1 The Ring of Formal Power Series

1.1 Basic Ring Properties

For now, let F be any arbitrary field (not an arbitrary ring). One might notice that if we take the so called **Fibonacci Power Series**

$$1 + x + 2x^2 + 3x^3 + 5x^4 + 8x^5 + 13x^6 + \cdots$$

we can claim a certain "formal" equality

$$1 + x + 2x^2 + 3x^3 + 5x^4 + \cdots = \frac{1}{1-x-x^2}.$$

Is this equation true in any sense? We might notice if we multiply by $1-x-x^2$ on both sides and join common monomial terms we get the trivial equality $1 = 1$. But what exactly does this equality mean? Note that this equation is true in the traditional power series sense. It has a non-zero radius of convergence. But what about other power series and formulae?

The point of introducing formal power series is to make equalities like the above actual equalities in a certain very well-defined sense, without worrying at all about issues like convergence, especially if the field we are working under is not the field of real or complex numbers.

Definition 1. Let F be a field. The formal power series $F[[x]]$ consists of sequences $(a_0, a_1, a_2, \dots)$ where the $a_i \in F$. We pretty much always represent these sequences by $a_0 + a_1x + a_2x^2 + \cdots$. Given two formal power series

$$a = \sum_{n=0}^{\infty} a_n x^n$$

and

$$b = \sum_{n=0}^{\infty} b_n x^n,$$

define their sum $a + b$ to be

$$a + b = \sum_{n=0}^{\infty} (a_n + b_n) x^n$$

and define their product ab to be

$$ab = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} a_k b_{n-k} \right) x^n$$

We can also define things like the formal derivative, which is unambiguous no matter the field F we are working in, since n can be interpreted as $1 + 1 + \cdots + 1$ repeated n times.

Proposition 1. *The set of formal power series $F[[x]]$ forms a commutative ring (with unity) under the operations defined above.*

Proof. Hopefully the following claims are fairly clear:

- The formal "zero" $0 + 0x + 0x^2 + \cdots$ is a zero element.
- The formal "one" $1 + 0x + 0x^2 + \cdots$ is a multiplicative identity.
- Every element of a formal power series has an additive inverse.
- Addition and multiplication in $F[[x]]$ are associative and commutative. This follows from commutativity and associativity of the base field F.

The first, fourth and third claims make clear that this ring is an abelian group under addition. The second and fourth items make clear the sense that the multiplication operation in $F[[x]]$ has the appropriate ring structure. Distributivity is straightforward, though a little tedious, to check. ■

In this ring, what is the so called group of units (that is, elements with a multiplicative inverse)?

Proposition 2. *The formal power series $a_0 + a_1x + a_2x^2 + \cdots$ has no multiplicative inverse if $a_0 = 0$ and has a unique multiplicative inverse otherwise.*

Proof. Given any such formal power series $a_0 + a_1x + a_2x^2 + \cdots$ we will suppose that such an inverse exists. Then the relations

$$\begin{aligned} a_0b_0 &= 1 \\ a_0b_1 + a_1b_0 &= 0 \\ &\vdots \\ \sum_{k=0}^{n} a_kb_{n-k} &= 0 \\ &\vdots \end{aligned}$$

must hold. It is clear from the first equation that no such inverse exists if $a_0 = 0$. Now suppose $a_0 \neq 0$. Then from the general equation for n we can solve for b_n as

$$b_n = \begin{cases} \frac{1}{a_0} & n = 0 \\ -\frac{1}{a_0}\sum_{k=1}^{n} a_kb_{n-k} & \text{otherwise.} \end{cases}$$

It follows inductively that if $a_0 \neq 0$ then b_n is always well defined in terms of the a_i and thus a well-defined inverse for $a_0 + a_1x + a_2x^2 + \cdots$ exists. Hence the claim follows, as desired. ■

The following proposition shows that $F[[x]]$ is an integral domain, which is a particularly nice property.

Theorem 1. *If f, g, and h are formal power series in $R[[x]]$, where R is an integral domain, then*

$$fh = gh \implies f = g.$$

So $R[[x]]$ is an integral domain.

