STERN NOTES, MATH 595, SPRING 2012

BRUCE REZNICK, UIUC

2. Generating functions

2.1. **Definitions.** One of the ways number theorists and combinatorists study a numerical sequence $a = (a_0, a_1, \dots)$ is to associate it with a *generating function*

$$f_a := f = \sum_{n=0}^{\infty} a_n X^n. \tag{2.1}$$

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 also desirable to treat it as an analytic function, and write $f(z)$.) Technically speaking, a generating function is a *formal power series.* The next few pages contain some of the necessary theoretical background for formal power series. Three excellent books which cover this topic and much, much more are: *Concrete Mathematics* by Graham, Knuth and Patashnik, the two volumes of *Enumerative Combinatorics* by Stanley, and *Generatingfunctionology* by Wilf, which is also available on-line.

We assume R is an integral domain, a commutative ring with identity 1_R and no zero divisors. The examples here will usually be $\mathbb{C}, \mathbb{Z}, \mathbb{Z}/p\mathbb{Z}$ for prime p. Let $R[[X]]$ denote the *ring of formal power series in* R.

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

$$f = \sum_{n=0}^{\infty} a_n X^n, \quad g = \sum_{n=0}^{\infty} b_n X^n \implies f + g = \sum_{n=0}^{\infty} (a_n + b_n) X^n,$$

$$fg = \sum_{n=0}^{\infty} c_n X^n, \quad \text{where} \quad c_n = \sum_{k=0}^{n} a_k b_{n-k}, \tag{2.2}$$

$$r \in R \implies rf = \sum_{n=0}^{\infty} r a_n X^n.$$

It is routine, and not very interesting, to prove that $R[[X]]$ is also an integral domain, with identity element $1_{R[[X]]} := 1 + \sum_{n=1}^{\infty} 0 \cdot X^n$, and we'll skip this, although the following result is useful.

Theorem 2.1. *If $f = \sum_{n=0}^{\infty} a_n X^n \in R[[X]]$ and a_0 is invertible in R, then f is invertible in $R[[X]]$.*

Sketch of proof. Taking f, g as above, we see that $fg = 1_{R[[X]]}$ if and only if this infinite system of equations is valid:

$$a_0 b_0 = 1, \quad a_0 b_1 + a_1 b_0 = 0, \quad a_0 b_2 + a_1 b_1 + a_2 b_0 = 0, \dots$$

If we define $b_0 = a_0^{-1}$, $b_1 = -a_0^{-1}(a_1 b_0)$, $b_2 = -a_0^{-1}(a_1 b_1 + a_2 b_0)$, etc, it's easy to see that the b_n's can be defined recursively. □

If f is invertible and $fg = h$, we write $g = f^{-1}h$ and $g = h/f$ interchangeably. One application is that if $f \in \mathbb{C}[[X]]$ with integer coefficients, so $f \in \mathbb{Z}[[X]]$ as well, and $a_0 = 1$, then $f^{-1} \in \mathbb{Z}[[X]]$.

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

$$\begin{aligned} X^k \cdot \sum_{n=0}^{\infty} a_n X^n &= \sum_{n=k}^{\infty} a_{n-k} X^n, \\ \sum_{n=0}^{\infty} X^n \cdot \sum_{n=0}^{\infty} a_n X^n &= \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} a_k \right) X^n, \\ \sum_{n=0}^{\infty} X^{tn} \cdot \sum_{n=0}^{\infty} a_n X^n &= \sum_{n=0}^{\infty} \left(\sum_{k=0}^{\lfloor n/t \rfloor} a_{n-kt} \right) X^n, \\ (1-X) \cdot \sum_{n=0}^{\infty} a_n X^n &= a_0 + \sum_{n=1}^{\infty} (a_n - a_{n-1}) X^n. \end{aligned} \tag{2.3}$$

The last equation above generalizes in an interesting way:

$$\begin{gathered} (1 - \lambda_1 X - \cdots - \lambda_d X^d) \cdot \sum_{n=0}^{\infty} a_n X^n = \\ \sum_{k=0}^{d-1} (a_k - \lambda_1 a_{k-1} - \cdots - \lambda_k a_0) X^k + \sum_{n=d}^{\infty} (a_n - \lambda_1 a_{n-1} - \cdots - \lambda_d a_{n-d}) X^n. \end{gathered} \tag{2.4}$$

Since $\mathbb{C}[[X]]$ is a vector space over $\mathbb{C}$, suppose $\lambda := (\lambda_1, \dots, \lambda_d) \in \mathbb{C}^d$ is fixed and let

$$\begin{aligned} A_\lambda &= \left\{ f = \sum_{n=0}^{\infty} a_n X^n \in \mathbb{C}[[X]] : a_n = \lambda_1 a_{n-1} + \cdots + \lambda_d a_{n-d}, \quad n \ge d. \right\} \\ &= \left\{ f : (1 - \lambda_1 X - \cdots \lambda_d X^d) f = \sum_{k=0}^{d-1} b_k X^k \right\}. \end{aligned} \tag{2.5}$$

That is, A_λ is the set of generating functions of sequences satisfying a given linear recurrence. Then (2.5) implies that A_λ is a d-dimensional subspace of $\mathbb{C}[[X]]$.

The *order* of a non-zero element $f \in R[[X]]$, $ord(f)$, is the smallest index n for which $a_n \neq 0$; in this case, we say that f has *leading term* a_nX^n, with *leading coefficient* a_n. It is customary to say that $ord(0_{R[[X]]}) = \infty$; don't tell the undergrads!

Put another way, $ord(f) \geq n$ if and only if $f = X^n g$ for some $g \in R[[X]]$; if f also defines an analytic function, then $ord(f)$ is the order of $z = 0$ as a zero of f. If f happens to be a polynomial (formally, if $a_n = 0$ for $n > d$), the order of f is the *smallest* degree of a non-zero monomial in f, not the *largest*.

More generally, $ord(fg) = ord(f) + ord(g)$, (so $ord(f^k) = k * ord(f)$); however, addition is trickier. If $ord(f) \neq ord(g)$, then $ord(f + g) = \min(ord(f), ord(g))$; if $ord(f) = ord(g) = m$, say, then $ord(f + g) \geq m$, with inequality occurring if the leading terms of f and g cancel. Since $ord(1_{R[[X]]}) = 0$, if f is invertible, then $ord(f) = ord(f^{-1}) = 0$. If $f - f'$ and $g - g'$ both have order $\geq n$, then so does $fg - f'g'$ (write $f = f' + h$ and $g = g' + k$ and multiply out.)

We impose the following topology on $R[[X]]$, based on the premise that we should assume nothing about the topology of R. For each $n \geq 1$, the open ball of radius $\frac{1}{n}$ centered at f consists of f, together with the set of g so that $ord(f - g) \geq n$. That is, the elements of this open ball are those g with the property that the first n terms of f and g agree. According to this topology, if $f_r \in R[[X]]$, then "$f_r \to f$" means precisely that for every $n \in \mathbb{N}$ there exists M_n so that if $r \geq M_n$ and $j \leq n$, then the coefficients of x^j are the same in f_r and f; that is, the coefficients *stabilize.* It is routine to verify that if $g \in R[[x]]$ and $f_r \to f$, then $gf_r \to gf$.

This is *not* the usual power series convergence. For example, every formal power series converges! That is, it's *always* true that

$$\lim_{N\to\infty} \sum_{n=0}^{N} a_nX^n = \sum_{n=0}^{\infty} a_nX^n. \tag{2.6}$$

However, if $R = \mathbb{C}$ and $f \neq 0$, then $(1 + \frac{1}{n})f$ *never* converges to f. Also, if

$$f_N = \sum_{n=0}^{\infty} a_nX^{nN},$$

then $f_N \to a_0$ as $N \to \infty$, (For analytic functions, $f_N(z) = f(z^N)$, and if f is analytic in a neighborhood of zero and $|z| < 1$, then it is true that $\lim_N f(z^N) = a_0$.)

