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 $a_0 + a_1 x + a_2 x^2 + \cdots$. 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 $\operatorname{ord}(f_k) \to \infty$ where*

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

(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 $\operatorname{ord}(f_k) \to \infty$. 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 $\operatorname{ord}(f_k) > n$ for $k \geq N$. It follows that $\operatorname{ord}(f_k) \to \infty$ 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 $\operatorname{ord}(f_k) = \operatorname{ord}(f_{k-1})$ 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 $(1-x)^n$.

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

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 Proof of the Formal Binomial Theorem

In this section we will prove what is known as the formal binomial theorem for power series. (To fill in later we will consider why we only consider power series of the form $(1 + xf(x))$ and why we can only hope to define a general exponentiation there).

For any formal power series $f(x)$ define

$$(1 + xf(x)) * r = 1 + rxf(x) + \left(\frac{r(r-1)}{2}\right) x^2 f(x)^2 + \left(\frac{r(r-1)(r-2)}{3!}\right) x^3 f(x)^3 + \cdots .$$

Theorem 6. *The following statements are all true:*

1. *The right hand side converges in the ring of formal power series.*
2. *When $r \geq 0$ is an integer, then $(1 + xf(x)) * r = (1 + xf(x))^r)$.*
3. *We have that for all r and s,*
$$((1 + xf(x)) * r)((1 + xf(x)) * s) = (1 + xf(x)) * (r + s).$$
4. *When $r < 0$ is an integer, then $(1 + xf(x)) * r = (1 + f(x))^r$.*
5. *For all r and s,*
$$((1 + xf(x)) * r) * s = (1 + xf(x)) * (rs)$$

Proof. The first statement follows readily from the fact that each succeeding term has order greater than the previous. The second statement is a reformulation of the usual binomial theorem.

For the third statement, for any integer $k \geq 0$, the coefficient of $f(x)^k x^k$ in the absolute product is a polynomial $P_k(r, s)$ of degree k in r and s. For example, when $k = 2$, we have that this coefficient is

$$1 \cdot \left(\frac{s(s-1)}{2}\right) + rs + \left(\frac{r(r-1)}{2}\right) \cdot 1.$$

Likewise the coefficient of $x^k f(x)^k$ in $(1+xf(x))*(r+s)$ is a polynomial $Q_k(r,s)$ of degree k in r and s. We know that $P_k(r,s) = Q_k(r,s)$ when $r, s \in \mathbb{N}$. So $P_k = Q_k$ for infinitely many points. So we use a $k+1$ point equality argument 3 times in order to conclude that $P_k = Q_k$ in general. This implies equality of the power series.

The fourth statement follows directly from the second and third statements.

The fifth statement has an argument which is the exact same reasoning as the third statement.

■

By this theorem, we can introduce the convention of just writing $g(x) * r$ as $g(x)^r$ for any g with $g(0) = 1$. In particular we have

$$\begin{aligned}(1-x)^{-m} &= 1 + (-m)(-x) + \left(\frac{(-m)(-m-1)}{2}\right)(-x)^2 + \left(\frac{-m(-m-1)(-m-2)}{3!}\right)(-x)^3 + \cdots \\ &= 1 + mx + \frac{m(m+1)}{2}x^2 + \frac{m(m+1)(m+2)}{6}x^3 + \cdots \\ &= \sum_{n=0}^{\infty} \binom{m+n-1}{n} x^n.\end{aligned}$$

This explains the connection with expansions of $(1-x)^{-m}$ and binomial coefficients that we saw earlier.

3 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 7. *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 8. *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 9. *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. ■

4 More comments on HLREs

The sequence $\{F_n\}$ of fibonacci numbers is annihilated by the operator

$$T^2 - T - I$$

where T is the left shift operator. The generating function has denominator $1 - x - x^2$. Note that the coefficients appear flipped. This is true in genreal, as the following theorem indicates.

