STERN NOTES, CHAPTER 2 (FIRST DRAFT)

BRUCE REZNICK, UIUC

Number theorists and combinatorists like to study sequences

$$a = (a_0, a_1, \dots) \in \mathbb{C}^N. \tag{1}$$

One standard technique is to associate with a its *generating function*

$$f := \sum_{n=0}^{\infty} a_n X^n. \tag{2}$$

We use the capital letter to emphasize that X is more a place-holder than a variable. We do not care about the convergence in making this definition. (If the series *does* have a positive radius of convergence, then it is desirable to treat it as an analytic function, and write $f(X)$.) Technically speaking a generating function is a *formal power series.* If R is a ring, then $R[[X]]$ is defined to be the *ring of formal power series in* R. Typically, $R = \mathbb{C}$, although it is not uncommon for all objects to live in $\mathbb{Z}[[X]]$. It is sensible to take $(\mathbb{Z}/d\mathbb{Z})[[X]]$, when one is interested in the a_k's mod d. If $a_n = 0$ for $n > N$, then f is a polynomial, and we will treat it as such.

The operations in $R[[X]]$ are the familiar natural ones; we act as if the elements are ordinary convergent power series, so

$$\begin{gathered} f = \sum_{n=0}^{\infty} a_n X^n, \quad g = \sum_{n=0}^{\infty} b_n X^n \implies f + g = \sum_{n=0}^{\infty} (a_n + b_n) X^n, \\ fg = \sum_{n=0}^{\infty} c_n X^n, \quad \text{where} \quad c_n = \sum_{k=0}^{n} a_k b_{n-k}. \end{gathered} \tag{3}$$

It should be noted that $\mathbb{C}[[X]]$ is also a vector space over $\mathbb{C}$, and this is extremely useful in discussing recurrences. Suppose $\lambda_1, \dots, \lambda_k$ are fixed and

$$A_\lambda = \{ f = \sum_{n=0}^{\infty} a_n X^n \ : \ a_n = \lambda_1 a_{n-1} + \cdots + \lambda_k a_{n-k}, \quad n \ge k \}. \tag{4}$$

That is, A_λ is the set of all generating functions of sequences satisfying a particular linear recurrence. Then it is easy to see that A_λ is a k-dimensional vector space, and if we can find k linearly independent elements in it, we will have gone a long way towards solving the recurrence.

One appeal of generating functions is that natural operations on the sequence are often easily expressed in the generating function. For example;

$$X^k \cdot \sum_{n=0}^{\infty} a_n X^n = \sum_{n=k}^{\infty} a_{n-k} X^n,$$

$$\sum_{n=0}^{\infty} X^n \cdot \sum_{n=0}^{\infty} a_n X^n = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} a_k \right) X^n,$$

$$\sum_{n=0}^{\infty} X^{tn} \cdot \sum_{n=0}^{\infty} a_n X^n = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{\lfloor n/t \rfloor} a_{n-kt} \right) X^n, \tag{5}$$

$$(1-X) \cdot \sum_{n=0}^{\infty} a_n X^n = a_0 + \sum_{n=1}^{\infty} (a_n - a_{n-1}) X^n;$$

$$(1 - \lambda_1 X - \cdots \lambda_k X^k) \cdot \sum_{n=0}^{\infty} a_n X^n = \cdots + \sum_{n=k}^{\infty} (a_n - \lambda_1 a_{n-1} - \cdots - \lambda_k a_{n-k}) X^n.$$

The *degree* of f, *deg*(f), is the smallest index n for which $a_n \neq 0$. Put another way, *deg*$(f) \geq n \iff f = x^n g$ for some $g \in R[[x]]$, or $f \in x^n R[[x]]$. It is customary to say that the 0 element has deg $= \infty$. It is routine to verify that, if $\deg(f - f'), \deg(g - g') \geq n$, then $\deg((f+g) - (f'+g')) \geq n$ and $\deg(fg - f'g') \geq n$. If a_0 is invertible, then $f = \sum_{n=0}^{\infty} a_n X^n$ is also invertible. In fact, letting $g = \sum_{n=0}^{\infty} b_n X^n$, take $b_0 = a_0^{-1}$ and, recursively, $b_n = -a_0^{-1} \sum_{k=1}^{n} a_k b_{n-k}$, to ensure that $fg = 1$.