Here is a proof that the geometric series converges in $R[[X]]$ to $(1-X)^{-1}$ according to this definition of convergence. (Since X is not assumed to take a value, there is no "circle of convergence".) Let

$$f = \sum_{n=0}^{\infty} X^n, \qquad f_N = \sum_{n=0}^{N} X^n. \tag{2.7}$$

Then $(1 - X)f_N = 1 - X^{N+1}$, and since $f_N \to f$ and $(1 - X)f_n \to 1_{R[[X]]}$, it follows that $(1 - X)f = 1_{R[[X]]}$; that is, $f = (1 - X)^{-1}$. It is routine to verify that if

$ord(h) \geq 1$, then

$$\sum_{n=0}^{\infty} h^n = (1-h)^{-1}. \tag{2.8}$$

More generally, if $h \in R[[X]]$ and $ord(h) \geq 1$, then functional composition can be unambiguously defined:

$$f = \sum_{i=0}^{\infty} a_i X^i \implies f \circ h = \sum_{i=0}^{\infty} a_i h^i \tag{2.9}$$

We violate our usual squeamishness about functional dependence when $h = X^t$:

$$f(X) = \sum_{n=0}^{\infty} a_n X^n \implies f(X^t) = \sum_{n=0}^{\infty} a_n X^{nt}. \tag{2.10}$$

One more pathology. There is nothing wrong with talking about

$$f = 1 - \sum_{n=1}^{\infty} n! X^n \in \mathbb{C}[[X]]$$

as a formal power series, and since its leading coefficient is 1, it is invertible. It would follow then from (2.8) that

$$f^{-1} = 1 + \sum_{n=1}^{\infty} e_n X^n = 1 + \sum_{k=1}^{\infty} \left(\sum_{m=1}^{\infty} m! X^m \right)^k. \tag{2.11}$$

For numerical $X = z$, the series for f only converges for $X = 0$. But each particular e_n, $n \geq 1$, can be calculated as a finite sum from (2.11):

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

where the block of terms satisfying the additional condition that $\sum j_\ell = k$ come from $(\sum_{m=1}^{\infty} m! X^m)^k$. There *are* circumstances in which this sort of sum arises.

2.2. **Infinite products.** We are particularly interested in infinite products. Suppose $ord(g_n) \to \infty$ and define

$$\prod_{n=1}^{\infty} (1+g_n) := \lim_{N\to\infty} \prod_{n=1}^{N} (1+g_n). \tag{2.12}$$

It is routine to verify that the coefficient of x^j in the partial products stabilizes once $ord(g_n) > j$, and so the limit in (2.12) is always defined.

Again, this definition is somewhat different from the analytic case. For example, as a formal power series

$$\prod_{n=1}^{\infty} (1 + n^n X^n)$$

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

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

is *not* convergent as a formal power series under this definition, because the coefficient of X^2 on the right hand side never stabilizes.

The most vital infinite product in number theory is quite simple, either as a formal power series or as a generating function.

$$\prod_{n=0}^{\infty}(1 + X^{2^n}) = (1 - X)^{-1}. \tag{2.13}$$

The proof of this formula uses a telescoping product:

$$\prod_{n=0}^{N}(1 + X^{2^n}) = \prod_{n=0}^{N} \frac{1 - X^{2^{n+1}}}{1 - X^{2^n}} = \frac{1 - X^{2^{N+1}}}{1 - X} = \sum_{n=0}^{2^{N+1}-1} X^n. \tag{2.14}$$

Thus, the partial products are a subsequence of the partial sums of $(1 - X)^{-1}$, and so converge to it. Alternatively,

$$\prod_{n=0}^{N}(1 + X^{2^n}) - (1 - X)^{-1} = -X^{2^{N+1}}(1 - X)^{-1}$$

and $ord(-X^{2^{N+1}}(1 - X)^{-1}) = 2^{N+1} \to \infty$.

One final point on pathologies. We want to define generating functions with two "variables":

$$\sum_{i,j} a_{i,j} X^i Y^j, \quad a_{i,j} \in R. \tag{2.15}$$

Define the order of the term $a_{i,j}X^iY^j$ to be $i + j$ and define convergence in the same way we did before. This gives the formal power series ring $R[[X, Y]]$. It is clear that we can sum for fixed i or for fixed j first and show that $R[[X, Y]] = (R[[Y]])[[X]] = (R[[X]])[[Y]]$; that is, a formal power series in one variable whose coefficients are formal power series in the other variables. Well, technically, no. These formal power series rings are *isomorphic*, but they're not *equal*, and the isomorphism is something awfully close to the identity map:

$$\sum_{i,j} a_{i,j} X^i Y^j \leftrightarrow \sum_{j=0}^{\infty} \left(\sum_{i=0}^{\infty} a_{i,j} X^i \right) Y^j \leftrightarrow \sum_{i=0}^{\infty} \left(\sum_{j=0}^{\infty} a_{i,j} Y^j \right) X^i.$$

(This is the sort of fine distinction that repelled me from algebra in grad school, until I realized that algebraists don't let these distinctions bother them.)

We give a simple expression of the power of generating functions. Observe that

$$\frac{1}{1-X-Y} = \sum_{n=0}^{\infty}(X+Y)^n = \sum_{n=0}^{\infty}\sum_{k=0}^{n}\binom{n}{k}X^kY^{n-k} = \sum_{k=0}^{\infty}\sum_{\ell=0}^{\infty}\frac{(k+\ell)!}{k!\ell!}X^kY^\ell. \tag{2.16}$$

At the same time, algebraic manipulation yields

$$\begin{aligned} \frac{1}{1-X-Y} &= \frac{1}{1-X}\cdot\frac{1}{1-Y(1-X)^{-1}} \\ &= \frac{1}{1-X}\cdot\sum_{m=0}^{\infty}\frac{Y^m}{(1-X)^m} = \sum_{m=0}^{\infty}\frac{Y^m}{(1-X)^{m+1}} \end{aligned} \tag{2.17}$$

On equating the coefficient of Y^m in (2.16) and (2.17), we see that

$$\frac{1}{(1-X)^{m+1}} = \sum_{k=0}^{\infty}\frac{(k+m)!}{k!m!}X^k = \sum_{k=0}^{\infty}\binom{k+m}{m}x^k. \tag{2.18}$$

This familiar and extremely useful expression can be readily derived in many different ways, both combinatorial and analytical, and will show up later in this chapter.

Complications show up when we take infinite products, and to avoid them, we'll visualize the summation as taking place over all terms of fixed order first; that is,

$$\sum_{i,j}a_{i,j}X^iY^j := \sum_{n=0}^{\infty}\left(\sum_{i=0}^{n}a_{i,n-i}X^iY^{n-i}\right). \tag{2.19}$$

A product such as

$$\sum_{i,j}a_{i,j}X^iY^j = \prod_{n=1}^{\infty}(1+X^n+Y^n)$$

only converges if thought of in this way, since upon viewing this as an element in $R([[Y]])[[X]]$, say, as $\prod(1+g_n)$, the order of $g_n = X^n + Y^n$ *as an element of* $R[[Y]]$ is 0, because Y^n is in the base ring, and so the infinite product does not converge according to our definition. (One way to resolve the conflict is to change the definition of convergence in $R([[Y]])[[X]]$, using the topology we've defined for $R[[Y]]$.) In (2.19), $a_{i,j}$ counts the number of partitions of i and j into distinct parts so that no part appears in both partitions.

2.3. **Partitions.** Partition generating functions are based on a simple idea. Suppose $A = \{0 = a_0 < a_1 \cdots < a_m\}$ is a finite subset of $\mathbb{N}$. We define the characteristic generating function I_A by

$$I_A = \sum_{a\in A}X^a = \sum_{j=0}^{m}X^{a_j} = 1 + \sum_{j=1}^{m}X^{a_j}. \tag{2.20}$$

If A and B are two such finite subsets, then I_A and I_B are finite sums; compute

$$\sum_{n=0}^{\infty} c_n X^n = I_A I_B = \sum_{j=0}^{m} X^{a_j} \sum_{k=0}^{\ell} X^{b_k} = \sum_{j=0}^{m}\sum_{k=0}^{\ell} X^{a_j+b_k}. \tag{2.21}$$

It follows from (2.21) that c_n is the number of ways to write $n = a+b, a \in A, b \in B$. (If, say, $a_0 > 0$, consider the set $A' = \{a_i - a_0\}$; the number of representations of $n - a_0$ from A' and B is equal to the number of n from A and B, etc.)

What if A and B are infinite? No problem. Fix n and let $A^{(n)} = A \cap \{0, 1, \ldots, n\}$ and $B^{(n)} = B \cap \{0, 1, \ldots, n\}$. If $n = a + b$ with $a \in A$ and $b \in B$, then $0 \le a, b \le n$, so $a \in A^{(n)}$, $b \in B^{(n)}$ and so c_n is the coefficient of X^n in $I_{A^{(n)}} I_{B^{(n)}}$. On the other hand, the orders of $I_A - I_{A^{(n)}}$ and $I_B - I_{B^{(n)}}$ are both larger than n, hence so is the order of $I_A I_B - I_{A^{(n)}} I_{B^{(n)}}$. Therefore, c_n is the coefficient of X^n in $I_A I_B$ as well.

