_n = d'_n = G_n.$$

So $H = G$, and the formula for F follows by dividing the equation $G = pF$ by the invertible element p. □

7.91. Example. Let us solve the recursion

$$a_n = 5a_{n-1} - 6a_{n-2} + 2^n \qquad (n \geq 2)$$

subject to the initial conditions $a_0 = 0$, $a_1 = 1$. Use the theorem with $k = 2$, $c_1 = 5$, $c_2 = -6$, $d_0 = 0$, $d_1 = 1$, and $g(n) = 2^n$. We find that $d'_0 = 0$, $d'_1 = 1$, $G = 0 + 1x + \sum_{n\geq 2} 2^n x^n = (1-2x)^{-1} - 1 - x$, $p = 1 - 5x + 6x^2 = (1-2x)(1-3x)$, and finally

$$F = \frac{(1-2x)^{-1} - 1 - x}{(1-2x)(1-3x)} = \frac{x + 2x^2}{(1-2x)^2(1-3x)}.$$

A tedious but routine partial fraction computation gives us

$$F = -2(1-2x)^{-2} - 3(1-2x)^{-1} + 5(1-3x)^{-1}.$$

Finally, using 7.76, we obtain

$$a_n = F_n = -2\binom{n+1}{1}2^n - 3 \cdot 2^n + 5 \cdot 3^n = 5 \cdot 3^n - (2n+5) \cdot 2^n.$$

Once this formula has been found, one may check the answer by proving by induction that it satisfies the initial condition and recursion.

7.15 Formal Exponentiation and Formal Logarithms

In 7.67, we introduced formal series

$$E = \sum_{n\geq 1} x^n/n! = e^x - 1, \qquad L = \sum_{n\geq 1} (-1)^{n-1} x^n/n = \log(1+x),$$

and showed that $L \bullet E = x = E \bullet L$. We can use these series and formal composition to define the exponential and logarithm of certain power series.

7.92. Definition: $\exp(G)$ and $\log(1+G)$. Suppose G is a formal power series with constant term zero. We define

$$e^G = \exp(G) = (E \bullet G) + 1 = \sum_{n=0}^{\infty} G^n/n!; \qquad \log(1+G) = L \bullet G = \sum_{n=1}^{\infty} (-1)^{n-1} G^n/n.$$

Similarly, if H is a formal power series with $H(0) = 1$, define

$$\log H = \log(1 + [H-1]) = L \bullet (H-1).$$

The combinatorial significance of exponentiating a formal power series will be revealed in the next chapter (see 8.32). For the moment, we will be content to prove some properties of the ordinary exponential and logarithm functions that are also satisfied by their formal counterparts.

7.93. Theorem: Sum-to-Product Rule for Exponentials. For all $G, H \in K[[x]]$ with $G(0) = 0 = H(0)$, we have

$$\exp(G+H) = \exp(G)\exp(H).$$

More generally, given a sequence $G_k \in K[[x]]$ with $G_k(0) = 0$ for all k and $\lim_{k\to\infty} G_k = 0$,

$$\exp\left(\sum_{k=1}^{\infty} G_k\right) = \prod_{k=1}^{\infty} \exp(G_k).$$

Proof. To prove the first identity, look at the coefficient of x^n on each side. For the left side, the binomial theorem gives

$$\begin{aligned}
\exp(G+H)_n &= \left(\sum_{k=0}^{\infty} \frac{(G+H)^k}{k!}\right)_n = \left(\sum_{k=0}^{n} \frac{(G+H)^k}{k!}\right)_n \\
&= \left(\sum_{k=0}^{n}\sum_{j=0}^{k} \binom{k}{j} \frac{G^j H^{k-j}}{k!}\right)_n = \sum_{(i,j)\in\mathbb{N}^2:\ i+j\leq n} \left(\frac{G^j H^i}{j!i!}\right)_n \\
&= \sum_{(i,j)\in\mathbb{N}^2:\ i+j\leq n} \sum_{m=0}^{n} \frac{G^j(m)}{j!} \cdot \frac{H^i(n-m)}{i!}.
\end{aligned}$$

On the right side, we get

$$\begin{aligned}[\exp(G)\exp(H)]_n &= \sum_{m=0}^{n} \exp(G)_m \exp(H)_{n-m} \\ &= \sum_{m=0}^{n} \left(\sum_{j\geq 0} \frac{G^j(m)}{j!}\right)\left(\sum_{i\geq 0} \frac{H^i(n-m)}{i!}\right) \\ &= \sum_{i,j,m=0}^{n} \frac{G^j(m)}{j!}\cdot\frac{H^i(n-m)}{i!}.\end{aligned}$$

The two answers almost agree, but the ranges of summation for i and j do not quite match. However, consider a triple (i,j,m) in the last summation for which $i+j>n$. This forces either $j>m$ or $i>n-m$, so either $G^j(m)=0$ or $H^i(n-m)=0$. In any case, the summand indexed by this triple (i,j,m) is zero. Dropping these summands, we get precisely the sum occurring in the earlier calculation.

Iteration of the result just proved shows that $\exp\left(\sum_{k=1}^{N} G_k\right) = \prod_{k=1}^{N}\exp(G_k)$ for any (finite) $N\in\mathbb{N}$. To prove the same formula with $N=\infty$, we check the coefficient of x^M on each side. One sees immediately from the definition that $\operatorname{ord}(F)>M$ implies $\operatorname{ord}(\exp(F)-1)>M$. Choose k_0 large enough that $\operatorname{ord}(G_k)>M$ or $G_k=0$ for all $k>k_0$. Taking $F=\sum_{k>k_0} G_k$, we then have $\operatorname{ord}(F)>M$ or $F=0$. Write $\exp(F)=1+H$ where $\operatorname{ord}(H)>M$ or $H=0$. Using the result for finite sums gives

$$\begin{aligned}\exp\left(\sum_{k=1}^{\infty} G_k\right)_M &= \exp\left(\sum_{k=1}^{k_0} G_k + F\right)_M \\ &= \left[\exp(F)\prod_{k=1}^{k_0}\exp(G_k)\right]_M = \left[\prod_{k=1}^{k_0}\exp(G_k)\right]_M.\end{aligned}$$

Now, for any $k_1\geq k_0$,

$$\left[\prod_{k=1}^{k_0}\exp(G_k)\right]_M = \left[\prod_{k=1}^{k_1}\exp(G_k)\right]_M$$

since, for $k_0<k\leq k_1$, $\exp(G_k)$ is 1 plus terms of order larger than M. We conclude finally that

$$\exp\left(\sum_{k=1}^{\infty} G_k\right)_M = \left[\prod_{k=1}^{\infty}\exp(G_k)\right]_M$$

for every $M\geq 0$. □

7.94. Theorem: Exponential and Logarithm are Inverses. If $H\in K[[x]]$ satisfies $H(0)=1$, then $\exp(\log(H))=H$. If $G\in K[[x]]$ satisfies $G(0)=0$, then $\log(\exp(G))=G$.