We impose a curiously strict topology on $R[[X]]$, based on the idea that R itself might not have many open sets – think $R = \mathbb{Z}$. Let (f_r) be a sequence of formal power series, then we say that f_r *converges* to f if *deg*$(f_r - f) \to \infty$. That is, for every n there exists M so that if $r \geq M$ and $j \leq n$, then the coefficient of x^j in f_r equals the coefficient of x^j in f. This is not the standard definition of convergence; for example, it is always true that

$$\lim_{N \to \infty} \sum_{n=0}^{N} a_n X^n = \sum_{n=0}^{\infty} a_n X^n. \tag{6}$$

Under this definition, if $\deg h \geq 1$, then $\sum_{n=0}^{\infty} h^n$ will converge, and its sum will be, as we might hope, $(1-h)^{-1}$. Indeed,

$$\sum_{n=0}^{N} h^n = \frac{1 - h^{N+1}}{1-h} = \frac{1}{1-h} + \Phi_N, \tag{7}$$

where $\deg \Phi_N = (N+1)(\deg h) \to \infty$. Another odd, and useful, fact is that if we let

$$f_N = \sum_{n=0}^{\infty} a_n X^{nN}, \tag{8}$$

then f_N converges to a_0. In the fortunate circumstance that f has a positive radius of convergence, it is sensible to write $f(X^n) \to a_0$.

As an example of the sort of question to which one might want a generating function which does not converge, consider this: Given $n \in \mathbb{N}$, compute

$$e_n := \sum_{j_1+2j_2+\cdots+nj_n=n} \frac{(j_1+\cdots+j_n)!}{j_1!\cdots j_n!} 1!^{j_1}\cdots n!^{j_n}. \tag{9}$$

This looks awful, but it has a reasonable interpretation. The sum, taken first over those j's with $j_1+\cdots+j_n=k$, is just the coefficient of x^n in

$$(1!x^1+2!x^2+\cdots+n!x^n+\cdots)^k, \tag{10}$$

and, after summing on k, we see that

$$1+\sum_{n=1}^{\infty} e_n X^n = \frac{1}{1-\sum\limits_{n=1}^{\infty} n!X^n}. \tag{11}$$

We are particularly interested in infinite products of a particular kind. Suppose $\deg(g_n) \to \infty$ and define

$$\prod_{n=1}^{\infty}(1+g_n) := \lim_{N\to\infty}\prod_{n=1}^{N}(1+g_n). \tag{12}$$

It is routine to verify that the coefficient of x^j stabilizes once $\deg(g_n) > j$. The most vital infinite product in number theory is quite simple:

$$\prod_{n=0}^{\infty}(1+X^{2^n}) = \frac{1}{1-X}. \tag{13}$$

The proof of this formula is simply a telescoping product

$$\prod_{n=0}^{N}(1+X^{2^n}) = \prod_{n=0}^{N}\frac{1-X^{2^{n+1}}}{1-X^{2^n}} = \frac{1-X^{2^{N+1}}}{1-X} = \frac{1}{1-X}+\Phi_N, \tag{14}$$

where $\deg \Phi_N = 2^{N+1} \to \infty$. It is important to note that this definition of infinite product is different from the standard definition in complex variables. For example

$$\prod_{n=1}^{\infty}(1+n^nX^n) \tag{15}$$

is a perfectly well-defined infinite product, even though, as a power series, it would converge only for $X=0$. On the other hand, Euler's famous infinite product,

$$\frac{sin(\pi X)}{\pi X} = \prod_{n=1}^{\infty}\left(1-\frac{X^2}{n^2}\right), \tag{16}$$

is *not* convergent under this definition.

Suppose $A = \{a_0 < a_1 \cdots < a_m\}$ is a finite subset of $\mathbb{N}$. We define the characteristic generating function I_A by

$$I_A = \sum_{a \in A} X^a = \sum_{j=0}^{m} X^{a_j}. \tag{17}$$

If A and B are two such finite subsets, then since these are finite sums,

$$I_A I_B = \sum_{j=0}^{m} X^{a_j} \sum_{k=0}^{\ell} X^{b_k} = \sum_{j=0}^{m} \sum_{k=0}^{\ell} X^{a_j + b_k}, \tag{18}$$

and we see that the coefficient of X^n in $I_A I_B$ is simply the number of ways to write $n = a_j + b_k$. This clearly generalizes to finite numbers of finite subsets of $\mathbb{N}$, and we would like to consider (possibly) infinite collections of (possibly) infinite subsets of $\mathbb{N}$.