What if there are r sets, $A_1, \ldots, A_r$? The same logic applies in terms of a finite sets, and the generalization to infinite sets A_k, $1 \le k \le r$, follows in the same way.

What if there are infinitely many sets A_k? Here we need to place a restriction on the smallest non-zero element, because we want c_n to be finite: for each n, there exist only finitely many A_k's which contain n. With this restriction, the computation of c_n becomes a count of representations of n as a sum from a finite number of sets.

To sum up, we have the following theorem.

Theorem 2.2. *Suppose there exist finite or infinite sets $A_k \subseteq \mathbb{N}$,*

$$A_k = \{0 = a_{k,0} < a_{k,1} < \cdots\},$$

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

$$\prod_k I_{A_k} = \sum_{n=0}^{\infty} c_n X^n \in \mathbb{Z}[[X]],$$

where c_n is the number of ways to write

$$n = a_{1,r_1} + a_{2,r_2} + \cdots, \qquad a_{k,r_k} \in A_k.$$

The same argument applies to subsets $A_k \subset \mathbb{N}^d$ containing 0, with $(a_1, \ldots, a_d)$ associated to $X_1^{a_1} \cdots X_d^{a_d}$, given that the minimum order of the non-constant terms is also going to ∞ as k increases. In this case, the generating function is in $\mathbb{Z}[[X_1, \ldots, X_d]]$. We skip the details.

In the most famous application of Theorem 2.2, let $A_k = \{0, 2^k\}$, $k \ge 0$. By (2.13),

$$\prod_{k=0}^{\infty} I_k = \prod_{k=0}^{\infty}(1 + X^{2^k}) = \frac{1}{1-X} = \sum_{n=0}^{\infty} X^n, \tag{2.22}$$

recovering the economically useful fact that every non-negative 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

$$\Psi(X,Y) = \prod_{k=0}^{\infty} (1 + X^{2^k} \cdot Y) = \sum_{i,j} a_{i,j} X^i Y^j. \tag{2.23}$$

(This is a natural example of convergence in $(\mathbb{C}[[Y]])[[X]]$ but not in $(\mathbb{C}[[X]])[[Y]]$.) Think of this as a partition problem from sets $\{(0,0),(2^k,1)\}$; each $X^n Y^m$ occurs exactly once as a sum, when $m = b(n)$. That is,

$$\begin{aligned} \Psi(X,Y) &= \sum_{n=0}^{\infty} Y^{b(n)} X^n \\ &= \sum_{m=0}^{\infty} a_m(X) Y^m, \quad \text{where} \quad a_m(X) = \sum_{0 \le i_1 < i_2 < \cdots < i_m} X^{2^{i_1}} + \cdots + X^{2^{i_m}}. \end{aligned} \tag{2.24}$$

Notice that if we replace Y by a numerical parameter λ, we get a valid expansion formula for a generating function in one variable:

$$\prod_{k=0}^{\infty} (1 + \lambda X^{2^k}) = \sum_{n=0}^{\infty} \lambda^{b(n)} X^n. \tag{2.25}$$

We'll apply this to the Stern sequence.

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 $i_0 \le i_1 \le \cdots$.. Let $p_A(n)$ be the number of such sums. Let m_k count the number of times that a_k appears in a given partition, so that $n = \sum m_k a_k$. We are thus in the situation of Theorem 2.2, with $A_k = a_k \mathbb{N} = \{0, a_k, 2a_k \ldots\}$. It follows that the generating function for $p_A(n)$ is

$$\prod_{k \ge 0} (1 + X^{a_k} + X^{2a_k} + \ldots) = \prod_{k \ge 0} \frac{1}{1 - X^{a_k}} = \sum_{n=0}^{\infty} p_A(n) X^n. \tag{2.26}$$

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

$$\prod_{k \ge 0} (1 + X^{a_k}) = \sum_{n=0}^{\infty} p_{A,d}(n) X^n. \tag{2.27}$$

One of the most beautiful classical theorems in partition theory goes back to Euler: $p_{2\mathbb{N}+1}(n) = p_{\mathbb{N},d}(n)$. In words, the number of partitions of n into odd parts is equal to the number of partitions of n into distinct parts. The best proof is bijective; the one below, however, uses the ideas of this section and the identity $1 + t = \frac{1-t^2}{1-t}$:

$$\prod_{k=1}^{\infty}(1+X^k) = \prod_{k=1}^{\infty}\frac{1-X^{2k}}{1-X^k} = \frac{\prod_{k=1}^{\infty}(1-X^{2k})}{\prod_{k=1}^{\infty}(1-X^k)} = \prod_{j=0}^{\infty}\frac{1}{1-X^{2j+1}}. \tag{2.28}$$

In the final step of (2.28), the terms with even exponents in the numerator cancel out in the denominator, leaving the terms with odd exponents.

2.4. **Return to Stern.** Remember the Stern sequence? Let

$$\begin{aligned} \mathcal{S}(X) &= \sum_{n=0}^{\infty} s(n)X^n = X\mathcal{T}(X); \\ \mathcal{S}(X) = X + X^2 + 2X^3 + X^4 + \cdots, &\quad \mathcal{T}(X) = 1 + X + 2X^2 + X^3 + \cdots. \end{aligned} \tag{2.29}$$

(We can define $\mathcal{T}(X)$ in this way because $s(0) = 0$ and $ord(\mathcal{S}(X)) = 1$.) We have already shown that $1 \le s(n) \le n$, hence $\lim(s(n))^{1/n} = 1$ and so $\mathcal{S}(z)$ has radius of convergence 1 as an analytic function, and similarly for $\mathcal{T}(z)$.

By breaking up the sum into even and odd indices, and using the recurrence, we obtain a functional equation satisfied by $\mathcal{S}(X)$:

$$\begin{aligned} \mathcal{S}(X) &= \sum_{n=0}^{\infty} s(2n)X^{2n} + \sum_{n=0}^{\infty} s(2n+1)X^{2n+1} \\ &= \sum_{n=0}^{\infty} s(n)X^{2n} + \sum_{n=0}^{\infty} s(n)X^{2n+1} + \sum_{n=0}^{\infty} s(n+1)X^{2n+1} \\ &= \mathcal{S}(X^2) + X\mathcal{S}(X^2) + X^{-1}\mathcal{S}(X^2). \end{aligned} \tag{2.30}$$

The expression $X^{-1}\mathcal{S}(X^2)$ is a legitimate formal power series, because $ord(\mathcal{S}(X^2)) = 2$. Rewrite (2.30) as:

$$\begin{aligned} \mathcal{S}(X) &= (1 + X + X^{-1})\mathcal{S}(X^2) \\ \implies X\mathcal{T}(X) &= (1 + X + X^{-1})X^2\mathcal{T}(X^2). \end{aligned} \tag{2.31}$$

It now follows that

$$\begin{aligned} \mathcal{T}(X) &= (1 + X + X^2)\mathcal{T}(X^2); \\ X\mathcal{S}(X) &= (1 + X + X^2)\mathcal{S}(X^2). \end{aligned} \tag{2.32}$$

The functional equation for $\mathcal{T}$ can be iterated N times to give

$$\mathcal{T}(X) = \left(\prod_{k=0}^{N-1}\left(1 + X^{2^k} + X^{2\cdot 2^k}\right)\right) \cdot \mathcal{T}(X^{2^N}), \tag{2.33}$$

Since $\mathcal{T}(X^{2^N}) = 1 + g_N$, where $ord(g_N) = 2^N$, it follows that $\mathcal{T}(X^{2^N}) \to 1$, and so

$$\mathcal{S}(X) = X\mathcal{T}(X) = X\prod_{k=0}^{\infty}(1 + X^{2^k} + X^{2^{k+1}}). \tag{2.34}$$

The coefficient of X^n in $\mathcal{T}(X)$ is $s(n-1)$ and by Theorem 2.2 and (2.33), $\mathcal{T}(X)$ is the generating function of sums from the sets $\{0, 2^k, 2\cdot 2^k\}$. This provides another proof of Theorem 1.8.