Proof. Recall from 7.67 that $E\bullet L = x = L\bullet E$. We can therefore compute

$$\begin{aligned}\exp(\log(H)) &= (E\bullet(L\bullet(H-1)))+1 = ((E\bullet L)\bullet(H-1))+1 \\ &= (x\bullet(H-1))+1 = H-1+1 = H;\end{aligned}$$

$$\begin{aligned}\log(\exp(G)) &= L\bullet(([E\bullet G]+1)-1) = L\bullet(E\bullet G) \\ &= (L\bullet E)\bullet G = x\bullet G = G.\end{aligned}$$ □

7.95. Theorem: Logarithm of a Product. For all $G, H \in K[[x]]$ with $G(0) = 1 = H(0)$,

$$\log(GH) = \log(G) + \log(H).$$

More generally, given a sequence $G_k \in K[[x]]$ with $G_k(0) = 1$ for all k and $\lim_{k\to\infty} G_k = 1$,

$$\log\left(\prod_{k=1}^{\infty} G_k\right) = \sum_{k=1}^{\infty} \log(G_k).$$

Proof. Since G and H have constant term 1, we know that

$$GH = \exp(\log G)\exp(\log H) = \exp[(\log G) + (\log H)].$$

Since GH has constant term 1, we can take logarithms to conclude that

$$\log(GH) = \log(G) + \log(H),$$

as desired. The formula for converting infinite products to infinite sums is proved similarly. □

Formal exponentials and logarithms obey formal differentiation rules entirely analogous to those learned in calculus.

7.96. Theorem: Derivative Rules for Exponentials and Logarithms. If $G \in K[[x]]$ satisfies $G(0) = 0$, then $(\exp(G))' = G' \exp G$. If $H \in K[[x]]$ satisfies $H(0) = 1$, then $(\log(H))' = H'/H$.

Proof. A direct calculation using the definition shows that $E' = E + 1$. Applying the formal chain rule, we conclude that

$$\begin{aligned}(\exp(G))' &= [(E \bullet G) + 1]' = (E' \bullet G)G' \\ &= ((E+1) \bullet G)G' = ((E \bullet G) + 1)G' = G' \exp(G).\end{aligned}$$

Use this result to differentiate the identity $\exp(\log(H)) = H$. We obtain

$$(\log(H))' \exp(\log(H)) = H',$$

or equivalently $(\log(H))'H = H'$. Since $H(0) = 1$, we can divide by H to conclude that $(\log(H))' = H'/H$. □

7.97. Theorem: Power Rule for Logarithms. If $H \in K[[x]]$ satisfies $H(0) = 1$ and $r \in K$, then $\log(H^r) = r\log(H)$.

Proof. On one hand, both $\log(H^r)$ and $r\log(H)$ have formal derivative equal to rH'/H (by 7.79, 7.75, 7.96, and the chain rule). On the other hand, both $\log(H^r)$ and $r\log(H)$ have zero constant term. Thus these two series must be equal (see 7.132). □

7.16 Multivariable Polynomials and Formal Series

So far we have discussed polynomials and formal power series involving a single indeterminate. One can generalize this setup to arrive at the notions of multivariable polynomials and series.

7.98. Definition: Formal Multivariable Power Series and Polynomials. A *formal power series in k variables with coefficients in K* is a function $F : \mathbb{N}^k \to K$. The set of all such series is denoted $K[[x_1, x_2, \ldots, x_k]]$. A series $f \in K[[x_1, \ldots, x_k]]$ is a *polynomial* iff $\{\vec{n} \in \mathbb{N}^k : f(\vec{n}) \neq 0\}$ is finite. The set of all such polynomials is denoted $K[x_1, \ldots, x_k]$.

The power series notation for a function $F : \mathbb{N}^k \to K$ is

$$F = \sum_{(n_1,\ldots,n_k)\in\mathbb{N}^k} F(n_1, \ldots, n_k) x_1^{n_1} x_2^{n_2} \cdots x_k^{n_k};$$

the function value $F(n_1, \ldots, n_k)$ is called the *coefficient of* $x_1^{n_1} \cdots x_k^{n_k}$ *in* F. F is a polynomial iff only a finite number of its coefficients are nonzero.

7.99. Definition: Algebraic Operations on Multivariable Series. Given $c \in K$ and series $F, G \in K[[x_1, \ldots, x_k]]$, the *sum* $F + G$ is defined by $(F + G)(\vec{n}) = F(\vec{n}) + G(\vec{n})$ for all $\vec{n} \in \mathbb{N}^k$. The *scalar multiple* cF is defined by $(cF)(\vec{n}) = c(F(\vec{n}))$ for all $\vec{n} \in \mathbb{N}^k$. The *product* FG is defined by

$$(FG)(\vec{n}) = \sum_{\substack{(\vec{i},\vec{j})\in(\mathbb{N}^k)^2:\\ \vec{i}+\vec{j}=\vec{n}}} F(\vec{i})G(\vec{j}) \qquad (\vec{n} \in \mathbb{N}^k).$$

7.100. Example. For $1 \leq i \leq k$, let $x_i : \mathbb{N}^k \to K$ be the polynomial defined by sending $(0, 0, \ldots, 1, \ldots, 0)$ (the 1 occurs in position i) to 1 and sending everything else to zero. One can check that $cx_1^{n_1} \cdots x_k^{n_k}$ is the series that sends $(n_1, \ldots, n_k)$ to c and everything else to zero. This justifies the notation used above for elements $F \in K[[x_1, \ldots, x_k]]$, at least when F is a polynomial.

The following theorem is proved by making the necessary adjustments to the proofs given in the one-variable case. Even more general results are sketched in the exercises.

7.101. Theorem: Algebraic Structure of Multivariable Series and Polynomials. $K[[x_1, \ldots, x_k]]$ and $K[x_1, \ldots, x_k]$ are commutative rings that are integral domains, as well as vector spaces over K containing a copy of K. The set $\{x_1^{n_1} \cdots x_k^{n_k} : (n_1, \ldots, n_k) \in \mathbb{N}^k\}$ is a basis for the vector space $K[x_1, \ldots, x_k]$.

Multivariable polynomial rings satisfy the following *universal mapping property* that generalizes 7.22. The proof is also left as an exercise.

7.102. Theorem: Evaluation Homomorphisms for Multivariable Polynomials. Let S be a commutative ring containing K. (a) For each $f \in K[x_1, \ldots, x_k]$, there is an associated function $P_f : S^k \to S$ given by

$$P_f(z_1, \ldots, z_k) = \sum_{(n_1,\ldots,n_k)\in\mathbb{N}^k} f(n_1, \ldots, n_k) z_1^{n_1} \cdots z_k^{n_k} \qquad (z_i \in S).$$

(b) For each k-tuple $\vec{z} = (z_1, \ldots, z_k) \in S^k$, there exists a unique ring homomorphism $E : K[x_1, \ldots, x_k] \to S$ such that $E(c) = c$ for all $c \in K$ and $E(x_i) = z_i$ for all $i \leq k$; namely, $E(f) = P_f(\vec{z})$ for $f \in K[x_1, \ldots, x_k]$. We write $E = \mathrm{ev}_{\vec{z}}$ and call it the *evaluation homomorphism determined by setting each* x_i *equal to* z_i.