Theorem 1. *Suppose*

$$A_k = \{0 = a_{k0} < a_{k1} < \cdots\} \subseteq \mathbb{N},$$

either for $k = 1, \cdots M$*, or for* $k \in \mathbb{N}$*, under the condition that* $\lim_{k \to \infty} a_{k1} = \infty$*. Then*

$$\prod_k I_{A_k} = \sum_{n=0}^{\infty} c_n X^n, \tag{19}$$

where c_n *is the number of ways to write*

$$n = a_{1,r_1} + a_{2,r_2} + \cdots . \tag{20}$$

Proof. It suffices to prove that each particular n, c_n has this interpretation. Let us define

$$A_k^{(n)} = A_k \cap \{0, 1, \ldots, n\}.$$

Clearly, in any representation of n as a sum of a_{j,r_j}'s, we will have $a_{j,r_j} \in A_j^{(n)}$, so the number of representations of n is not affected by this restriction. In the same way, replacing $\prod_k I_{A_k}$ with $\prod_k I_{A_k^{(n)}}$ will only change terms with exponents larger than n. Furthermore, even if we started with an infinite set of A_k's, since $a_{k1} \to \infty$, only finitely many factors $I_{A_k^{(n)}}$ will not be equal to 1. Thus, in any case, it suffices to consider the coefficient of X^n in the *finite* product

$$\prod_k I_{A_k^{(n)}},$$

and this is what we have already done. □

In the most notorious example, let $A_k = \{0, 2^k\}$, $k \geq 0$ and since

$$\prod_{k=0}^{\infty} I_k = \prod_{k=0}^{\infty} (1 + X^{2^k}) = \frac{1}{1 - X} = \sum_{n=0}^{\infty} X^n, \tag{21}$$

we recover the economically useful fact that every integer n has a unique representation of the form

$$n = \sum_{k=0}^{\infty} \epsilon_k(n) 2^k, \quad \epsilon_k(n) \in \{0, 1\}.$$

Now let

$$b(n) := \sum_{k=0}^{\infty} \epsilon_k(n)$$

denote the sum of the binary digits of n, and consider the infinite product

$$\Theta(X, Y) = \prod_{k=0}^{\infty} (1 + Y \cdot X^{2^k}). \tag{22}$$

Interlude

It is not hard to see that $(R[[X]])[[Y]]$, $(R[[Y]])[[X]]$ and $R[[X, Y]]$ all represent the set of all objects; namely, formal series in two variables that look like

$$\sum_{j=0}^{\infty} \sum_{k=0}^{\infty} a_{jk} X^j Y^k, \tag{23}$$

with $a_{jk} \in R$. However, the topology in the three cases is different. Note that

$$\sum_{n=1}^{\infty} (X^n + Y^n) \tag{24}$$

does not converge in either $(R[[X]])[[Y]]$ or $(R[[Y]])[[X]]$. For example, in $(R[[X]])[[Y]]$, X^n is in the base-ring, so it has degree 0, and so the degree of the summands is always 0, and doesn't go to ∞. This is silly, because the sum *obviously* converges, and as an element in $R[[X, Y]]$, the monomial has degree n, which does go to ∞. To make convergence work in $R[[X, Y]]$, you have to take the total degree of the monomials and show that, for all N, there are only finitely many monomials with degree $< N$. In the example of interest, the product for Θ converges as an element in $(R[[Y]])[[X]]$ (the factors are 1 plus a term of degree 2^k), but the product does not converge as an element in $(R[[X]])[[Y]]$, since each factor is only linear in Y. This is the sort of mathematics that led me as a grad student to have evil, and incorrect thoughts about the value and beauty of algebra. In this class, we'll say that Θ converges.