Proof. It suffices to prove that if f and g are non-zero power series then $fg \neq 0$ (for then, in the language of our original statement, we would have $(f-g)h = 0$ with $h \neq 0$, implying $f = g$). Let $f = a_0 + a_1x + a_2x^2 + \cdots$ and let $g = b_0 + b_1x + b_2x^2 + \cdots$. Let i be the least index such that $a_i \neq 0$ and let j be the least index such that $b_j \neq 0$. Then it follows that the coefficient of x^{i+j} of gh is

$$\begin{aligned}\sum_{n=0}^{i+j} a_k b_{i+j-k} &= \sum_{k=0}^{i-1} a_k b_{i+j-k} + a_i b_j + \sum_{k=i+1}^{i+j} a_k b_{i+j-k} \\ &= \sum_{k=0}^{i-1} a_k b_{i+j-k} + a_i b_j + \sum_{k=0}^{j-1} a_{i+j-k} b_k \\ &= 0 + a_i b_j + 0 \neq 0\end{aligned}$$

since $a_i \neq 0$, $b_j, \neq 0$, and $a_k = 0$ for all $k < i$ and $b_l = 0$ for all $l < j$ (by definition). So fg is not the zero element, as desired. ∎

1.2 Convergence of formal power series and an underlying topology on $F[[x]]$

One deficit of the ring of formal power series is that unlike the usual notion of power series, which have radius of convergence, there might not be a useful way to evaluate a power series outside of $x = 0$. Such evaluation might not make sense if F is not $\mathbb{R}$ or $\mathbb{C}$. However, we still do want to say something about "formal" convergence of a power series. As an example, we would like to say that the sequence of power series $a_0, a_0 + a_1x, a_0 + a_1x + a_2x^2, \ldots$ converges to the associated power series $a_0 + a_1x + a_2x^2 + \cdots$. To make this notion precise we will precisely define the notion of convergence (which will hopefully be enough for our purposes, since strictly speaking convergence of sequences does not determine a topology if your underlying space is not metrizable or first countable *I might flesh this out a little bit later, considering that the convergence I've defined is just R^ω where R gets the discrete topology and we endow the product with the product topology*). We will also use this notion of convergence to give a criterion about whether or not an infinite sum of *formal power series* converges.

Definition 2. We will say a sequence $r_1, r_2, r_3, \ldots$ of elements in F *converges sharply* to some limit s if r_n is eventually constant and equal to s.

Definition 3. A seqeunce of formal power series $f_1, f_2, \ldots$ converges to a formal power series g if the sequence (r_n^i) obtained by taking the x^i coefficient of each f_n converges sharply to the x^i coefficient in g.

For example, the sequence of formal power series'

$$\begin{aligned}
&1 + x + x^2 + x^3 + \cdots \\
&1 + 2x + 2x^2 + 2x^3 + \cdots \\
&1 + 2x + 3x^2 + 3x^3 + \cdots \\
&1 + 2x + 3x^2 + 4x^3 + \cdots \\
\ddots&
\end{aligned}$$

converges to the formal power series $1 + 2x + 3x^2 + 4x^3 + 5x^4 + \cdots$.

Lemma 1. *In $F[[x]]$ the sequence $a_0, a_0 + a_1x, a_0 + a_1x + a_2x^2, \ldots$ converges to the formal power series $f(x) = a_0 + a_1 + a_2x^2 + \cdots$.*

The proof is very simple.

Proof. For any $k \geq 0$, the degree k term of $a_0 + \cdots + a_nx^n$ is equal to the degree term of $f(x)$ for sufficiently large n (specifically if $n \geq k$). ■

Theorem 2. *Let $f_1, f_2, f_3, \ldots$ be formal power series. Then the infinite series $f_1+f_2+f_3+\cdots$ exists if and only if $\text{ord}(f_k) \to \infty$ where*

$$\text{ord}(f) = \min(k \mid a_k \neq 0)$$

(provided that $f = a_0 + a_1x + a_2x^2 + \cdots$).