We can now play with the generating function and derive a number of new, and rather unexpected, identities involving $\Psi(X,Y)$, cf. (2.24). Define

$$\mathcal{B}(X) = \Psi(X,-1) = \prod_{j=0}^{\infty}(1 - X^{2^j}) = \sum_{n=0}^{\infty}(-1)^{b(n)}X^n. \tag{2.35}$$

Since $1 + t + t^2 = \frac{1-t^3}{1-t}$, it follows that

$$\mathcal{S}(X) = X\prod_{j=0}^{\infty}(1 + X^{2^j} + X^{2^{j+1}}) = X\prod_{j=0}^{\infty}\frac{1 - X^{3\cdot 2^j}}{1 - X^{2^j}} = X\cdot\frac{\mathcal{B}(X^3)}{\mathcal{B}(X)}. \tag{2.36}$$

Thus,

$$\mathcal{B}(X)\mathcal{S}(X) = X\mathcal{B}(X^3), \tag{2.37}$$

We read off the coefficient of X^{3k+r} on both sides of (2.37), $r = 0, 1, 2$, to obtain some peculiar recurrences:

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

The so-called *binary partition function*, $b(n,\infty)$, has been studied since Euler, with revived interest by Churchhouse and others since the 1960s. Let $A_2 = \{2^k : k \ge 0\}$, and let $b(n,\infty) = p_{A_2}(n)$. Then

$$\mathcal{B}_\infty(X) := \sum_{n=0}^{\infty} b(n,\infty)X^n = \prod_{k=0}^{\infty}\frac{1}{1 - X^{2^k}} = \frac{1}{\mathcal{B}(X)}. \tag{2.38}$$

In Stern terms,

$$\mathcal{S}(X) = X\cdot\frac{\mathcal{B}_\infty(X)}{\mathcal{B}_\infty(X^3)} \implies \mathcal{S}(X)\mathcal{B}_\infty(X^3) = X\cdot\mathcal{B}_\infty(X). \tag{2.39}$$

Again, taking the coefficient of X^n on both sides of (2.39), we get

$$\sum_{j=0}^{\lfloor n/3\rfloor} s(n-3j)b(j,\infty) = b(n-1,\infty). \tag{2.40}$$

The binary partition functions have an interesting alternative interpretation due to Neil Sloane and James Sellers as "non-squashing stacks of boxes". Suppose one has an unlimited supply of boxes labeled with positive integers, so that a box labeled i

both weighs i units and can support a stack of boxes above it of total weight i. How many different "non-squashing" stacks are there of total weight n? In other words, how many partitions are there of n in which each part is at least as large as the sum of the previous parts. Putting this symbolically,

$$n = a_1 + a_2 + \cdots + a_r; \qquad a_1 + \cdots + a_j \le a_{j+1}, \quad 1 \le j \le r-1. \tag{2.41}$$

We first fix the number of parts, r, and reparameterize:

$$\begin{gathered} a_1 = b_1, \quad a_2 = b_1 + b_2, \quad a_3 = a_1 + a_2 + b_3 = 2b_1 + b_2 + b_3, \\ a_4 = b_1 + b_2 + b_3 + b_4 = 4b_1 + 2b_2 + b_3 + b_4, \cdots \end{gathered} \tag{2.42}$$

The conditions of the problem require $b_1 \ge 1$ and $b_i \ge 0$ for $i \ge 2$. A comparison of (2.42) with (2.41) and an omitted inductive argument imply that

$$n = 2^{r-1}b_1 + 2^{r-2}b_2 + \cdots + 2b_{r-1} + b_r.$$

This is a partition of n into powers of 2 with largest part 2^{r-1}. Summing over r shows that the number of partitions of n satisfying (2.41) is equal to $b(n, \infty)$.

Another application of the functional equation (2.32) leads to a rapid proof of Exercise 7 from Chapter 1:

$$\begin{aligned} X\mathcal{S}(X) = (1 + X + X^2)\mathcal{S}(X^2) &\implies X(1-X)\mathcal{S}(X) = (1 - X^3)\mathcal{S}(X^2) \\ &\implies \frac{\mathcal{S}(X)}{1 - X^3} = \frac{1}{X} \cdot \frac{\mathcal{S}(X^2)}{1 - X}. \end{aligned}$$

These expressions can be identified via (2.3):

$$\begin{aligned} \frac{\mathcal{S}(X)}{1-X^3} &= \mathcal{S}(X)(1 + X^3 + X^6 + \cdots) = \sum_{n=0}^{\infty} \left(\sum_{j=0}^{\lfloor n/3 \rfloor} s(n - 3j) \right) X^n; \\ \frac{1}{X} \cdot \frac{\mathcal{S}(X^2)}{1-X} &= \frac{1}{X} \cdot (s(0) + s(1)X^2 + s(2)X^4 + \cdots)(1 + X + X^2 + \cdots) \\ &= \frac{1}{X} \cdot (S(0)(1+X) + S(1)(X^2 + X^3) + S(2)(X^4 + X^5) + \cdots) \\ &= S(1)(X + X^2) + S(2)(X^3 + X^4) + \cdots = \sum_{n=0}^{\infty} S(\lceil \tfrac{n}{2} \rceil) X^n. \end{aligned} \tag{2.43}$$

It follows from (2.43) that

$$\sum_{j=0}^{\lfloor N/3 \rfloor} s(N - 3j) = S(\lceil \tfrac{n}{2} \rceil).$$

The following observation is due to Richard Stanley, from a conversation with the author 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

$$\mathcal{S}(X) = X\prod_{j=0}^{\infty}(1+\epsilon X^{2^j})\prod_{j=0}^{\infty}(1+\epsilon^{-1}X^{2^j}) = X\sum_{i=0}^{\infty}\epsilon^{b(i)}X^i\sum_{j=0}^{\infty}\epsilon^{-b(j)}X^j$$
$$\implies s(n) = \sum_{k=0}^{n-1}\epsilon^{b(k)-b(n-1-k)}. \tag{2.44}$$

Thus,

$$2s(n) = \sum_{k=0}^{n-1}\epsilon^{(b(k)-b(n-1-k))} + \epsilon^{-(b(k)-b(n-1-k))}. \tag{2.45}$$

Now $\epsilon^j + \epsilon^{-j} = 2, 1, -1, -2$ when $j \equiv 0, \pm 1, \pm 2, 3 \pmod 6$, and the sum on the right is not *a priori* positive. This suggests some unexpected patterns in $(b(m)) \bmod 6$.

Replacing $\{0,1,2\}$ with $\{0,1,2,3\}$ gives a much easier problem to analyze. Let $f_4(n)$ denote the number of ways to write n as

$$n = \sum_{i=0}^{\infty}\epsilon_i 2^i, \qquad \epsilon_i \in \{0,1,2,3\}. \tag{2.46}$$

As we have seen,

$$\sum_{n=0}^{\infty} f_4(n)X^n = \prod_{j=0}^{\infty}(1 + X^{2^j} + X^{2\cdot 2^j} + X^{3\cdot 2^j})$$
$$= \prod_{j=0}^{\infty}\frac{1-X^{2^{j+2}}}{1-X^{2^j}} = \frac{\prod_{j=2}^{\infty}(1-X^{2^j})}{\prod_{j=0}^{\infty}(1-X^{2^j})} = \frac{1}{(1-X)(1-X^2)}. \tag{2.47}$$

Here are two ways to look at this sum to get the exact value: $f_4(n) = \lfloor \frac{n}{2} \rfloor + 1$:

$$\frac{1}{(1-X)(1-X^2)} = \frac{1}{1-X}\left(1 + X^2 + X^4 + \cdots\right)$$
$$= 1 + X + 2X^2 + 2X^3 + 3X^4 + \cdots;$$

$$\frac{1}{(1-X)(1-X^2)} = \frac{1}{(1-X)^2(1+X)} = \frac{1/4}{1+X} + \frac{1/4}{1-X} + \frac{1/2}{(1-X)^2}$$
$$= \sum_{n=0}^{\infty}\left(\tfrac{1}{4}(-1)^n + \tfrac{1}{4} + \tfrac{n+1}{2}\right)X^n = \sum_{n=0}^{\infty}\left(\tfrac{n}{2} + \tfrac{3+(-1)^n}{4}\right)X^n$$
$$= \sum_{n=0}^{\infty}\left(\left\lfloor\frac{n}{2}\right\rfloor + 1\right)X^n.$$

A combinatorial explanation for the value of $f_4(n)$ is to write $\epsilon_i = 2\alpha_i + \beta_i$ in (2.46), with α_i, β_i taken independently in $\{0, 1\}$, and then observe that

$$n = \sum_{i=0}^{\infty} \epsilon_i 2^i = \sum_{i=0}^{\infty} (2\alpha_i + \beta_i) 2^i = 2 \sum_{i=0}^{\infty} \alpha_i 2^i + \sum_{i=0}^{\infty} \beta_i 2^i.$$

Thus a representation of n in (2.46) can be bijectively associated with a representation $n = 2n' + n''$ for $n', n'' \geq 0$; there are $\lfloor \frac{n}{2} \rfloor + 1$ possible choices for n'. The computation of $f_4(n)$ was Problem B2 on the 1983 Putnam.