Theorem 10. *Suppose $a_0, \ldots, a_d$ are non-zero coefficients, and that f is some sequence. The following statements are equivalent:*

- *f satisfies the equation*

$$(a_d T^d + a_{d-1} T^{d-1} + \cdots + a_1 T + a_0) f = 0.$$

- *The generating function for f can be expressed as a rational function whose numerator is of degree less than d, with denominator*

$$a_d + a_{d-1} x + \cdots + a_0 x^d$$

- *$f(n)$ is expressible as a linear combination of functions of the form $n^j r^n$ where r is a root fo $a_d t^d + \cdots + a_1 t + a_0 = 0$ and where $j \geq 0$ is less than the multiplicity of r.*

Observe the following: if the roots of such a polynomial are ± 1, then the solutions to such a recurrence relation are simply polynomial equations in the variable n.

Let's consider some of these results in action. Suppose we'd like to find either a recurrence or closed form formula for some sequence, say,

$$1, 2, 5, 12, 29, 70, \cdots$$

satisfying a linear recurrence relation.

If we suspect that such a relation is second order, we can take the system of equations

$$\begin{aligned} 5 &= A(2) + B(1) \\ 12 &= A(5) + B(2) \end{aligned}$$

From which we get $A = 2$ and $B = 1$. This gives us the candidate linear recurrence relation $f(n) = 2f(n-1) + f(n-2)$, which can be seen to work. But what if we suspected that the relation was third order instead? Then we get the linear recurrence

$$f(n) = 3f(n-1) - f(n-2) - f(n-3).$$

Is there any relation between these two recurrence relations? It seems like the former is "simpler" because it involves less previous terms. It turns out that this intuition is correct. Here are some facts about the kinds of sequences you get.

Proposition 3. *1. Suppose p and q are polynomials such that $p(T)f = 0$ and $p \mid q$. Then $q(T)f = 0$.*

2. For any sequence f satisfying an LRE with constant coefficients, there exists a unique monic polynomial $p(t)$ such that $p(T)f = 0$ and $q(T)f = 0$ implies that $p \mid q$.

3. Any sequence satisfying an LRE satisfies a linear recurrence relation satisfies a recurrence relation with a minimal number of terms and which is unique up to scalar multiples.

Proof. The first item is more or less obvious. The second item follows from the fact that the set of polynomials p such that $p(T)f = 0$ form an ideal in the polynomial ring and there is a unique monic polynomial which is the generator for said ideal. The third item more or less follows from the second item. ■

For example, in the above example, the polynomial operators for each recurrence relation are

$$\begin{aligned} & T^2 - 2T - I \\ T^3 - 3T^2 + T + I &= (T^2 - 2T - I)(T - I) \end{aligned}$$

Let's consider a less trivial example. Let's find a recurrence relation for the sequence $g(n) = F_n + 2^n$, where F_n is the nth Fibonacci number. If we suspect that this sequence satisfies a third order linear recurrence relation, we can solve a 3 dimensional system to get that

$$g(n) = 3g(n-1) - g(n-2) - 2g(n-3).$$

One can confirm that $g(n)$ indeed satisfies the recurrence relation above.

Here is a slicker method: We might use linearity, and observe that $(T^2 - T - I)(T - 2I)$ annihilates the $g(n)$. This gives us the desired recurrence relation. Or similarly, we add

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

To test the minimality of this recurrence relation, we can show that $\gcd(2 - 3x - x^2, 1 - 3x + x^2 + 2x^3) = 1$, or we can show that the determinant of cycles of the first equation is non-zero (that there are no non-trivial dependence relations in this recurrence, showing it is minimal).

Finally, we conclude this section with some combinatorics: we will tile a $2 \times n$ rectangle by dominos (known as a domino tiling).

This is a classic problem. We can solve it with the following observation. Either the $2 \times n$ rectangle begins with a vertical rectangle or two horizontal rectangles. Thus for $n \geq 2$ the number of rectangles D_n satisifies the recurrence relation

$$D_n = D_{n-1} + D_{n-2}.$$

It follows that the number of tilings is counted by the Fibonacci numbers.

What about a $3 \times n$ rectangle? (to be illustrated later).