7.103. Definition: Formal Partial Derivatives. For $1 \leq i \leq k$, define a map $D_i : K[[x_1, \ldots, x_k]] \to K[[x_1, \ldots, x_k]]$ by

$$D_i(F) = \frac{\partial F}{\partial x_i} = \sum_{(n_1,n_2,\ldots,n_k)\in\mathbb{N}^k} (n_i + 1) F(n_1, \ldots, n_i + 1, \ldots, n_k) x_1^{n_1} \cdots x_i^{n_i} \cdots x_k^{n_k}.$$

D_i is called the *formal partial derivative operator with respect to* x_i.

It is routine to check that the analogues of the one-variable differentiation rules (§7.8) extend to the partial derivative operators D_i. There are also formal versions of the multivariable chain rule. We now prove one such rule for multivariable polynomials, which will be used in §10.16.

7.104. Theorem: Multivariable Chain Rule. Let $h \in K[y_1, \ldots, y_n]$ and $g_1, \ldots, g_n \in K[x_1, \ldots, x_m]$. Let $h \bullet g \in K[x_1, \ldots, x_m]$ denote the polynomial obtained by setting each $y_i = g_i$ in h. For $1 \leq k \leq m$,

$$D_k(h \bullet g) = \sum_{j=1}^{n} ((D_j h) \bullet g) D_k(g_j).$$

Informally, we may write

$$\frac{\partial (h \bullet g)}{\partial x_k} = \frac{\partial h}{\partial y_1} \cdot \frac{\partial g_1}{\partial x_k} + \cdots + \frac{\partial h}{\partial y_n} \cdot \frac{\partial g_n}{\partial x_k},$$

with all of the partial derivatives $\partial h / \partial y_j$ being evaluated at $(y_1, \ldots, y_n) = (g_1, \ldots, g_n)$.

Proof. Both sides of the claimed identity are K-linear functions of h. So it suffices to check the identity when h has the form $y_1^{e_1} \cdots y_n^{e_n}$. In this case, $h \bullet g = g_1^{e_1} \cdots g_n^{e_n}$. Viewing this as a product of $e_1 + \cdots + e_n$ factors, each equal to some g_j, the multivariable product rule leads to

$$D_k(h \bullet g) = \sum_{j=1}^{n} e_j g_1^{e_1} \cdots g_j^{e_j - 1} (D_k(g_j)) \cdots g_n^{e_n}.$$

On the other side, $(D_j h) \bullet g = e_j g_1^{e_1} \cdots g_j^{e_j - 1} \cdots g_n^{e_n}$. Multiplying by $D_k(g_j)$ and summing over j gives the same answer as before, so the proof is complete. □

Summary

Table 7.1 reviews the definitions of concepts and operations involving formal power series and polynomials. Table 7.2 lists some rules and formulas arising in computations with formal power series (some hypotheses on constant terms are omitted in this table). Let us also recall the following results.

- *Algebraic Structure of $K[[x]]$ and $K[x]$.* Both $K[[x]]$ and $K[x]$ are commutative rings, integral domains, and vector spaces over K containing a copy of K. The same holds for $K[[x_1, \ldots, x_k]]$ and $K[x_1, \ldots, x_k]$. The set $\{x^i : i \geq 0\}$ is a basis for $K[x]$, whereas the set $\{x_1^{n_1} \cdots x_k^{n_k} : (n_1, \ldots, n_k) \in \mathbb{N}^k\}$ is a basis for $K[x_1, \ldots, x_k]$.
- *Degree and Order.* For polynomials $f, g \in K[x]$, $\deg(f+g) \leq \max(\deg(f), \deg(g))$ and $\deg(fg) = \deg(f) + \deg(g)$ whenever both sides are defined. For series $F, G \in K[[x]]$, we have $\operatorname{ord}(F+G) \geq \min(\operatorname{ord}(F), \operatorname{ord}(G))$ and $\operatorname{ord}(FG) = \operatorname{ord}(F) + \operatorname{ord}(G)$ whenever both sides are defined.
- *Multiplicative Inverses in $K[x]$ and $K[[x]]$.* A polynomial f is invertible (a unit) in $K[x]$ iff $\deg(f) = 0$. A series F is invertible in $K[[x]]$ iff $F(0) \neq 0$. In this case, one can use the formal geometric series to invert F or the recursive formula $(F^{-1})_n = -(1/F_0) \sum_{k=1}^{n} F_k (F^{-1})_{n-k}$. A nonzero formal quotient $F/G \in K((x))$ can be written in a unique way as a Laurent series $\sum_{n=m}^{\infty} a_n x^n$ where $m \in \mathbb{Z}$, each $a_n \in K$, and $a_m \neq 0$; $m = \operatorname{ord}(F) - \operatorname{ord}(G)$ is the *order* of this Laurent series.

TABLE 7.1
Definitions concerning formal power series and polynomials.

Term	Brief Definition
formal power series	function $F : \mathbb{N} \to K$, denoted $\sum_{n=0}^{\infty} F_n x^n$
formal polynomial	formal series f with $\{n : f_n \neq 0\}$ finite
ring	set with addition and multiplication satisfying axioms in 2.2
integral domain	nonzero commutative ring with no zero divisors
field	nonzero commutative ring with all nonzero elements invertible
$K[[x]]$	set of formal power series with coefficients in K
$K[x]$	set of formal polynomials with coefficients in K
$K(x)$	$\{f/g : f, g \in K[x], g \neq 0\}$ =field of fractions of $K[x]$
$K((x))$	$\{F/G : F, G \in K[[x]], G \neq 0\}$ =field of fractions of $K[[x]]$, or the set of formal Laurent series $\sum_{n=m}^{\infty} a_n x^n$ $(m \in \mathbb{Z}, a_n \in K)$
the series x	$x = X_1 = (0, 1, 0, 0, \ldots) = \sum_{n=0}^{\infty} \chi(n = 1) x^n$
equality of series	$F = G$ iff $F_n = G_n$ for all $n \in \mathbb{N}$
sum of series	$(F + G)(n) = F(n) + G(n)$ for all $n \in \mathbb{N}$
product of series	$(FG)(n) = \sum_{(i,j) \in \mathbb{N}^2:\ i+j=n} F(i)G(j)$ for $n \in \mathbb{N}$ (convolution)
derivative of series	$F'(n) = (dF/dx)_n = (n + 1)F(n + 1)$ for $n \in \mathbb{N}$
kth derivative	$F^{(k)}(n) = (d^k F/dx^k)_n = (n + 1) \cdots (n + k)F(n + k)$ $(n \in \mathbb{N})$
formal limit	$F_n \to L$ iff $\forall m \in \mathbb{N}, \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, (n \geq N \Rightarrow F_n(m) = L(m))$
infinite sum of series	$\sum_{n=0}^{\infty} F_n = \lim_{N \to \infty} \sum_{n=0}^{N} F_n$ (if limit exists)
infinite product	$\prod_{n=0}^{\infty} F_n = \lim_{N \to \infty} \prod_{n=0}^{N} F_n$ (if limit exists)
formal composition	$F \bullet G = \sum_{n=0}^{\infty} F_n G^n$ (need $F \in K[x]$ or $G(0) = 0$)
formal e^x	$e^x = \sum_{n=0}^{\infty} x^n/n! \in K[[x]]$
formal $\sin x$	$\sin x = (0, 1, 0, -1/3!, 0, 1/5!, 0, \ldots) \in K[[x]]$
formal $\cos x$	$\cos x = (1, 0, -1/2!, 0, 1/4!, 0, -1/6!, \ldots) \in K[[x]]$
formal $\log(1 + x)$	$\log(1 + x) = \sum_{n=1}^{\infty} (-1)^{n-1} x^n/n \in K[[x]]$
formal $(1 + x)^r$	$\mathrm{Pow}_r = (1 + x)^r = \sum_{n=0}^{\infty} ((r)\downarrow_n /n!) x^n \in K[[x]]$ $(r \in K)$
formal power	$F^r = \sum_{n=0}^{\infty} ((r)\downarrow_n /n!)(F - 1)^n$ (need $F(0) = 1$)
formal exponential	$\exp(G) = \sum_{n=0}^{\infty} G^n/n!$ (need $G(0) = 0$)
formal logarithm	$\log(1 + G) = \sum_{n=1}^{\infty} (-1)^{n-1} G^n/n$ (need $G(0) = 0$) or $\log(F) = \sum_{n=1}^{\infty} (-1)^{n-1} (F - 1)^n/n$ (need $F(0) = 1$)
degree	$\deg(f) = \max\{n : f_n \neq 0\}$ for nonzero $f \in K[x]$
order	$\mathrm{ord}(F) = \min\{n : F_n \neq 0\}$ for nonzero $F \in K[[x]]$
polynomial function	for $f \in K[x]$, $P_f : S \to S$ sends $z \in S \supseteq K$ to $\sum_{n=0}^{\deg(f)} f_n z^n$
ring homomorphism	map between rings preserving $+$, $\times$, and 1
evaluation hom.	$\mathrm{ev}_z(f) = P_f(z)$ for $f \in K[x]$, $z \in S \supseteq K$
$K[[x_1, \ldots, x_k]]$	set of k-variable formal series $F : \mathbb{N}^k \to K$
$K[x_1, \ldots, x_k]$	set of k-variable polynomials
partial derivative	$D_i F(n_1, \ldots, n_k) = (n_i + 1)F(n_1, \ldots, n_i + 1, \ldots, n_k)$