This discussion can be generalized by taking $2^r - 1$ for 3, and we will later show that $b(n, \infty)$ grows more rapidly than any polynomial.

Finally, as another harbinger of a later chapter, we look at the generating function $\mathcal{S}(X)$ over $R = \mathbb{Z}/2\mathbb{Z}$ to get a quick proof that $2 \mid s(n) \iff 3 \mid n$. Keep in mind that $1 = -1$ in R, so $1 + X + X^2 = 1 - X + X^2$ and it's possible to rewrite (2.39) as

$$\begin{aligned} \mathcal{S}(X) &= X \prod_{j=0}^{\infty} (1 + X^{3 \cdot 2^j}) \prod_{j=0}^{\infty} \frac{1}{1 + X^{\cdot 2^j}} \\ &= X \cdot \frac{1}{1 - X^3} \cdot (1 - X) = \frac{X + X^2}{1 - X^3} = (X + X^2)(1 + X^3 + X^6 + \dots). \end{aligned} \tag{2.48}$$

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

$$\mathcal{S}(X) \equiv X \left(\sum_{n=0}^{\infty} (-1)^{b(n)} X^n \right)^2 \pmod 3.$$

2.5. **Some asymptotics.** We want to discuss the behavior of $\mathcal{S}(z)$ and we first need some general asymptotic facts about power series with positive real coefficients. In this section, for an integer $m \geq 0$, let

$$f(z) = \sum_{n=0}^{\infty} a_n z^n, \quad f_m(z) = \sum_{n=m}^{\infty} a_n z^n, \qquad a_n > 0. \tag{2.49}$$

We are particularly interested in statements such as

$$\lim_{x \to 1^-} (1 - x)^\lambda f(x) = c > 0,$$

in which the limit is taken over real $x \to 1$ and real $\lambda > 0$.

Fact 1: If m is a positive integer, then

$$\lim_{x \to 1^-} (1 - x)^\lambda f(x) = c \iff \lim_{x \to 1^-} (1 - x)^\lambda f_m(x) = c. \tag{2.50}$$

The reason is that the omitted m terms are a polynomial which is being multiplied by something going to 0; this *finite* sum does not affect the limit. The same result holds if lim is replaced by lim inf or lim sup and for the same reasons.

Now let

$$g(z) = \sum_{n=0}^{\infty} b_n z^n, \quad g_m(z) = \sum_{n=m}^{\infty} b_n z^n, \qquad b_n > 0. \tag{2.51}$$

Fact 2: If $a_n \le b_n$, then $f(x) \le g(x)$ pointwise, and so $(1-x)^\lambda f(x) \le (1-x)^\lambda g(x)$. Taking the various limits, we find that

$$\begin{aligned} \liminf_{x\to 1^-} (1-x)^\lambda f(x) &\le \liminf_{x\to 1^-} (1-x)^\lambda g(x), \\ \limsup_{x\to 1^-} (1-x)^\lambda f(x) &\le \limsup_{x\to 1^-} (1-x)^\lambda g(x). \end{aligned} \tag{2.52}$$

If one or both of the limits actually exists, then these statements become stronger.

Lemma 2.3. *If* $a_n, b_n > 0$, (2.49) *and* (2.51) *hold and*

$$\lim_{n\to\infty} \frac{b_n}{a_n} = 1, \tag{2.53}$$

then

$$\lim_{x\to 1^-} (1-x)^\lambda f(x) = c \implies \lim_{x\to 1^-} (1-x)^\lambda g(x) = c. \tag{2.54}$$

Proof. Pick $\epsilon > 0$ and assume $\epsilon < 1$. There exists N so that for $n \ge N$, $a_n(1-\epsilon) \le b_n \le a_n(1+\epsilon)$, hence for $x \in (0,1)$,

$$(1-\epsilon)(1-x)^\lambda f_N(x) \le (1-x)^\lambda g_N(x) \le (1+\epsilon)(1-x)^\lambda f_N(x).$$

Taking the limit as $x \to 1^-$, it follows from Fact 2 that

$$c(1-\epsilon) \le \liminf_{x\to 1^-} (1-x)^\lambda g_N(x) \le \limsup_{x\to 1^-} (1-x)^\lambda g_N(x) \le c(1+\epsilon).$$

Thus, by Fact 1,

$$c(1-\epsilon) \le \liminf_{x\to 1^-} (1-x)^\lambda g(x) \le \limsup_{x\to 1^-} (1-x)^\lambda g(x) \le c(1+\epsilon). \tag{2.55}$$

Since $\epsilon > 0$ is arbitrary in (2.55), (2.54) is established. □

We have already seen the power series for $(1-x)^{-m}$ for $m \in \mathbb{N}$ in (2.18), but will need it for other m; Taylor series come to the rescue. Observe that $((1-x)^{-\nu})' = \nu(1-x)^{-(\nu+1)}$, from which it follows that for $\kappa \in \mathbb{R}$,

$$\frac{1}{(1-z)^\kappa} = 1 + \sum_{n=1}^{\infty} \frac{\kappa\cdot(\kappa+1)\cdots(\kappa+(n-1))}{n!} z^n := 1 + \sum_{n=1}^{\infty} A(\kappa;n) z^n. \tag{2.56}$$

We are interested in the growth of the coefficient $A(\kappa;n)$ in n for fixed κ. When $\kappa = m+1 \in \mathbb{N}$, this is clear:

$$\begin{aligned} A(m+1;n) &= \frac{(m+n)!/m!}{n!} = \frac{(n+m)!/n!}{m!} \\ &= \frac{(n+1)(n+2)\cdots(n+m)}{m!} \implies \lim_{n\to\infty} \frac{A(m+1;n)}{n^m/m!} = 1. \end{aligned} \tag{2.57}$$

The Gamma function with positive arguments seems unavoidable, and comes with good asymptotics. Recall that for $t > 0$,

$$\Gamma(t) = \int_0^\infty x^{t-1}e^{-x}\,dx, \quad \Gamma(t+1) = t\Gamma(t), \quad \Gamma(m+1) = m!;$$
$$\lim_{t\to\infty} \frac{\Gamma(t+1)}{\sqrt{2\pi}\cdot e^{-t}t^{t+1/2}} = 1. \tag{2.58}$$

The natural Gamma function generalization of (2.57) is valid.

Theorem 2.4.

$$\lim_{n\to\infty} \frac{A(\lambda+1;n)}{n^\lambda/\Gamma(\lambda+1)} = 1. \tag{2.59}$$

Proof. First observe that

$$A(\lambda+1;n) = \frac{(\lambda+1)(\lambda+2)\cdots(\lambda+n)}{n!} = \frac{\Gamma(\lambda+n+1)/\Gamma(\lambda+1)}{n!}, \tag{2.60}$$

so (2.59) is equivalent to

$$\lim_{n\to\infty} \frac{\Gamma(\lambda+n+1)}{n^\lambda n!} = 1. \tag{2.61}$$

By multiplying limits and using (2.58), we see that the left-hand side of (2.61) becomes

$$\lim_{n\to\infty} \left(\left(\sqrt{2\pi}\cdot e^{-(n+\lambda)}(n+\lambda)^{n+\lambda+1/2}\right) n^{-\lambda} \left(\frac{1}{\sqrt{2\pi}\cdot e^{-n}n^{n+1/2}} \right) \right)$$
$$= \lim_{n\to\infty} \left(e^{-\lambda} \left(\frac{n+\lambda}{n}\right)^n \left(\frac{n+\lambda}{n}\right)^{\lambda+1/2} \right) = e^{-\lambda}\cdot e^{\lambda}\cdot 1 = 1.$$

□

Corollary 2.5.