Back to our story

In fact, it's easy to see that

$$\Theta(X, Y) = \sum_{n=0}^{\infty} Y^{b(n)} X^n. \tag{25}$$

If you want to take it the other way around, then

$$\Theta(X,Y) = \sum_{m=0}^{\infty} a_m(X) Y^m, \quad \text{where} \quad a_m(X) = \sum_{0 \le i_1 < i_2 < \cdots < i_m} X^{2^{i_1}} + \cdots + X^{2^{i_m}}. \tag{26}$$

If $A = \{1 \le a_0 < a_1 < \cdots\} \subseteq \mathbb{Z}$, then a *partition* of n from A is a sum $n = a_{i_0} + a_{i_1} + \cdots$ in which $i_0 \le i_1 \le \cdots$.. One counts the number of times a given element a_j appears in a summation and so we are writing n as a sum taken from the sets $a_j\mathbb{N}$, and in the terminology of the theorem, the generating function for the number of partitions of n from A is simply

$$\prod_{j \ge 0} (1 + X^{a_j} + X^{2a_j} + \dots) = \prod_{j \ge 0} \frac{1}{1 - X^{a_j}}. \tag{27}$$

A partition of n into *distinct parts* is one in which each a_j occurs at most once, so $A_j = \{0, .a_j\}$. The generating function for the number of partitions of n into distinct parts from A is

$$\prod_{j \ge 0} (1 + X^{a_j}). \tag{28}$$

One of the most beautiful classical theorems in partition theory is that the number of partitions of n into odd parts (i.e., $A = 2\mathbb{N} + 1$) is equal to the number of partitions of n into distinct parts from $\mathbb{N}$. One proof of this is that the two generating functions are equal (there are other proofs). Indeed,

$$\prod_{k=1}^{\infty} (1 + X^k) = \prod_{k=1}^{\infty} \frac{1 - X^{2k}}{1 - X^k} = \frac{\prod_{k=1}^{\infty} (1 - X^{2k})}{\prod_{k=1}^{\infty} (1 - X^k)} = \prod_{j=0}^{\infty} \frac{1}{1 - X^{2j+1}} \tag{29}$$

since the terms with even exponents in the numerator cancel out in the denominator, leaving the terms with odd exponents. This proof *is* rigorous, since you can equate the coefficients of the partial products of both sides, up to ever-increasing degree.

Remeber the Stern sequence? Let

$$\mathcal{S}(X) = \sum_{n=0}^{\infty} s(n) X^n = X\mathcal{T}(X). \tag{30}$$

(We can define $\mathcal{T}(X)$ in this way because $s(0) = 0$.) We have already shown that $1 \le s(n) \le n$, hence $\lim(s(n))^{1/n} = 1$ and so $\mathcal{S}(x)$ has radius of convergence 1, and is analytic on the open unit disk. It is therefore reasonable to talk about $\mathcal{S}(X^r)$, as we shall below.

By breaking up the sum into even and odd indices, we obtain an interesting equation satisfied by $\mathcal{S}$.

$$\begin{aligned}\mathcal{S}(X) &= \sum_{n=0}^{\infty} s(2n)X^{2n} + \sum_{n=0}^{\infty} s(2n+1)X^{2n+1} \\ &= \sum_{n=0}^{\infty} s(n)X^{2n} + \sum_{n=0}^{\infty} s(n)X^{2n+1} + \sum_{n=0}^{\infty} s(n+1)X^{2n+1}\end{aligned} \tag{31}$$

By reindexing the third sum, we see that

$$\begin{aligned}\mathcal{S}(X) &= (1+X+X^{-1})\mathcal{S}(x^2) \\ \implies X\mathcal{T}(X) &= (1+X+X^{-1})X^2\mathcal{T}(X^2) \\ \implies \mathcal{T}(X) &= (1+X+X^2)\mathcal{T}(X^2).\end{aligned} \tag{32}$$

Now feed this last equation into itself repeatedly, to get

$$\mathcal{T}(X) = \left(\prod_{j=0}^{N-1}(1+X^{2^j}+X^{2^{j+1}})\right)\cdot \mathcal{T}(X^{2^N}), \tag{33}$$

and since $\mathcal{T}(X^{2^N}) \to 1$, we conclude that, as an analytic function as well as a formal power series,

$$\mathcal{S}(X) = X\mathcal{T}(X) = X\prod_{j=0}^{\infty}(1+X^{2^j}+X^{2^{j+1}}). \tag{34}$$

The coefficient of X^n in $\mathcal{T}(X)$ is $s(n-1)$ and $\mathcal{T}(X)$ is the generating function of sums from the sets $\{0, 2^j, 2\cdot 2^j\}$. Thus, we obtain another proof of the last theorem in the first section.