TABLE 7.2
Rules for calculating with formal power series.

Product of k series:	$(F_1 \cdots F_k)(n) = \sum_{i_1+\cdots+i_k=n} F_1(i_1)\cdots F_k(i_k)$.
Positive powers:	$G^m(n) = \sum \binom{m}{k_0,\ldots,k_n} G(0)^{k_0}\cdots G(n)^{k_n}$, summed over $(k_0,\ldots,k_n)$ with $\sum_i k_i = m$ and $\sum_i i k_i = n$.
Geometric series:	$(1-G)^{-1} = \sum_{n=0}^{\infty} G^n$ when $G(0) = 0$.
Negative powers:	$(1-G)^{-m} = \sum_{n=0}^{\infty} \binom{n+m-1}{n,m-1} G^n$ when $m \in \mathbb{N}^+$ and $G(0) = 0$.
Limit rules:	$F_n \to P$ and $G_n \to Q$ imply $F_n + G_n \to P + Q$, $\quad F_n G_n \to PQ$, $\quad F_n \bullet G_n \to P \bullet Q$.
Derivative rules:	$(F+G)' = F' + G'$, $\quad (FG)' = (F')G + F(G')$, $\quad (x^k)' = kx^{k-1}$, $(F \bullet G)' = (F' \bullet G)G'$, $\quad (F^r)' = rF^{r-1}F'$, $(\exp(G))' = G'\exp(G)$ $[G(0) = 0]$, $\quad (\log(H))' = H'/H$ $[H(0) = 1]$, $(H_n \to F) \Rightarrow (H_n' \to F')$, $\quad (\sum_{n=1}^{\infty} H_n)' = \sum_{n=1}^{\infty} H_n'$, $F = \sum_{k=0}^{\infty} (F^{(k)}(0)/k!)x^k$; $\quad D_k(h \bullet g) = \sum_j ((D_j h) \bullet g) D_k(g_j)$.
Laws of exponents:	$\mathrm{Pow}_{r+s} = \mathrm{Pow}_r \mathrm{Pow}_s$; $\quad F^{r+s} = F^r F^s$; $(F^r)^n = F^{rn}$ for $r, s \in K, n \in \mathbb{Z}$.
Exp and Log:	$\exp(G+H) = \exp(G)\exp(H)$; $\quad \exp\left(\sum_{k=1}^{\infty} G_k\right) = \prod_{k=1}^{\infty} \exp(G_k)$; $\log(GH) = \log(G) + \log(H)$; $\quad \log\left(\prod_{k=1}^{\infty} G_k\right) = \sum_{k=1}^{\infty} \log(G_k)$; $\log(\exp(G)) = G$ $[G(0) = 0]$; $\quad \exp(\log(H)) = H$ $[H(0) = 1]$; $\log(H^r) = r\log(H)$ $[H(0) = 1]$.

- *Compositional Inverses in $K[[x]]$.* Formal composition is associative and has x as a two-sided identity. For fixed G, the map $F \mapsto F \bullet G$ is a ring homomorphism fixing K. A series $F \in K[[x]]$ has an inverse G relative to formal composition if $F(0) = 0$ and $F(1) \neq 0$. The set of all such series is closed under composition and forms a group under this operation. The inverse G of F may be found by Lagrange inversion (see 8.15) or by the recursive formula $G_n = -(1/F_1^n)(\sum_{m=0}^{n-1} G_m F^m)_n$.

- *Evaluation Homomorphisms.* Let S be a commutative ring containing K and $z_1, \ldots, z_k \in S$. There exists a unique ring homomorphism $E : K[x_1, \ldots, x_k] \to S$ such that $E(x_i) = z_i$ for $1 \leq i \leq k$ and $E(c) = c$ for all $c \in K$.

- *Density of Polynomials in $K[[x]]$.* For each $F \in K[[x]]$, $F = \lim_{N\to\infty} \sum_{n=0}^{N} F_n x^n$, so that any formal power series is a limit of formal polynomials.

- *Existence Criteria for Formal Limits.* For nonzero series $F_k \in K[[x]]$:
  $F_k \to 0$ in $K[[x]]$ iff $\mathrm{ord}(F_k) \to \infty$ in $\mathbb{R}$;
  $\sum_{k=0}^{\infty} F_k$ exists in $K[[x]]$ iff $\mathrm{ord}(F_k) \to \infty$ in $\mathbb{R}$;
  if $F_k(0) = 0$ for all k, $\prod_{k=0}^{\infty}(1+F_k)$ exists in $K[[x]]$ iff $\mathrm{ord}(F_k) \to \infty$ in $\mathbb{R}$.