$$\lim_{x\to 1^-} (1-x)^{\lambda+1} \left(\sum_{n=0}^{\infty} n^\lambda x^n \right) = \Gamma(\lambda+1). \tag{2.62}$$

Proof. By (2.56),

$$(1-x)^{\lambda+1} \left(\Gamma(\lambda+1) + \sum_{n=0}^{\infty} \Gamma(\lambda+1)A(\lambda+1;n)x^n \right) = \Gamma(\lambda+1). \tag{2.63}$$

By Lemma 2.3 and Theorem 2.4, we may replace $\Gamma(\lambda+1)A(\lambda+1;n)$ in (2.63) by n^λ without affecting the limit. □

2.6. **Applications to the Stern sequence.** Remember the Stern sequence? Consider

$$\mathcal{S}(z) = \sum_{n=0}^{\infty} s(n)z^n. \tag{2.64}$$

It's far from clear how to apply this discussion to $\mathcal{S}(z)$, because of the irregular behavior of the growth of $(s(n))$. However,

$$(1-z)^{-1}\mathcal{S}(z) = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} s(k)\right) z^n = \sum_{n=0}^{\infty} S(n)z^n, \tag{2.65}$$

and we can say something about the growth of $S(n)$. In fact, $S(2^r) = \frac{1}{2}(3^r+1)$, so by the monotonicity of S, if $n \in I_r$, then

$$2^r \le n \le 2^{r+1}, \qquad \tfrac{1}{2}(3^r+1) \le S(n) \le \tfrac{1}{2}(3^{r+1}+1). \tag{2.66}$$

Since $\log_2 n - 1 \le r \le \log_2 n$, without trying to be very careful, we see that

$$\tfrac{1}{6}n^\gamma + \tfrac{1}{2} \le S(n) \le \tfrac{3}{2}n^\gamma + \tfrac{1}{2}, \qquad \gamma = \frac{\log 3}{\log 2} \approx 1.585. \tag{2.67}$$

Why aren't we careful in the estimate? We have seen that $S(2^r) = \frac{1}{2}(3^r+1)$ and it will be an exercise (easy!) to show that $S(3\cdot 2^{r-1}) = 3^r + 1$. This means that

$$\frac{S(2^r)}{(2^r)^\gamma} \to \frac{1}{2} = .5, \qquad \frac{S(3\cdot 2^{r-1})}{(3\cdot 2^{r-1})^\gamma} \to \left(\frac{2}{3}\right)^\gamma \approx .525899. \tag{2.68}$$

In other words, $\lim_{n\to\infty} n^{-\gamma}s(n)$ does not exist; (2.66) implies that there exist $\alpha > 0$ and β so that

$$\alpha \le \frac{s(n)}{n^\gamma} \le \beta. \tag{2.69}$$

The best we can hope from Corollary 2.5 are upper and lower bounds on the growth. By combining (2.69) with the results of the last two sections, and keeping in mind that the estimates apply to $(1-z)^{-1}\mathcal{S}(z)$. we obtain the following corollary, which is improved in the next section as Theorem 2.10.

Corollary 2.6. *We have the following estimate on $\mathcal{S}(x)$ as real $x \to 1^-$:*

$$\frac{\alpha\Gamma(\gamma+1)(1+o(1))}{(1-x)^\gamma} \le \mathcal{S}(x) \le \frac{\beta\Gamma(\gamma+1)(1+o(1))}{(1-x)^\gamma}. \tag{2.70}$$

Proof. It follows from (2.69) and Lemma 2.3 that

$$\alpha\sum_{n=0}^{\infty} n^\gamma x^n \le \sum_{n=0}^{\infty} S(n)x^n \le \beta\sum_{n=0}^{\infty} n^\gamma x^n.$$

Now use Corollary 2.5 with $\lambda = \gamma$, and multiply by $(1-x)^{-1}$. $\square$

To provide a numerical version of (2.70), Mathematica tells us that $\Gamma(\gamma+1) \approx 1.41364$. Numerical evidence suggests that for t close to 1, $(1-t)^\gamma \mathcal{S}(t)$ oscillates in the range $.725189 \pm .000003$. Comparison with (2.70) suggests that

$$\frac{S(n)}{n^\gamma} \approx .512992 \pm .000002. \tag{2.71}$$

This is amazingly close to halfway between the values in (2.68).

We could also guess the order of growth of $\mathcal{T}(x)$, using the functional equation $\mathcal{T}(z) = (1+z+z^2)\mathcal{T}(z^2)$. Let $z = 1-\epsilon/2$ where $\epsilon \approx 0$, so that, in practical terms, $z^2 = 1-\epsilon$ and $1+z+z^2 \approx 3$. Then $\mathcal{T}(1-\epsilon/2) \approx 3\mathcal{T}(1-\epsilon)$. The "nice" function $f(1-t) = \phi(t)$ satisfying $f(1-\epsilon/2) = 3f(1-\epsilon)$ is $\phi(t) = c(1-t)^\gamma$ for some c.

In the next section, we will make the previous remarks more rigorous as we give a more detailed analysis of the behavior of $|\mathcal{S}(te^{2\pi i\alpha})|$ for fixed α as real $t \to 1^-$.

2.7. **Computations. Warning: the material in this section is subject to improvement, revision and/or retraction!** First note that $|\mathcal{S}(z)| = |z||\mathcal{T}(z)|$, so as $|z| \to 1$, it doesn't matter much which function is used; $|\mathcal{T}(z)|$ will be estimated here. In this section t is always a real number in $(0,1)$.

Our first step is to give an admittedly peculiar-looking "ruler" to measure $|z|$. Let

$$\sigma(t) = -\log_2(\log_2(t^{-1})); \qquad \sigma(t) = m \iff t := t_m = 2^{-2^{-m}}. \tag{2.72}$$

Reading from the inside out, σ maps $(0,1)$ to $(\infty,1)$ to $(\infty,0)$ to $(\infty,-\infty)$ to $(-\infty,\infty)$ in a monotone way, with the delightfully useful property that

$$\sigma(t^2) = \sigma(t) - 1 \implies t_m^2 = t_{m-1}. \tag{2.73}$$

It follows from (2.32) and (2.33) that for any positive integer v,

$$\mathcal{T}(t_m) = \prod_{j=0}^{\infty}(1+t_{m-j}+t_{m-j-1}) = \prod_{j=0}^{v-1}(1+t_{m-j}+t_{m-j-1}) \cdot \mathcal{T}(t_{m-v}). \tag{2.74}$$

It's worth noting that as $m \to -\infty$, $t_m \to 0$ at a doubly exponential rate: $t_{-5} \approx 2.33 \times 10^{-10}$. But as $m \to \infty$, the growth to 1 is only singly exponential:

$$1 - \frac{\log 2}{2^m} + \frac{(\log 2)^2/2}{2^{2m}} > t_m > 1 - \frac{\log 2}{2^m}. \tag{2.75}$$

The midpoint of the ruler is $t_0 = \frac{1}{2}$.

Our second step is to show that in numerical work, we may safely ignore the terms in the infinite product involving t_m when m is very negative.

Lemma 2.7. *For* $|z| < 1$,

$$|\mathcal{T}(z) - 1| \le \frac{|z|}{(1-|z|)^2}; \tag{2.76}$$

in particular,

$$|z| < \frac{1}{4} \implies |\mathcal{T}(z) - 1| \le 2|z|.$$

Proof. It is easy to show that for $n \ge 2$, $s(n) \le n-1$, hence

$$|\mathcal{T}(z) - 1| = \left| \sum_{n=1}^{\infty} s(n+1) z^n \right| \le \sum_{n=1}^{\infty} n |z|^n = \frac{|z|}{(1-|z|)^2}.$$

If $|z| \le \frac{1}{4}$, then $(1-|z|)^2 \ge \frac{9}{16} > \frac{1}{2}$. ☐

What this means is that stopping the infinite product in (2.74) when $v = \lceil m \rceil + 5$ gives $\mathcal{T}(t_m)$ as a finite product of polynomials times $\mathcal{T}(t_{m-v})$, where $m - v \le -5$, so that $t_{m-v} < \frac{1}{2} \cdot 10^{-9}$ and so $|1 - \mathcal{T}(t_{m-v})| \le 10^{-9}$. For most numerical purposes, $|\mathcal{T}(t_m)|$ can be identified as this finite product.