Proof. Let $p_n(f_k)$ denote the coefficient of x^n in f_k.

First suppose $\text{ord}(f_k) \to \infty$. This means that for any $n \geq 0$ there exists $N > 0$ such that $p_n(f_k) = 0$ for all $k \geq N$. It follows that $p_n(f_1 + \cdots + f_k)$ is constant for all $k \geq N$, implying for all n that each coefficient converges sharply. Hence by definition $f_1 + f_2 + f_3 + \cdots$ exists.

Conversely, suppose that the series $f_1+f_2+f_3+\cdots$ exists. Then for all $n \geq 0$ there exists $N_n \geq 0$ such that $p_n(f_1+\cdots+f_k)$ is constant for all $k \geq N_n$. But this implies that $p_n(f_k) = 0$ for all $k > N_n$ (because p_n is actually a ring homomorphism). Taking $N = \max_{1\leq i\leq n}(N_i)$, we see that $\text{ord}(f_k) > n$ for $k \geq N$. It follows that $\text{ord}(f_k) \to \infty$ as desired. ■

Some examples which illustrate this theorem is appropriate. The sum of series

$$1 + (x + x^2 + x^3 + \cdots) + (x + x^2 + x^3 + \cdots)^3 + \cdots$$

converges, since we can see that $\text{ord}(f_k) = \text{ord}(f_{k-1})$ for $k \geq 1$. However, the series

$$1 + \frac{1}{2} + \frac{1}{4} + \cdots$$

does not, which is a potential limitation of the way we defined convergence. However, for most things we need these power series for (combinatorics) these limitations do not come up in practice.

Theorem 3. *Suppose $a_i \to a$ and $b_i \to b$, where a_i, b_i, a, b live in $F[[x]]$. Then we have that*

$$a_i + b_i \to a + b$$

and

$$a_i b_i \to ab.$$

It follows that $F[[x]]$ is a topological ring: that is, the addition and multiplication operators are continuous operations that respect convergence.

Proof. Omitted for now. ■

Theorem 4. *Any series which is convergent in $F[[x]]$ is absolutely convergent. In particular, given any convergent infinite sum of power series we can rearrange any number of terms without changing the equality.*

Proof. Omitted for now. ■

With the above facts, we are ready to do some interesting things with formal power series which are completely rigorous.

Theorem 5. *Suppose $f \in F[[x]]$ is some power series where the constant term is 0 (so $\operatorname{ord}(f) = 1$). Then*

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

Proof. First observe by hypothesis ($\operatorname{ord}(f) = 1$ and $f(0) = 0$) that the terms on the left hand and right hand sides are well defined formal power series. By definition, g is the limit of the power series

$$g_n(x) = \sum_{k=0}^{n} (f(x))^k.$$

We observe that

$$(1-f)(g_n) = 1 - f^{n+1}.$$

Since $\operatorname{ord} f \geq 1$ we can deduce that $f^{n+1} \to 0$ and hence $(1-f)g_n \to 1$. But since $g_n \to g$ it follows that $(1-f)g = 1$, as desired.

Alternatively, using absolute convergence we have that

$$(1-f)g = (1-f) + (f - f^2) + (f^2 - f^3) + \cdots = 1,$$

as desired. ■

From this theorem we can deduce some interesting patterns. For example, we obtain the usual geometric series formula

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

We can use this to generate power series for the inverses of $(1-x)^n$.

$$\begin{aligned}
\frac{1}{1-x} &= 1+x+x^2+x^3+\cdots\\
\frac{1}{(1-x)^2} &= 1+2x+3x^2+4x^3+\cdots\\
\frac{1}{(1-x)^3} &= 1+3x+6x^2+10x^3+\cdots\\
\frac{1}{(1-x)^3} &= 1+4x+10x^2+20x^3+\cdots
\end{aligned}$$

Something very interesting is going on here! It appears as if the terms of these expansions line up with binomial coefficients. But why are the expansions like this?