4.1 A short comment on various bases

Let's consider the subspace of solutions to say the equation $(T - rI)^m f = 0$ where T is the left shift operator. One such basis is given by

$$f(n) = r^n, f(n) = nr^n, \ldots, f(n) = n^{m-1}r^n.$$

Instead, taking the generating function point of view, A natural basis for the space of generating functions these sequences is

$$\frac{1}{(1-rx)^m}, \frac{x}{(1-rx)^m}, \frac{x^2}{(1-rx)^m}, \ldots, \frac{x^{m-1}}{(1-rx)^m}.$$

Another natural basis is

$$\frac{1}{1-rx}, \frac{1}{(1-rx)^2}, \ldots, \frac{1}{(1-rx)^m}.$$

These basis elements correspond to the solutions $f_{r,k}(n) = \binom{n+k-1}{n} r^n$.

5 The Calculus of Finite Differences

In this section we investigate some special properties of the basis

$$1, x, \frac{x(x-1)}{2}, \frac{x(x-1)(x-2)}{6}, \ldots$$

of polynomial space.

Suppose for the first example that we're given a sequence

$$0, 1, 5, 15, 34, 65, 111, \ldots$$

and we suspect its formula is given by a cubic polynomial. How do we find it?

One primitive way is to use undetermined coefficients, or Lagrange interpolation. Here is another way. First, we observe that if $f(n)$ is a polynomial of degree d, then $(T - I)f$ is degree $d - 1$. So we can construct a difference table.

(difference table here).

Conversely, if the nth difference is 0, then f is an nth degree polynomial and lower. We observe the terms on the very right of the difference table. If we look at the difference tables for the polynomials above, we find that we can represent our cubic polynomial using this basis in a natural way. It will follow that

$$p(x) = 0 \cdot 1 + 1 \cdot x + 3\left(\frac{x(x-1)}{2}\right) + 3\left(\frac{x(x-1)(x-2)}{6}\right) = \frac{x^3 + x}{2}.$$

This basis above is called the basis of falling factorials. We'll denote the kth element (k starts index at 0) as $(x)_k$.

The falling factorial polynomials form an integral basis in $P[x]$ with the property that $p(n)$ is an integer for all integers n. In other words, any polynomial $p(n)$ can be represented as a linear combination of finitely many basis polynomials where the coefficients are integers.

Proof. The converse is hopefully obvious, we'll prove the forward direction. Suppose $p(n)$ is an integer for all n. Then

$$p(x) = a_0(x)_0 + a_1(x)_1 + \cdots + a_n(x)_n.$$

The idea is that plugging in $x = 0$ we will find that $a_0 \in \mathbb{Z}$. Then we continue inductively to deduce that $a_j \in \mathbb{Z}$ for all coefficients a_j. ■

5.1 Application of the falling factorial basis: Dobinski's formula for the Bell Numbers

Let B_n be the number of **partitions** of an n-element set (called Bell numbers). This is the same thing as dividing the set into chunks such that every element is in exactly one chunk. For example, when $n = 3$, $S = \{a, b, c\}$ we can represent the partitions as follows:

$$\{\{a, b, c\}\}, \{\{a, b\}, \{c\}\}, \{\{a, c\}, \{b\}\}, \{\{c, b\}, \{a\}\}, \{\{a\}, \{b\}, \{c\}\}.$$

It is easy to prove that I've listed all of them.

So it's simple to show that $B_3 = 5$, $B_2 = 2$, and $B_1 = 1$. We would like to try and find a formula for the B_n. This will involve a couple of magic tricks.

The first thing to do is to consider a seemingly unrelated question: given an n element set N and an m element set M, how many functions $f : N \to M$ are there? We can use a left-right approach here. First we notice that the answer is obviously m^n. Alternatively, we might notice that given any function $f : N \to M$ we can create a partition π of N by taking the non-empty preimages $f^{-1}(\{i\}), i \in M$.

Let $S_{n,k}$ represent the number of partitions of N of size k. Note that any such partition does not uniquely determine a function f. We also need to determine the values in M which each partition is sent to, there are $m!/(m-k)!$ such ways we can do this for a partition of size k. Hence we obtain the identity

$$m^n = \sum_{k=1}^{n} S_{n,k} \frac{m!}{(m-k)!}.$$

Since this is true for all positive integer values of m and it is a polynomial identity we conclude it actually holds for general x and hence we have the formula

$$x^n = \sum_{k=1}^{n} S_{n,k}[x]_k,$$

where $[x]_0 = 1, [x]_k = x(x-1)(\cdots)(x-k+1)$.