- *Polynomial Factorization in $\mathbb{C}[x]$ and Partial Fractions.* A monic polynomial $p \in \mathbb{C}[x]$ factors uniquely as $(x-r_1)^{n_1}\cdots(x-r_k)^{n_k}$ with $r_i \in \mathbb{C}$. If instead $p(0) = 1$, we can write

$p = (1 - s_1 x)^{n_1} \cdots (1 - s_k x)^{n_k}$ with $s_i \in \mathbb{C}$. Given $q \in \mathbb{C}[x]$, there is a unique expression

$$\frac{q}{p} = h + \sum_{i=1}^{k} \sum_{j=1}^{n_i} \frac{a_{ij}}{(1 - s_i x)^j},$$

where h is the remainder when q is divided by p; each a_{i,n_i} is found by multiplying all terms by $(1 - s_i x)^{n_i}$ and setting $x = 1/s_i$; and the remaining a_{ij}'s are found by subtracting the previously recovered terms and iterating.

- *Recursions with Constant Coefficients.* Suppose $F_n = \sum_{i=1}^{k} c_i F_{n-i} + H_n$ for all $n \geq k$, where $c_1, \ldots, c_k \in K$ and $H \in K[[x]]$ are given. F is uniquely determined by the k initial values $F_0, \ldots, F_{k-1}$. The series F has the form G/p, where $p = 1 - c_1 x - \cdots - c_k x^k$ and G is a series with $G_n = H_n$ for $n \geq k$.

Exercises

7.105. Let $f = x - x^2 + 3x^4$ and $g = 1 - 2x + 3x^4$. Compute $f + g$, fg, and the degrees and orders of f, g, $f + g$, and fg.

7.106. Let $F = (1, 0, 1, 0, 1, 0, \ldots)$ and $G = \sum_{n \geq 0} n x^n$. Compute $F + G$, FG, $F(1 + x)$, $F(1 - x^2)$, $G(1 + x)$, F', G', and the orders of all these series.

7.107. Let $f = x^2 + 4x - 1$ and $g = x^3 + x$. Compute $P_f(2)$, $P_g(\sqrt{5})$, $P_f(x)$, $P_f(g)$, $P_g(f)$, and $P_f(f)$.

7.108. Compute the coefficient of x^n for $0 \leq n \leq 6$ for each of the following formal series: (a) $e^x + \sin x$; (b) $e^x \sin x$; (c) $(\cos x) \log(1 + x)$; (d) $(\log(1 + x))^2$.

7.109. (a) Find necessary and sufficient conditions for strict inequality to hold in the formula $\deg(f + g) \leq \max(\deg(f), \deg(g))$. (b) Find necessary and sufficient conditions for strict inequality to hold in the formula $\operatorname{ord}(F + G) \geq \min(\operatorname{ord}(F), \operatorname{ord}(G))$.

7.110. Use (7.3) in the proof of 7.40 to find the first five terms in the multiplicative inverse of each of the following series: (a) e^x; (b) $1 - 2x + x^3 + 3x^4$; (c) $1 + \log(1 + x)$.

7.111. Use 7.41 to find the first five terms in $(1 - x + x^3)^{-1}$.

7.112. Compute the multiplicative inverse of $\sum_{n=0}^{\infty} n^2 x^n$ in $K((x))$.

7.113. Convert the following expressions to formal Laurent series: (a) $(x^2 + 3)/(x^3 - x^2)$; (b) $x/(x^3 - 5x^2 + 6x)$.

7.114. Formal Hyperbolic Sine and Cosine Functions. Define formal series $\sinh x = (e^x - e^{-x})/2$ and $\cosh x = (e^x + e^{-x})/2$. (a) Find $(\sinh x)_n$ and $(\cosh x)_n$ for all $n \in \mathbb{N}$. (b) Show $(\sinh x)' = \cosh x$ and $(\cosh x)' = \sinh x$.

7.115. Complete the proof of 7.8(a) by verifying the remaining ring axioms for $K[[x]]$. Indicate which of the ring axioms for K are used in each part of your proof.

7.116. Let R be any ring. Verify that the sum and product operations in 7.6 (with K replaced by R) make $R[x]$ and $R[[x]]$ rings, which are commutative if R is commutative.

7.117. This exercise shows that the characterization of units in $K[x]$ given in 7.38 can fail if K is not a field. (a) Give an example of a commutative ring R and $f \in R[x]$ such that $\deg(f) = 0$ but f is not a unit of $R[x]$. (b) Give an example of a commutative ring R and $f \in R[x]$ such that $\deg(f) > 0$, yet f is a unit of $R[x]$. (c) Show that for any $n \in \mathbb{N}^+$, there exists f as in part (b) with n nonzero coefficients.

7.118. Continuity of Exp and Log. (a) Assume $F_k(0) = G(0) = 0$ and $F_k \to G$ in $K[[x]]$. Prove $\exp(F_k) \to \exp(G)$. (b) Assume $F_k(0) = G(0) = 1$ and $F_k \to G$. Prove $\log(F_k) \to \log(G)$.

7.119. Prove the following general version of the *universal mapping property for polynomial rings.* Let $f : L \to R$ be a given ring homomorphism between two commutative rings. Given $(z_1, \ldots, z_k) \in R^k$, there exists a unique ring homomorphism $E : L[x_1, \ldots, x_k] \to R$ such that $E(x_i) = z_i$ for all i and $E(c) = f(c)$ for all $c \in L$. Point out any steps in your proof that require the assumption that the rings are commutative.

7.120. (a) Show that, because K is an infinite field, the map $\pi : K[x] \to {}^K K$, given by $\pi(f) = P_f$ for $f \in K[x]$, is injective. (b) Give an example of a commutative ring R such that the map $\pi : R[x] \to {}^R R$ is not injective.

7.121. Prove 7.20(a),(c).

7.122. Let $F_k, G_k, P, Q \in K[[x]]$ satisfy $F_k \to P$ and $G_k \to Q$. Prove $F_k + G_k \to P + Q$.

7.123. Prove 7.54(a),(b),(c),(g).

7.124. Complete the following outline to give a new proof of the formal product rule $(FG)' = (F')G + F(G')$ for $F, G \in K[[x]]$. (a) Show that the result holds when $F = x^i$ and $G = x^j$, for all $i, j \in \mathbb{N}$. (b) Deduce from (a) that the result holds for all $F, G \in K[x]$. (c) Use a continuity argument to obtain the result for all $F, G \in K[[x]]$.

7.125. Prove 7.62(b),(c).

7.126. Use a continuity argument to deduce the formal chain rule for formal power series (see 7.64) from the chain rule for polynomials.

7.127. Prove the following formal derivative identities: (a) $(\sin x)' = \cos x$; (b) $(\cos x)' = -\sin x$; (c) $[\log(1+x)]' = (1+x)^{-1}$.

7.128. Formal Quotient Rule. Suppose $F, G \in K[[x]]$ where $G(0) \neq 0$. Prove the derivative rule $(F/G)' = (GF' - FG')/G^2$.

7.129. Formal Integrals. The *formal integral* or *antiderivative* of a series $F \in K[[x]]$ is the series

$$\int F\,dx = \sum_{n \geq 1} \frac{F_{n-1}}{n} x^n \in K[[x]],$$

which has constant term zero. Compute the formal integrals of the following formal power series: (a) $3 + 2x - 7x^2 + 12x^5$; (b) $\sum_{n \geq 0} n^2 x^n$; (c) $\sum_{n \geq 0} (n+1)! x^n$; (d) e^x; (e) $\sin x$; (f) $\cos x$; (g) $(1+x)^{-1}$; (h) $\frac{3+2x}{1-3x+2x^2}$.