2 Homogeneous Linear Recurrence Equations and the magic of Linear Algebra

One of the most basic combinatorial questions that one can ask is "What is the number/how many objects are there?" In the last section we set up the machinery of *generating functions*, where we put a parametrized family into a sequence and we operated on it using a so called "generating function" approach.

Here is an example of the power of this approach. Consider the problem of finding a formula for the sequence

$$1, 2, 5, 14, 41, \ldots$$

satisfying $f(n+2) = 4f(n+1) - 3f(n)$ for $n \geq 0$.

Setting up the generating function

$$F(x) = \sum_{k=0}^{\infty} f(n)x^n$$

we rearrange and we find

$$\begin{aligned}
F(x) &= \sum_{k=0}^{\infty} f(n)x^n\\
&= f(0) + f(1)x + \sum_{k=2}^{\infty} f(n)x^n\\
&= f(0) + f(1)x + \sum_{k=0}^{\infty}(4f(n+1) - 3f(n))x^{n+2}\\
&= f(0) + f(1)x - 4x + 4x\sum_{k=0}^{\infty} f(n)x^n - 3x^2\sum_{k=0}^{\infty} f(n)x^n\\
&= f(0) + (f(1) - 4)x + (4x - 3x^2)F(x)
\end{aligned}$$

which implies that

$$F(x)(1-4x+3x^2) = f(0) + (f(1)-4)x.$$

Conversely, any formal power series of the form $\frac{ax+b}{1-4x+3x^2}$ is a solution to the recurrence relation. Since $1-4x+3x^2 = (1-x)(1-3x)$, by partial fraction decomposition we have

$$\frac{ax+b}{1-4x+3x^2} = \frac{c}{1-x} + \frac{d}{1-3x}.$$

By the geometric series formula of last chapter, this is equal to $c(1+x+x^2+x^3+\cdots) + d(1+3x+9x^2+27x^3+\cdots)$. We conclude that

$$f(n) = c + d3^n$$

for arbitrary c and d are all the solutions to this recurrence relation.

In this section, instead of using the ring structure of the formal power series $F[[x]]$ we will put a vector space structure on the set of infinite sequences.

Observe that the equation $f(n+2) = 4f(n+1) - 3f(n)$ is **linear**. That is, if f_1, and f_2 satisfy this recurrence relation, so does $g = Af_1 + bf_2$, where A and B are arbitrary constants. This implies that the set of solutions to this linear recurrence relation is a vector space (namely, a subspace of the $\mathbb{F}^\omega$ sequence space). Here are some elements of this vector space:

$$\begin{aligned}(1,1,1,1,1,\dots)\\(1,3,9,27,\dots)\\(1,0,-3,-12,-39,\dots)\\(0,1,4,13,40,\dots)\end{aligned}$$

This vector space is 2-dimensional: every sequence is completely determined by the first two elements. One might notice that either the first two elements above or the last two elements above form a basis. The first two elements are nice in that they have simple formulas, namely 1^n and 3^n. The last two are nice because they can be seen to be like a standard basis for the vector space, and you can easily express any other sequence in the vector space in terms of a linear combination of the two.

Definition 4. Any recurrence relation of the form

$$\sum_{j=0}^{k} A_j f(n+j) = 0$$

where the A_j are constants is called a kth order homogeneous linear recurrence relation with constant coefficients.

There are certainly important recurrence relations where the coefficients are non-constant, consider the simple example

$$(n+1)! = (n+1)(n!).$$

For the rest of this section, let S be the vector space of all sequences

$$f = (f(0), f(1), f(2), \dots)$$

called the *sequence space.* For any sequence f, define $T(f)$ (or Tf) as the sequence whose value at n is $f(n+1)$. In other words, T is the "left shift" operator. It is easy to see that T is linear. If we view it as an infinite diagonal matrix, then T is zero everywhere except 1 on the diagonal just above the main diagonal.

Now observe that the recurrence relation

$$f(n+2) - 4f(n+1) + 3f(n) = 0$$

can be written as

$$T^2(f) - 4T(f) + 3f = 0.$$

Note that the 0 in the above equation is the zero sequence $(0, 0, 0, \dots)$, and not the zero of the field. T^2 denotes $T \circ T$.