We observe that the set of polynomials $[x]_k$ actually form a basis for the set of real valued polynomials (note that as defined $[x]_k$ has degree k). So we might define a linear functional $L : \mathbb{R}[x] \to \mathbb{R}$ as being the unique functional with the property that

$$L([x]_k) = 1$$

for all k. Applying L to the formula above we get

$$L(x^n) = \sum_{k=1}^{n} S_{n,k} = B_n,$$

for the $S_{n,k}$ counts the number of partitions of N of size k, and we are summing over all the possible values of k.

Here is where the real magic happens: We observe that

$$L([x]_n) = 1 = \frac{1}{e}e = \frac{1}{e}\sum_{k=0}^{\infty}\frac{1}{k!} = \frac{1}{e}\sum_{k=0}^{\infty}\frac{[k]_n}{k!}.$$

To justify the last equality just observe that the infinite sum is just the previous one shifted over n places to the right (using the cancellation of $[k]_n$ and $k!$). It follows by the properties of the basis and the fact that any polynomial $p(x)$ is a finite sum of the basis elements $[x_k]$ that we get

$$L(p(x)) = \frac{1}{e}\sum_{k=0}^{\infty}\frac{p(k)}{k!}.$$

In particular, when $p(x) = x^n$, we get

$$B_n = L(x^n) = \frac{1}{e}\sum_{k=0}^{\infty}\frac{k^n}{k!}$$

which is called *Dobinski's formula* for the Bell numbers.

6 Weighted Enumeration

Previously, we looked at the number of ways to tile a $2 \times n$ rectangle with dominoes, and we found that the number of such tilings was a Fibonacci number. Here is another question: If we tile a $2 \times n$ rectangle with dominoes, then what fraction of the tiles will be vertical? This question is not exactly formed, as it depends on the tiling. Here is a better question: What is the expected number of tilings? Here's the magic technique that we will use: we will count with polynomials.

Definition 5. Given any $2 \times n$ tiling, give it weight w^k if it has exactly k vertical dominoes. The **number** of tilings by dominoes is defined as the sum of the weights. We denote this polynomial by $p_n(w)$.

For example, for a 2×3 tiling the number of tilings is $w^3 + 2w$. If we evaluate using $w = 1$ we get the actual *number* of tilings. Here are the first few examples of tilings.

$$\begin{aligned} p_0(w) &= 1(?) \\ p_1(w) &= w \\ p_2(w) &= w^2 + 1 \\ p_3(2) &= w^3 + 2w. \end{aligned}$$

Can we find a recurrence for $p_n(w)$ similarly to how we found a recurrence for ordinary tilings of the $2 \times n$ rectangle? It turns out that we can. Here is the reasoning: We can decompose any $2 \times n$ tiling as a tiling where there is one vertical tile on the left or 2 horizontal tiles on the right. But for each tiling of a $2 \times (n-1)$ rectangle the corresponding $2 \times n$ tiling will have one more vertical tile. For the $2 \times (n-2)$ rectangles no new vertical tiles are added. It follows that the recurrence for these polynomials is

$$p_n(w) = wp_{n-1}(w) + p_{n-2}(w).$$

This is the essence of weighted enumeration.

We have observed above that $p_n(1)$ is the total number of tilings. Observe that $p_n'(1)$, the formal derivative of $p_n(w)$ evaluated at 1, is the total number of **all** vertical tiles in all the $2 \times n$ tilings. To see this we need only observe that w^k represents a tiling with k vertical dominoes. When you take the derivative you get kw^{k-1}, which evaluated at $w = 1$ gives you the number of vertical dominoes in that tiling.

It follows that the average or expected number of vertical tiles in a $2 \times n$ tiling is $\frac{p_n'(1)}{p_n(1)}$. We already know a formula for $p_n(1)$ (as a Fibonacci number). What about the numerator? Here is the magic trick.