7.130. Prove the following facts about formal integrals (as defined in 7.129).
(a) (sum rule) $\int F + G\,dx = \int F\,dx + \int G\,dx$ for $F, G \in K[[x]]$.
(b) (scalar rule) $\int cF\,dx = c \int F\,dx$ for $c \in K$ and $F \in K[[x]]$.
(c) (linear combination rule) $\int \sum_{i=1}^n c_i H_i\,dx = \sum_{i=1}^n c_i \int H_i\,dx$ for $c_i \in K$ and $H_i \in K[[x]]$.
Can you formulate a similar statement for infinite sums?

(d) (power rule) $\int x^k\,dx = \frac{1}{k+1}x^{k+1}$ for all $k \geq 0$.
(e) (general antiderivatives) For all $F, G \in K[[x]]$, $G' = F$ iff there exists $c \in K$ with $G = \int F\,dx + c$.
(f) (formal fundamental theorems of calculus) $F = \frac{d}{dx}\int F\,dx$ and $\int F'\,dx = F - F(0)$ for $F \in K[[x]]$.
(g) (continuity of integration) If $F_k, H \in K[[x]]$ and $F_k \to H$, then $\int F_k\,dx \to \int H\,dx$.

7.131. Formulate and prove an "integration by parts" rule and a "substitution rule" for formal integrals (as defined in 7.129).

7.132. (a) Given $F, G \in K[[x]]$, prove that $F = G$ iff $F' = G'$ and $F(0) = G(0)$. (b) State and prove an analogous statement for multivariable series.

7.133. (a) Prove that $(\sin x)^2 + (\cos x)^2 = 1$ in $K[[x]]$ by computing the coefficient of x^n on each side. (b) Prove that $(\sin x)^2 + (\cos x)^2 = 1$ in $K[[x]]$ by invoking 7.132 and derivative rules.

7.134. Prove that $(\cosh x)^2 - (\sinh x)^2 = 1$ in $K[[x]]$.

7.135. Cancellation in an Integral Domain. Let R be a nonzero commutative ring. Prove that R is an integral domain iff the following *cancellation axiom* holds: for all $a, b, c \in R$, $ab = ac$ and $a \neq 0$ imply $b = c$.

7.136. Product Rule for Multiple Factors. Let $F_1, \ldots, F_k \in K[[x]]$. Prove that

$$\frac{d}{dx}(F_1 F_2 \cdots F_k) = \sum_{j=1}^{k} F_1 \cdots F_{j-1}(\frac{d}{dx}F_j)F_{j+1}\cdots F_k.$$

Does a version of this rule hold for infinite products?

7.137. Use 7.11 to prove the differentiation rule $\frac{d}{dx}(G^m) = mG^{m-1}G'$ for $G \in K[[x]]$ and $m \in \mathbb{N}^+$ without using the formal chain rule.

7.138. For $m, n \in \mathbb{N}^+$, evaluate the sum

$$\sum_{\substack{(k_0,k_1,\ldots,k_n)\in\mathbb{N}^{n+1}:\\ \sum_i k_i = m, \sum_i i k_i = n}} \frac{m!}{k_0!0!^{k_0}k_1!1!^{k_1}\cdots k_n!n!^{k_n}}.$$

7.139. Let $F = \prod_{k=1}^{\infty}(1-x^k)$. Find F_n for $0 \leq n \leq 22$. Can you see a pattern?

7.140. Carefully justify the following calculation:

$$\prod_{n=1}^{\infty}(1-x^{2n-1})^{-1} = \prod_{i=1}^{\infty}(1-x^{2i})\prod_{j=1}^{\infty}(1-x^j)^{-1} = \prod_{k=1}^{\infty}(1+x^k).$$

In particular, explain why all the infinite products appearing here exist.

7.141. Find a necessary and sufficient condition on series $F_k \in K[[x]]$ so that the infinite product $\prod_{k=1}^{\infty}(1+F_k)^{-1}$ exists.

7.142. Evaluate $\prod_{n=0}^{\infty}(1+x^{2^n})$.

7.143. Verify the partial fraction expansion of F given in 7.91.

7.144. Write out the formal series for each of the following expressions: (a) $(1-x)^{-5}$; (b) $\sqrt{1+x}$; (c) $1/\sqrt{1-x^2}$; (d) $\sqrt[3]{1+3x}$.

7.145. Compute the first four nonzero terms in the following series: (a) $\sqrt{1+x+3x^2}$; (b) $\sqrt[3]{\cos x}$; (c) $(\sum_{n\geq 0}(n+1)^2x^n)^{-5/2}$.

7.146. Compute the first four nonzero terms in: (a) $\exp(\sin x)$; (b) $\log(\cos x)$.

7.147. Find the partial fraction decomposition of $F = (10+2x)/(1-2x-8x^2)$, and use this to determine $F(n)$ for all n.

7.148. Find the partial fraction decomposition of $F = (1-7x)/(15x^2-8x+1)$, and use this to determine $F(n)$ for all n.

7.149. Find the partial fraction decomposition of $F = (2x^3-4x^2-x-3)/(2x^2-4x+2)$, and use this to determine $F(n)$ for all n.

7.150. Find the partial fraction decomposition of $F = (15x^6+30x^5-15x^4-35x^4-15x^2-12x-8)/(15(x^4+2x^3-2x-1))$, and use this to determine $F(n)$ for all n.

7.151. (a) Solve the recursion $a_n = 3a_{n-1}$ $(n \geq 1)$, given that $a_0 = 2$. (b) Solve the recursion $a_n = 3a_{n-1}+3n$ $(n \geq 1)$, given that $a_0 = 2$. (c) Solve the recursion $a_n = 3a_{n-1}+3^n$ $(n \geq 1)$, given that $a_0 = 2$.

7.152. Solve the recursion $a_n = 6a_{n-1} - 8a_{n-2} + g(n)$ (for $n \geq 2$) with initial conditions $a_0 = 0$, $a_1 = 2$ for the following choices of $g(n)$: (a) $g(n) = 0$; (b) $g(n) = 1$; (c) $g(n) = 2^n$; (d) $g(n) = n4^n$.

7.153. The *Lucas numbers* are defined by setting $L_0 = 1$, $L_1 = 3$, and $L_n = L_{n-1} + L_{n-2}$ for $n \geq 2$. Use formal series to find a closed formula for L_n.

7.154. Solve the recursion $a_n = -3a_{n-1} + 2a_{n-2} + 6a_{n-3} - a_{n-4} - 3a_{n-5}$ (for $n \geq 5$) with initial conditions $a_k = k$ for $0 \leq k < 5$.