An important consequence of Lemma 2.7 is that $\mathcal{T}(z) \ne 0$ in the open unit disk:

Corollary 2.8. *If $|z| < 1$, then $\mathcal{T}(z) \ne 0$.*

Proof. If $|z| < \frac{1}{3}$, then (2.76) implies that $|\mathcal{T}(z) - 1| \le \frac{3}{4}$, so $\mathcal{T}(z) \ne 0$ for $|z| < \frac{1}{3}$. Suppose $\mathcal{T}(z) \ne 0$ for $|z| < \rho$. Since $1 + z + z^2$ has no zeros inside the unit circle, $\mathcal{T}(z) = (1+z+z^2)\mathcal{T}(z^2)$ implies that $\mathcal{T}(z) \ne 0$ for $|z| < \rho^{1/2}$, and hence by induction for $|z| < \rho^{1/2^n}$. This is true for all n and completes the proof. ☐

A more precise version of Corollary 2.6 requires a classical lemma.

Lemma 2.9. *If $\sum_{n=0}^{\infty} |a_n| = M < \infty$, then $\prod_{n=0}^{\infty}(1 + a_n) \to p > 0$.*

Proof. First define the partial products

$$p_n = \prod_{k=0}^{n}(1 + a_k) \iff \log p_n = \sum_{k=0}^{n} \log(1 + a_n).$$

Since $a_n \to 0$ and $\sum |a_n| < \infty$, $\sum_n \log(1+a_n)$ converges by the Bounded Comparison test. It follows that $(e^{\log p_n})$ converges to a positive value. ☐

We now discuss $\mathcal{T}(t)$ for real $t \to 1^-$; a first observation is that (2.34) implies that $\mathcal{T}(t)$ is an increasing positive real function in t. For real $m \ge 0$, let

$$h(m) = (1 - t_m)^{\gamma} \mathcal{T}(t_m). \tag{2.77}$$

We saw in Corollary 2.6 that $h(m)$ is a bounded function as $m \to \infty$. We now show that, in the limit, it is a periodic function with period 1.

Theorem 2.10. *There is a positive function ψ, defined on $[0, 1)$ so that for fixed $\alpha \in [0, 1)$ and integral k,*

$$\lim_{k \to \infty} h(k + \alpha) = \psi(\alpha). \tag{2.78}$$

Proof. Write

$$h(k + \alpha) = h(\alpha) \prod_{j=1}^{k} \frac{h(j + \alpha)}{h(j - 1 + \alpha)},$$

and consider the convergence of the infinite product

$$\prod_{j=1}^{\infty} \frac{h(j+\alpha)}{h(j-1+\alpha)}. \tag{2.79}$$

Let $u = t_{j+\alpha}$ in (2.79), so $u^2 = t_{j-1+\alpha}$, and recall that $\frac{3}{2} < \gamma < 2$ by (2.67). Then

$$\frac{h(j+\alpha)}{h(j-1+\alpha)} = \frac{(1-t_{j+\alpha})^{\gamma}\mathcal{T}(t_{j+\alpha})}{(1-t_{j-1+\alpha})^{\gamma}\mathcal{T}(t_{j-1+\alpha})} = \frac{(1-u)^{\gamma}}{(1-u^2)^{\gamma}} \cdot \frac{\mathcal{T}(u)}{\mathcal{T}(u^2)} = \frac{1+u+u^2}{(1+u)^{\gamma}}. \tag{2.80}$$

Since $j+\alpha \ge 0$, $u \le \frac{1}{2}$, so $1+u+u^2 \le 1+\frac{3}{2}u < 1+\gamma u < (1+u)^{\gamma}$, hence each factor in the infinite product is < 1. For the other inequality, write $u = 1-2w$. Then

$$\frac{1+u+u^2}{(1+u)^{\gamma}} = \frac{3-6w+4w^2}{(2-2w)^{\gamma}} > \frac{3(1-w)^2}{2^{\gamma}(1-w)^{\gamma}} = (1-w)^{2-\gamma} > 1-w. \tag{2.81}$$

But by (2.75),

$$1-w = \frac{1+u}{2} > u > 1 - \frac{\log 2}{2^{j+\alpha}}. \tag{2.82}$$

The factors in (2.79) are then $1+\mathcal{O}(2^{-j})$, and so the infinite product converges to a positive limit by Lemma 2.9. □

The numerical evidence suggests that

$$.7251918 \ge \psi(\alpha) \ge .7251858. \tag{2.83}$$

The next estimate we wish to consider is $|\mathcal{T}(\omega t)|$ for $\omega = e^{2\pi i/3}$, which has a radically different behavior. Since $s(n)$ is real, $\mathcal{T}(\bar{z}) = \overline{\mathcal{T}(z)}$, hence $|\mathcal{T}(\omega t)| = |\mathcal{T}(\omega^2 t)|$, and it doesn't matter which primitive cube root we use.

How does $\mathcal{T}(\omega t)$ behave as $t \to 1^-$? For $k \in \mathbb{Z}$, $1 + \omega^{2^k} + \omega^{2^{k+1}} = 1 + \omega^{(-1)^k} + \omega^{(-1)^{k+1}} = 0$. It is plausible that $\mathcal{S}(\omega t)$ should go to zero rapidly. Helpfully,

$$|1+t\omega+t^2\omega^2|^2 = |1+t\omega^2+t^2\omega|^2 = \left(1-\tfrac{t+t^2}{2}\right)^2 + \tfrac{3}{4}\left(t-t^2\right)^2 = (1-t)(1-t^3). \tag{2.84}$$

It follows that $|1+t\omega+t^2\omega^2|$ is decreasing quadratically to 0 as t increases to 1^-. In particular, the factor is decreasing for increasing m in t_m, so that

$$|\mathcal{T}(\omega t_m)| = \prod_{j=0}^{\infty}(1-t_{m-j})^{1/2}(1-t_{m-j}^3)^{1/2} \tag{2.85}$$

is decreasing in increasing m. The asymptotics is clumsy; it is easier to describe $|\mathcal{T}(\omega t_m)|$ as a function of m than to describe $|\mathcal{T}(\omega t_m)|$ as a function of t_m. Let

$$\Upsilon(m) = \frac{(3/2)^{m/2}(\log 2)^m}{2^{m^2/2}}, \tag{2.86}$$

and let

$$v(m) = \Upsilon(m)^{-1}|\mathcal{T}(\omega t_m)|. \tag{2.87}$$

Theorem 2.11. *There is a positive function η, defined on $[0,1)$ so that for fixed $\alpha \in [0,1)$ and integral k,*

$$\lim_{k\to\infty} v(k+\alpha) = \eta(\alpha). \tag{2.88}$$

Proof. As before, fix α and write

$$v(k+\alpha) = v(\alpha)\prod_{j=1}^{k}\frac{v(j+\alpha)}{v(j-1+\alpha)};$$

again, we wish to show that the infinite product converges. But

$$\frac{v(j+\alpha)}{v(j-1+\alpha)} = \frac{\Upsilon(j-1+\alpha)}{\Upsilon(j+\alpha)}\cdot\frac{|\mathcal{T}(\omega t_{j+\alpha})|}{|\mathcal{T}(\omega t_{j-1+\alpha})|}$$

$$= \frac{2^{j+\alpha-1/2}}{\sqrt{3/2}(\log 2)}(1-u)(1+u+u^2)^{1/2} = \frac{1-u}{(\log 2)/2^{j+\alpha}}\cdot\frac{\sqrt{1+u+u^2}}{\sqrt{3}},$$

where $u = t_{j+\alpha} \approx 1-(\log 2)/2^{(j+\alpha)}$. A computation, which we omit, shows that $\frac{v(j+\alpha)}{v(j-1+\alpha)} = 1+\mathcal{O}(2^{-j})$, hence, as was the case with Theorem 2.10, Lemma 2.9 implies that the series converges. □

For numerical reference, $|\mathcal{T}(\omega/2)| \approx .549$ and $\sqrt{3/2}(\log 2) \approx .849$. Theorem 2.11 implies that $|\mathcal{T}(\omega t_m)|$ goes to zero faster than any polynomial in $1-t_m$. The numerical evidence suggests that the range of $\eta(\alpha)$ is roughly $.2838218 \pm .0000002$.

Let $\zeta_d = e^{2\pi i/d}$ be a primitive d-th root of unity. There is a strong connection between $\mathcal{T}(z)$ and $\mathcal{T}(\zeta_{2^r}^{\ell}z)$. For convenience, note that $\zeta_{2^r}^{\ell 2^j} = \zeta_{2^{r-j}}^{\ell}$.