Let

$$\begin{aligned} P(w,x) &= \sum_{n=0}^{\infty} p_n(w)x^n \\ &= 1 + wx + (w^2+1)x^2 + (w^3+2w)x^3 + \cdots \end{aligned}$$

as a formal power series in $\mathbb{Q}[w][[x]]$. Using the recurrence for the weighted polynomials we can calculate that

$$P(w,x)(q - wx - x^2) = 1 \implies P(w,x) = \frac{1}{1 - wx - x^2}$$

as the generating function for $P(w,x)$. Observe now (by differentiation with respect to w in the formal power series) that

$$\begin{aligned} \sum_{n=0}^{\infty} p_n'(w)x^n &= \frac{d}{dw} \sum_{n=0}^{\infty} p_n(w)x^n \\ &= \frac{d}{dw} P(w,x) \\ &= \frac{x}{(1 - wx - x^2)^2}. \end{aligned}$$

So we get a formula for $p_n'(1)$. Letting α and β be the reciprocal roots for $1 - x - x^2$ we get a formula of the form $A\alpha^n + B\beta^n + A'n\alpha^n + B'n\beta^n$ for $p_n'(1)$. Letting α be the larger root (in magnitude) this means $p_n'(1) \sim Cn\alpha^n$ for some constant C. It follows that approximately,

$$\frac{p_n'(1)}{p_n(1)} = \frac{C_1 n\alpha^n}{C_2\alpha^n} = Cn$$

for some constant C.

7 Path Enumeration, Directed Acyclic Graphs

Here is another problem to consider: How many length n sequences of 0's and 1's are there such that no two ones appear consecutively? We might calculate them out for short sequences and we might discover that they are equal to a shift in the Fibonacci numbers. We will solve this by considering paths through a network. A sequence is specified by paths through the network. It should be clear that each sequence is specified by a path in the graph.

(add the graph here)

We will solve this essentially using dynamic programming. Suppose we would like to calculate the paths from the source vertices on the left to the sink vertex s. We can first calculate the number of paths from the vertices directly adjacent to s, for each of those vertices there is only one path. Then for the vertices directly left of those, for the top vertex we can go to either vertex, so there are $1+1$ paths to s. For the bottom vertex we can only travel to the top right vertex so there is only 1 path to s. And so on.

We might notice that we can represent the following operation by the matrix multiplication

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

and we get the number of paths of the second right most vertices to s and to t respectively. Then we can iterate by using the matrix multiplication

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^k \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

and the total number of paths to either s or t is given by

$$\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

In this section we will generalize this result sufficiently to work with graphs that are similar to the one above.

Definition 6. A **finite directed acyclic graph** (DAG) is a directed graph with only a finite number of paths between vertices and no infinite paths (no cycles). For vertices x, y in G we denote of paths from x to y by $N(x, y)$. Note that in general $N(x, y) \neq N(y, x)$ (and in fact, if $N(x, y) \neq 0$ it follows that $N(y, x) = 0$).

Theorem 11. *Suppose G is a finite DAG and $x, y \in G$. Suppose that B is a non-empty set of vertices with the property that every path from x to y passes through exactly one vertex in B. Then*

$$N(x, y) = \sum_{b \in B} N(x, b)N(b, y)$$

More generally, suppose A is the set of sources and C is the set of sinks of some DAG, and that some set B has the property such that for every path from A to C passes exactly once through B. Define the gransfer matrix $N(A, B)$ as the matrix whose (a, b)th entry is $N(a, b)$ and define $N(B, C)$ and $N(A, C)$ respectively. Then

$$N(A, C) = N(A, B)N(B, C).$$

Proof. The first result follows from the divide and conquer principle and the principle of multiplication. The second claim directly follows from the first essentially by the definition of matrix multiplication. ■

Theorem 12. *Suppose the vertex set of the DAG G is of the form $V(G) = V_0 \cup V_1 \cup \cdots \cup V_m$ where V_i and V_j are disjoint for $i \neq j$ and every edge goes from V_i to V_{i+1} for some i. Suppose also that each path from V_0 to V_m passes through exactly one vertex in each of the sets V_j, and that the pattern of connections from V_i to V_{i+1} is independent of i. That is, there exists a graph isomorphism $f : V \to (V_1 \cup V_2 \cup \cdots)$, such that*

$$V_0 \mapsto V_1 \mapsto V_2 \mapsto \cdots .$$

Then the transfer matrix from V_i to V_{i+1} is independent of i, and the transfer matrix from V_0 to V_m is equal to the mth power of this matrix.