7.155. Repeat 7.154 with initial conditions $a_k = 3$ for $0 \leq k < 5$.

7.156. Suppose $b_0 = 1$ and $b_n = b_0 + b_1 + \cdots + b_{n-1} + 1$ for all $n \geq 1$. Find $\sum_{n\geq 0} b_n x^n$.

7.157. Suppose $(c_n : n \in \mathbb{Z})$ satisfies $c_0 = 0$, $c_1 = 1$, and $c_n = (c_{n-1} + c_{n+1})/L$ for all $n \in \mathbb{Z}$, where $L \in \mathbb{R}^+$ is a constant. Find an explicit formula for c_n.

7.158. Differentiation of Laurent Series. Define a version of the formal derivative operator for the ring $K((x))$ of formal Laurent series. Extend the derivative rules (in particular, the quotient rule) to this ring.

7.159. Formal Tangent and Secant Functions. Define formal series $\sec x = 1/\cos x$ and $\tan x = \sin x/\cos x$. (a) Compute $(\sec x)_n$ and $(\tan x)_n$ for $0 \leq n \leq 9$. (A combinatorial interpretation of these coefficients is described in §12.8.) (b) Show that $(\tan x)^2+1 = (\sec x)^2$. (c) Show that $(\tan x)' = (\sec x)^2$ and $(\sec x)' = \tan x \sec x$. (d) Show that $(\tan x)_n = 0$ for all even n and $(\sec x)_n = 0$ for all odd n. (e) Can you give similar definitions and results for $\cot x$ and $\csc x$?

7.160. Substitution of rx for x. Given $F \in K[[x]]$ and nonzero $r \in K$, define $F(rx)$ to be the formal composition $F \bullet (rx)$. Prove: (a) $\sin(2x) = 2\sin x \cos x$; (b) $\cos(2x) = (\cos x)^2 - (\sin x)^2$; (c) $\exp(rx) = \exp(x)^r$ for nonzero $r \in K$; (d) $\exp(ix) = \cos x + i \sin x$, $\cos x = (\exp(ix) + \exp(-ix))/2$, and $\sin x = (\exp(ix) - \exp(-ix))/2i$ (assuming $i = \sqrt{-1} \in K$).

7.161. Even and Odd Formal Series. A series $F \in K[[x]]$ is *even* iff $F(-x) = F$ (see 7.160); F is *odd* iff $F(-x) = -F$. (a) Show that F is even iff $F_n = 0$ for all odd n, and F is odd iff $F_n = 0$ for all even n. (b) Which formal trigonometric and hyperbolic trigonometric series are odd? Which are even? (c) Give rules for determining the parity (even or odd) of $F + G$, FG, and (when defined) F^{-1}, given the parity of F and G.

7.162. Let R be a commutative ring. Recall that $x \in R$ is a *unit of* R iff there exists $y \in R$ with $xy = yx = 1_R$; x is *nilpotent* iff there exists $n \in \mathbb{N}^+$ with $x^n = 0_R$. (a) Suppose $x \in R$ is arbitrary and $z \in R$ is nilpotent. Prove xz is nilpotent. (b) Suppose $x \in R$ is a unit of R and $y \in R$ is nilpotent. Prove $x + y$ is a unit of R. (c) Suppose $x, y \in R$ are both nilpotent. Prove $x+y$ is nilpotent. (d) Which results in (a), (b), and (c) hold if R is a non-commutative ring?

7.163. Let R be a nonzero commutative ring. Prove that $f \in R[x]$ is a unit of $R[x]$ iff f_0 is a unit of R and f_n is nilpotent in R for all $n > 0$.

7.164. Use (7.6) in the proof of 7.65 to find the first several coefficients in the compositional inverses of each of the following series: (a) $\sin x$; (b) $\tan x$; (c) $x/(1-x)$.

7.165. Use the formal Maclaurin formula 7.55 to deduce the series expansions of e^x, $\sin x$, $\cos x$, and $(1 - rx)^{-1}$ starting from the rules for differentiating these formal series.

7.166. Taylor's formula states that (under suitable hypotheses on $f : \mathbb{R} \to \mathbb{R}$) $f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$ for all x sufficiently close to a. Give two reasons why this formula is not meaningful (as written) for formal power series, when $a \in K$ is nonzero.

7.167. (a) Show that $\sum_{n\geq 0} x^{3n}/(3n)! = (1/3)e^x + (2/3)\cos(x\sqrt{3}/2)e^{-x/2}$ in $\mathbb{C}[[x]]$. (b) Try to find similar formulas for $\sum_{n\geq 0} x^{3n+1}/(3n+1)!$ and $\sum_{n\geq 0} x^{3n+2}/(3n+2)!$.

7.168. Exponential Generating Functions. Given a sequence $F = (F_n : n \geq 0) \in K[[x]]$, the *exponential generating function* of this sequence is $F^* = \sum_{n\geq 0}(F_n/n!)x^n \in K[[x]]$. Prove that, for all $F, G \in K[[x]]$: (a) $(F+G)^* = F^* + G^*$; (b) $n!(F^*G^*)_n = \sum_{k=0}^{n} \binom{n}{k} F_k G_{n-k}$; (c) $\frac{d}{dx}(F^*) = ((F - F(0))/x)^*$.

7.169. Sum of Squares via Formal Series. The goal of this problem is to use series to derive a formula for $\sum_{k=0}^{n} k^2$ without guessing the answer in advance. (a) Express the series $\sum_{n\geq 0} n^2 x^n$ as a linear combination of the series $(1-x)^{-1}$, $(1-x)^{-2}$, and $(1-x)^{-3}$. (b) Perform a suitable operation on the series in (a) to obtain an algebraic formula for the series $\sum_{n\geq 0}(\sum_{k=0}^{n} k^2)x^n$. (c) Extract the coefficient of x^n in (b) to obtain a formula for $\sum_{k=0}^{n} k^2$ that is a polynomial of degree 3 in n. (d) Explain another way to solve this problem based on 7.90.

7.170. Use the method of 7.169 to evaluate the following sums for all n: (a) $\sum_{k=0}^{n} k$; (b) $\sum_{k=0}^{n} k^3$; (c) $\sum_{k=0}^{n} 3^k$.

7.171. Prove that for all $k, n \in \mathbb{N}^+$,

$$1^k + 2^k + \cdots + n^k = \sum_{j=0}^{k} \frac{S(k+1, j+1)}{j+1} (-1)^{k+j} (n+1)\uparrow_{j+1} .$$

7.172. State and prove a version of the quadratic formula for solving $AF^2 + BF + C = 0$, where $A, B, C \in K[[x]]$ are known series and $F \in K[[x]]$ is unknown. What hypotheses must you impose on A, B, C? Is the solution F unique?