This act of representing recurrence relations as polynomials of this shift operator is very powerful, as we shall see. Let's see what $T - 3I$ does to a vector of the form $f(n) = r^n$. We have that

$$\begin{aligned}(T - 3I)f(n) &= Tf(n) - 3If(n)\\ &= r^{n+1} - 3r^n\\ &= (3 - r)r^n.\end{aligned}$$

That is, f is an eigenvector of $T - 3I$ with eigenvalue $r - 3$. In particular, if $r = 3$, then f is annihilated by $T - 3I$. Observe that $(T - I)(T - 3I)$ annihilates $f(n) = 3^n$ and $(T - 3I)(T - I)$ annihilates $f(n) = 1^n$.

But by linearity, both these operators are equal! Hence any linear combination of $f_2(n) = 3^n$ and $f_1(n) = 1$ is in the kernel of $T^2-4T+3I$, and therefore satisfies the recurrence relation we were interested in above.

Let's introduce some new terminology. The operator $T - I$ (also sometimes written $T-1$) is called the *difference operator* (compare with the derivative operator). The equation $p(T)f = 0$ for any polynomial p is called a *difference equation.*

Here is the first result of a more general theorem:

Theorem 6. *Let $p(t) = a_d t^d + a_{d-1}t^{d-1} + \cdots + a_0$, and consider the equation $p(T)f = 0$, a homogeneous linear recurrence equation with constant coefficients.*

Suppose $p(t)$ has d distinct roots. Then the general solution to the linear recurrence relation $p(T)f = 0$ is of the form

$$f(n) = A_1 r_1^n + A_2 r_2^n + \cdots + A_d r_d^n,$$

where the r_i are the roots of $p(t)$.

Proof. It is easy to check that $f(n) = r_i^n$ is a solution to the linear recurrence equation we are considering. These d functions are linearly independent, since they are all eigenvectors of the left shift operator T with distinct eigenvalues. Snce we know that the subspace of the sequence space consisting of $\ker(p(T))$ is d-dimensional, we are done. ■

Note that even if the terms of the sequence $f(n)$ are integers, the numbers r_i and A_i need not be.

But what if our polynomial doesn't have repeated roots? For example, take $p(t) = t^2 - 2t + 1$. Using a standard method (such as the one in the beginning of this section), we get a 2 parameter family of generating functions

$$\frac{ax+b}{x^2-2x+1}.$$

When $a = -1$ and $b = 1$ we can simplify to get the formal power series

$$1 + x + x^2 + x^3 + \cdots .$$

Put $a = 1$, $b = 0$ to get

$$\frac{x}{1-2x+x^2} = x(1-x)^{-2} = x(1+2x+3x^2+4x^3+\cdots) = x+2x^2 = 3x^3 + \cdots .$$

then we have two "fundamental solutions" $f(n) = 1$, $g(n) = n$.

In general, it turns out that if $p(T) = (T - rI)^d$, there are d fundamental solutions $f(n) = n^k r^n$ for $k = 0$ to $d - 1$. When $r = 1$, these fundamental solutions are polynomials and the general solution is a polynomial of degree less than or equal to d.

Theorem 7. *If $p(t)$ is degree d, with leading coefficient and constant term non-zero, the solutions of $p(T)f = 0$ form a d-dimensional space with basis elements of the form $n^k r^n$, where r is any root of $p(t) = 0$ and k is any non-negative integer less than the multiplicity of r in p.*

Proof. omitted for now ■

Theorem 8. *If f satisfies a homogeneous linear recurrence equation of order d, then $f, Tf, T^2 f, \ldots, T^d f$ are linearly dependent.*

Conversely, if $f, Tf, T^2 f, \ldots$ span a d-dimensional sequence space, then f satisfies a linear recurrence equation of order d.

Proof. Exercise to reader; follows from basic properties of dimension in finite dimensional vector spaces. ■