We close this section with a few applications of generating function. We suspect there we have missed some interesting ones. Using the previous definition, let

$$\Theta(X,-1) = \prod_{k=0}^{\infty}(1-X^{2^k}) = \sum_{n=0}^{\infty}(-1)^{b(n)}X^n. \tag{35}$$

Observe that

$$\mathcal{S}(X) = X\prod_{j=0}^{\infty}(1+X^{2^j}+X^{2^{j+1}}) = X\prod_{j=0}^{\infty}\frac{1-X^{3\cdot 2^j}}{1-X^{\cdot 2^j}} = X\cdot\frac{\Theta(X^3,-1)}{\Theta(X,-1)}. \tag{36}$$

Upon cross-multiplying, we find that

$$\Theta(X,-1)\mathcal{S}(X) = X\Theta(X^3,-1), \tag{37}$$

and by taking the coefficient of X^{3k+r} on both sides, we find peculiar recurrences:

$$\sum_{j=0}^{3k+r} s(j)b(3k+r-j) = 0, (r=0,2); \qquad \sum_{j=0}^{3k+1} s(j)b(3k+1-j) = b(n). \tag{38}$$

It is also true that

$$\frac{1}{\Theta(X,-1)} = \prod_{k=0}^{\infty} \frac{1}{1-X^{2^k}} = \sum_{n=0}^{\infty} b(n,\infty)X^n \tag{39}$$

gives the generating function for the so-called *binary partition function*, which counts the number of partitions from the set of powers of 2. (These were studied by Churchhouse and others in the 60's.) From this point of view, we have

$$\mathcal{S}(X) = X \cdot \frac{\Theta(X^3,-1)}{\Theta(X,-1)} = \frac{\sum\limits_{n=0}^{\infty} b(n,\infty)X^n}{\sum\limits_{n=0}^{\infty} b(n,\infty)X^{3n}}, \tag{40}$$

and so we get another recurrence:

$$\sum_{j=0}^{n} s(n-3j)b(j,\infty) = b(n-1,\infty). \tag{41}$$

We may look at the generating function $\mathcal{S}(X)$ over $(\mathbb{Z})/(2\mathbb{Z})$:

$$\mathcal{S}(X) = X\prod_{j=0}^{\infty}(1+X^{3\cdot 2^j})\prod_{j=0}^{\infty}\frac{1}{1+X^{\cdot 2^j}} = X\cdot\frac{1}{1-X^3}\cdot(1-X) = \frac{X+X^2}{1-X^3}, \tag{42}$$

from which it is easy to see that $s(n)$ is even only when $3 \mid n$. We haven't found any *useful* versions mod d for $d \geq 3$; however, $1+X+X^2 \equiv (1-X)^2$ mod 3, so

$$\mathcal{S}(X) \equiv X\Theta(X,-1)^2 \pmod{3}. \tag{43}$$

One final application is an observation due to Richard Stanley, which appeared on the second floor of Illini Hall in the 1980's. Let $\epsilon = e^{\pi i/3} = \frac{1}{2} + \frac{i\sqrt{3}}{2}$; ϵ is a primitive 6-th root of unity, and $(1+\epsilon x)(1+\epsilon^{-1}x) = 1+x+x^2$. It follows that

$$\begin{gathered} \mathcal{S}(X) = X\prod_{j=0}^{\infty}(1+\epsilon X^{2^j})\prod_{j=0}^{\infty}(1+\epsilon^{-1}X^{2^j}) = X\Theta(X,\epsilon)\Theta(X,\epsilon^{-1}) \\ \implies s(n) = \sum_{k=0}^{n-1}\epsilon^{b(k)-b(n-1-k)}, \end{gathered} \tag{44}$$

The sum above is not *a priori* positive, and suggests an unexpected pattern in $b(n)$ (mod 6).

Notes I, p.6, (32): the formula at the very end is wrong! Please make this change:

$$b + \frac{1}{a+\frac{s(n')}{s(n'+1)}} \to b + \frac{1}{a+\frac{1}{\frac{s(n')}{s(n'+1)}}}.$$

Notes I (supp), p.3, l.7-: "helpfulin" → "helpful in".

Notes II (for the rest of this list), p. 1, l.6-: insert "$\lambda = (\lambda_1,\ldots,\lambda_k)$".