7.173. Recursion for Divide-and-Conquer Algorithms. Many algorithms use a "divide-and-conquer" approach in which a problem of size n is divided into a subproblems of size n/b, and the solutions to these subproblems are then combined in time cn^k to give the solution to the original problem. Letting $T(n)$ be the time needed to solve a problem of size n, $T(n)$ will satisfy the recursion $T(n) = aT(n/b) + cn^k$ and initial condition $T(1) = d$ (where $a, b, c, d > 0$ and $k \geq 0$ are given constants). Assume for simplicity that n ranges over powers of b. (a) Find a recursion and initial condition satisfied by $S(m) = T(b^m)$, where m ranges over $\mathbb{N}$. (b) Use formal series to solve the recursion in (a). Deduce that, for a suitable constant C and large enough n,

$$T(n) \leq \begin{cases} Cn^k & \text{if } a < b^k \text{ (combining time dominates);} \\ Cn^k \log_2 n & \text{if } a = b^k \text{ (dividing and combining times balance);} \\ Cn^{\log_b a} & \text{if } a > b^k \text{ (time to solve subproblems dominates).} \end{cases}$$

7.174. Merge Sort. Suppose we wish to sort a given sequence of integers $x_1, \ldots, x_n$ into increasing order. Consider the following recursive method: if $n = 1$, the sequence is already sorted. For $n > 1$, divide the list into two halves, sort each half recursively, and merge the resulting sorted lists. Let $T(n)$ be the time needed to sort n objects using this algorithm. Find a recursion satisfied by $T(n)$, and use 7.173 to show that $T(n) \leq Cn \log_2 n$ for some constant C. (You may assume n ranges over powers of 2.)

7.175. Fast Binary Multiplication. (a) Given $x = ak + b$ and $y = ck + d$ (where $a, b, c, d, k \in \mathbb{N}$), verify that $xy = (ak + b)(ck + d) = ack^2 + bd + ((a + b)(c + d) - ac - bd)k$. Take $k = 2^n$ in this identity to show that one can multiply two $2n$-bit numbers by recursively computing three products of n-bit numbers and doing several binary additions. (b) Find a recursion describing the number of bit operations needed to multiply two n-bit numbers by the recursive method suggested in (a). (c) Solve the recursion in (b) to determine the time complexity of this recursive algorithm (you may assume n is a power of 2).

7.176. Formal Linear Ordinary Differential Equations. Suppose $P, Q \in K[[x]]$ are given formal series, and we wish to find a formal series $F \in K[[x]]$ satisfying the "linear ODE" $F' + PF = Q$ and initial condition $F(0) = c \in K$. Solve this ODE by multiplying by the "integrating factor" $\exp(\int P\,dx)$ and using the product rule to simplify the left side.

7.177. Formal ODEs with Constant Coefficients. For fixed $c_1, \ldots, c_k \in \mathbb{C}$, let V be the set of all formal series $F \in \mathbb{C}[[x]]$ satisfying the ODE

$$F^{(k)} + c_1 F^{(k-1)} + c_2 F^{(k-2)} + \cdots + c_k F = 0. \tag{7.12}$$

The *characteristic polynomial* for this ODE is $q = x^k + c_1 x^{k-1} + c_2 x^{k-2} + \cdots + c_k \in \mathbb{C}[x]$. Suppose q factors as $(x - r_1)^{k_1} \cdots (x - r_s)^{k_s}$ for certain $k_i > 0$ and distinct $r_i \in \mathbb{C}$. (a) Show that the k series $x^j \exp(r_i x)$ (for $1 \leq i \leq s$ and $0 \leq j < k_i$) lie in V. (b) Show that V is a complex vector space, and the k series in (a) form a basis for V. (c) Describe a procedure for expressing a given sequence $F \in V$ as a linear combination of the sequences in the basis from part (a), given the "initial conditions" $F(0), F'(0), \ldots, F^{(k-1)}(0)$. (d) Let W be the set of formal series $G \in \mathbb{C}[[x]]$ satisfying the non-homogeneous ODE

$$G^{(k)} + c_1 G^{(k-1)} + c_2 G^{(k-2)} + \cdots + c_k G = H,$$

where $H \in \mathbb{C}[[x]]$ is a given series . If G^* is one particular series in W, show that $W = \{F + G^* : F \in V\}$.

7.178. Characteristic Polynomial of a Recursion. This problem sets up an analogy between recursions with constant coefficients and ordinary differential equations with

constant coefficients. For fixed $c_1, \ldots, c_k \in \mathbb{C}$, let V be the set of all formal series $(A_n : n \geq 0) \in \mathbb{C}[[x]]$ satisfying the recursion

$$A_n = c_1 A_{n-1} + c_2 A_{n-2} + \cdots + c_k A_{n-k} \qquad (n \geq k). \tag{7.13}$$

The *characteristic polynomial* for this recursion is $q = x^k - c_1 x^{k-1} - c_2 x^{k-2} - \cdots - c_k \in \mathbb{C}[x]$. Suppose q factors as $(x - r_1)^{k_1} \cdots (x - r_s)^{k_s}$ for certain $k_i > 0$ and distinct $r_i \in \mathbb{C}$. (a) Show that the k sequences $(n^j r_i^n : n \geq 0)$ (for $1 \leq i \leq s$ and $0 \leq j < k_i$) lie in V. (b) Show that V is a complex vector space, and the k sequences in (a) form a basis for V. (c) Describe a procedure for expressing a given sequence $A \in V$ as a linear combination of the sequences in the basis from part (a). (Use the "initial conditions" $A_0, \ldots, A_{k-1}$.) (d) Let W be the set of sequences $(B_n : n \geq 0)$ satisfying

$$B_n = c_1 B_{n-1} + c_2 B_{n-2} + \cdots + c_k B_{n-k} + g(n) \qquad (n \geq k),$$

where $g(n)$ is a given function. If B^* is one particular sequence in W, show that $W = \{A + B^* : A \in V\}$.

7.179. Fill in the details of the construction of the field of fractions of an integral domain, which was sketched in 7.44. Specifically, show that: (a) the relation $\sim$ on X is reflexive, symmetric, and transitive; (b) addition and multiplication on F are well defined; (c) F, with these operations, is a field; (d) the map $i : D \to F$ is an injective ring homomorphism; (e) F satisfies the universal mapping property stated in 7.44.

7.180. Localization. The construction of fields of fractions can be generalized as follows. Let R be a commutative ring (not necessarily an integral domain), and let $S \subseteq R$ be a subset such that $1 \in S$ and $xy \in S$ whenever $x, y \in S$. Our goal is to use R to construct a new ring in which every element of S becomes a unit.

Define an equivalence relation on $X = R \times S$ by setting $(a, s) \sim (b, t)$ iff there exists $u \in S$ with $u(at - bs) = 0$. (a) Show that $\sim$ is an equivalence relation on X; let T be the set of equivalence classes. (b) Define addition and multiplication operations on T; show that these operations are well defined and make T into a commutative ring. (c) Define $i : R \to T$ by letting $i(r)$ be the equivalence class of $(r, 1)$ for $r \in R$. Show that i is a ring homomorphism (which may not be injective, however) such that $i(s)$ is a unit in T for every $s \in S$. (d) Show T has the following universal property: if U is any commutative ring and $j : R \to U$ any ring homomorphism such that $j(s)$ is a unit in U for every $s \in S$, then there exists a unique ring homomorphism $f : T \to U$ such that $j = f \circ i$.

Notes

A detailed but rather technical treatment of the algebraic theory of polynomials and formal power series is given in Bourbaki [19, Ch. IV]. More facts concerning polynomials, integral domains, and related aspects of ring theory may be found in algebra texts [8, 70, 71]. Discussions of formal power series from a more combinatorial perspective may be found in Stanley [127, Ch. 1] and Wilf [139].