Theorem 13. *Let V be a semi-infinite (ie, extending infinitely rightwards but not leftwards) with the symmetry property of the last theorem. Fix a vertex x in V_0 and y_0, y_1, y_2 where $y_j \in V_j$, respectively, which are all symmetrical vertices. The $N(x, y_0), N(x, y_1), \ldots$ is a sequence that satisfies a dth order linear recurrence equation where $d = |V_i|$.*

To prove this we will use the Cayley Hamilton

Theorem 14. *Let M be a matrix with characteristic polynomial $p(x) = \det(xI - M)$. Then $p(M) = 0$.*

We will not prove this theorem, it is a fundamental theorem of linear algebra. A nice computation to do is to directly calculate it in the case when M is a 2×2 matrix. Here is a corollary which can be thought of as a combinatorial interpretation of the Cayley Hamilton Theorem.

Corollary 1. *Write $p(t) = t^m + a_{m-1}t^{m-1} + \cdots + a_0$ where $p(t)$ is as above. Then the sequence of matrices $I, M, M^2, M^3, \ldots$ satisfy a linear recurrence relation*

$$M^{n+m} + a_{m-1}M^{n+m-1} + \cdots + a_0 M^n = 0$$

for all $n \geq 0$.

It can clearly be seen that each entry of the matrix as a component also satisfies this recurrence relation. Now we are ready to prove the theorems above:

Proof. Let M be the transfer matrix of connections between V_0 and V_1 (M is also the transfer matrix from V_i to V_{i+1}). Then $N(x, y_k)$ is the (x, y)th entry of M^k (by definition of the transfer matrix). Applying the corollary above gives us the desired results. ■

8 Change of Basis and the Stirling Numbers

In these notes we have introduced (one implicitly, one explicitly) two different bases for the vector space of polynomials $P[x]$. First we have the standard monomial basis

$$1, x, x^2, x^3, \dots$$

and then we introduced the falling factorial basis

$$1, (x), (x)_2, (x)_3, \dots$$

where

$$(x)_k = x(x-1)(x-2)\cdots(x-k+1).$$

It is clear that both sets of polynomials form bases. Moreover, due to the nature of the monomial basis, the change of basis "matrix" (in the sense of an infinite matrix) is upper triangular (so working with it we have no analytic difficulties). It makes sense to consider the coefficients of a change of basis formula. We first define $s(n,k)$ to be the coefficients such that

$$(x)_n = \sum_{k=0}^{\infty} s(n,k)x^k$$

is true. Note that $s(n,k) = 0$ if $k > n$. Similarly, we can write

$$x^n = \sum_{k=0}^{\infty} S(n,k)(x)_k.$$

Note as well that $S(n,k) = 0$ if $k > n$. If we substitute the first equation into the second we get

$$\begin{aligned} x^n &= \sum_{k=0}^{\infty} S(n,k)(x)_k \\ &= \sum_{k=0}^{\infty} S(n,k) \sum_{j=0}^{\infty} s(k,j)x^j \\ &= \sum_{j=0}^{\infty} \left[\sum_{k=0}^{\infty} S(n,k)s(k,j) \right] x^j. \end{aligned}$$

By linear independence it follows that

$$\sum_{k=0}^{\infty} S(n,k)s(k,j) = \begin{cases} 1 & n = j \\ 0 & \text{otherwise} \end{cases}.$$

This is perhaps not so surprising on its own: This actually always happens when you consider any change of basis from one basis back to itself. What will be surprising is that s and S have combinatorial meaning. For example, when we derived a formula for the Bell Numbers, we realized $S(n,k)$ as the number of ways of partitioning $\{1, \dots, n\}$ into exactly k disjoint subsets. We will understand the combinatorial meaning of these formulas soon enough.