Lemma 2.12. *For $r\in\mathbb{N}$ and $\ell\in\mathbb{Z}$,*

$$\mathcal{T}(\zeta_{2^r}^{\ell}z) = \prod_{j=0}^{r-1}\left(\frac{1+\zeta_{2^r}^{2^j\ell}z+\zeta_{2^r}^{2^{j+1}\ell}}{1+z^{2^j}+z^{2^{j+1}}}\right)\mathcal{T}(z). \tag{2.89}$$

Proof. This follows from (2.33) and $\zeta_{2^r}^{2^k} = 1$ for $k\ge r$; all but the first r factors in the infinite products for $\mathcal{T}(\zeta_{2^r}^{\ell}z)$ and $\mathcal{T}(z)$ are the same. □

Theorem 2.13. *For any fixed (ℓ,r), and real $t\to 1^-$, there exist non-zero constants $\alpha(\ell,r),\beta(\ell,r)$ so that $\alpha(\ell,r)\le(1-t)^{\gamma}\mathcal{T}(\zeta_{2^r}^{\ell}t)\le\beta(\ell,r)$.*

Proof. Write (2.89) as

$$\mathcal{T}(\zeta_{2^r}^{\ell}z) = \frac{A(z)}{B(z)}\mathcal{T}(z). \tag{2.90}$$

Observe that $A(z)$ and $B(z)$ are both (bounded) polynomials, and $B(\zeta_{2^m}^{\ell})\ne 0$, since $1+z^{2^j}+z^{2^{j+1}}=0$ implies that $z^{2^j}\in\{\omega,\omega^2\}$. Thus, $M\ge|\frac{A(z)}{B(z)}|>\epsilon>0$ for suitable $M,\epsilon>0$, and the result follows from Theorem 2.10. □

It follows that $|\mathcal{T}(e^{i\alpha}t)| \to \infty$ on a dense set of rays, those with angles $\alpha = \frac{2\pi\ell}{2^r}$. A similar argument would show that the asymptotic behavior of $|\mathcal{T}|$ is the same (up to multiplicative constants) on any two rays whose angles differ by a multiple of $\frac{2\pi}{2^r}$, unless $B(z)$ might be zero. Fortunately, these cases are covered by Theorem 2.11.

Theorem 2.14. *Suppose* $gcd(\ell, 3) = 1$. *Then*

$$\lim_{x\to 1^-} |\mathcal{T}(e^{2\pi i\cdot \frac{\ell}{3\cdot 2^r}}x)| = 0. \tag{2.91}$$

Proof. Observe that $e^{2\pi i\cdot\frac{\ell}{3}} = \omega$ or ω^2. Without loss of generality, choose the former. Since each of the factors in (2.33) is bounded by 3 in absolute value,

$$|\mathcal{T}(e^{2\pi i\cdot \frac{\ell}{3\cdot 2^r}}x)| \le 3^r|\mathcal{T}(\omega x^{2^r})|. \tag{2.92}$$

The upper bound goes to zero quite rapidly. □

Thus, $|\mathcal{T}(e^{i\alpha}x)| \to 0$ on a different dense set of rays, those whose angles are $\alpha = \frac{2\pi\ell}{3\cdot 2^r}$.

The behavior on other rays is likely to be difficult to understand. If $\alpha = \frac{2\pi p}{q}$, where $q \ge 5$ is odd and $\gcd(p, q) = 1$, we can say something, because $2^{\phi(q)}p \equiv p \pmod q$. Observe that, as $m \to \infty$, $1 + t_m e^{i\theta} + t_{m-1}e^{2i\theta} \to 1 + e^{i\theta} + e^{2i\theta}$ and

$$\frac{\mathcal{T}(t_{m+\phi(q)}e^{i\alpha})}{\mathcal{T}(t_m e^{i\alpha})} \to \prod_{j=0}^{\phi(q)-1}(1 + e^{2\pi i(2^j p/q)} + e^{2\pi i(2^{j+1}p/q)}) = \prod_{j=0}^{\phi(q)-1}\frac{1 - e^{2\pi i\cdot 3\cdot(2^j p/q)}}{1 - e^{2\pi i\cdot(2^j p/q)}}. \tag{2.93}$$

It is difficult to say much about the set of q to which the hypothesis of the following observation hold. A careful analysis is still to be written.

Suppose $q \ge 5$ is odd and there exists s so that $2^s \equiv 3 \pmod q$ or $2^s \equiv -3 \pmod q$. Then $|\mathcal{T}(xe^{2\pi ip/q})|$ should be bounded as $x \to 1^-$. Consider the quotient in (2.93),

$$\prod_{j=0}^{\phi(q)-1}\frac{1 - e^{2\pi i\cdot 3\cdot 2^j p/q)}}{1 - e^{2\pi i\cdot(2^j p/q)}}. \tag{2.94}$$

If $2^s \equiv 3 \pmod q$, then the factors in the numerator of (2.94) are a permutation of the factors in the denominator, and so the product is 1, which suggests that, asymptotically, $\mathcal{T}(e^{i\alpha}t_m)$ should approach periodicity in m with period dividing $\phi(q)$. If $2^s \equiv -3 \pmod q$, then the factors in the numerator of (2.94) are a permutation of the conjugates of the factors in the denominator, so the absolute value of the product is 1, which suggests that, asymptotically, $|\mathcal{T}(e^{i\alpha}t_m)|$ should approach periodicity in m with period dividing $\phi(q)$. These hypotheses are satisfied if 2 is a primitive root mod q, the study of which is a famous and extremely difficult question, and they obviously *cannot* be satisfied if $3 \mid m$. Of the odd integers $6t \pm 1 \le 100$, only $17, 31, 41, 43, 65, 73, 85, 89$ do not satisfy this criterion.

The discussion of $|\mathcal{T}(z)|$ on rays is complemented by a discussion on $|z| = r$. Since $\mathcal{T}(z)$ is analytic and non-zero on the unit disk, we have by Jensen's formula,

$$0 = \log|\mathcal{T}(0)| = \frac{1}{2\pi}\int_0^{2\pi} \log|\mathcal{T}(re^{i\theta})|\,d\theta, \quad 0 \le r < 1. \tag{2.95}$$

It all balances out. Let $\chi(r)$ denote the number of $\theta \in [0, 2\pi)$ for which $|\mathcal{T}(re^{i\theta})| = 1$. One expects the number of these "crossings" to grow quite rapidly as $r \to 1^-$.

Finally, it follows from Corollary 2.8 that the reciprocal of $\mathcal{T}(z)$ is also an analytic function in the disk. Let

$$\begin{aligned} \mathcal{U}(x) := \frac{1}{\mathcal{T}(X)} &= \prod_{k=0}^{\infty} \frac{1}{1 + X^{2^k} + X^{2^{k+1}}} := \sum_{n=0}^{\infty} u(n)X^n \\ &= \prod_{k=0}^{\infty} \frac{1 - x^{2^k}}{1 - x^{3\cdot 2^k}} = \prod_{k=0}^{\infty}(1 - X^{2^k} + X^{3\cdot 2^k} - X^{4\cdot 2^k} + \dots) \\ &= 1 - x - x^2 + 2x^3 - 2x^4 + 4x^6 - 4x^7 - 2x^8 + 6x^9 - 4x^{10} - 2x^{11} + \cdots . \end{aligned} \tag{2.96}$$

A few brief facts about the mysterious (u_n). It is easy to show that the generating function for $\mathcal{U}$ over $\mathbb{Z}/2\mathbb{Z}$ is $\frac{1-X^3}{1-X} = 1 + X + X^2$, so $u(n)$ is odd for $n \le 2$ and even for $n \ge 3$. (No other congruence properties seem to be easy.) We also have

$$\begin{aligned} (1 + X + X^2)\mathcal{U}(X) = \mathcal{U}(X^2) &= \sum_{n=0}^{\infty} u(n)X^{2n}, \\ \implies u(2k) + u(2k-1) + u(2k-2) &= u(k) \\ \text{and} \qquad u(2k+1) + u(2k) + u(2k-1) &= 0 \\ \implies u(2n) - u(2n-3) = u(n); \qquad & u(2n+1) - u(2n-2) = -u(n) \end{aligned} \tag{2.97}$$

Numerical data strongly suggest that $u(n) > 0$ iff $3 \mid n$, and that $u(3n) \ge |u(3n+1)| \ge |u(3n+2)|$. It would be interesting to find a combinatorial interpretation for $u(n)$.