STERN NOTES, CHAPTER 4 (FIRST DRAFT)

BRUCE REZNICK, UIUC

1. Simple continued fractions

In this set of notes, we talk about simple continued fractions (numerators = 1, *scf* for short) and their relationship to the Stern sequence. There is a close relationship between simple continued fractions and the Euclidean algorithm. As a representative example,

$$2+\cfrac{1}{3+\cfrac{1}{7}} = 2+\frac{1}{\frac{22}{7}} = 2+\frac{7}{22} = \frac{51}{22}, \qquad \begin{array}{ccccccc} 51 & = & 2 & \times & 22 & + & 7 \\ 22 & = & 3 & \times & 7 & + & 1 \\ 7 & = & 7 & \times & 1 & + & 0 \end{array}. \tag{1}$$

Somewhat more formally, if $z = \frac{p}{q}$, $p, q \in \mathbb{N}$, $gcd(p, q) = 1$, then either $q = 1$ and $z = p$, or $p = a_0 q + r$, with $1 \le r \le q - 1$, and

$$z = \frac{p}{q} = \frac{a_0 q + r}{q} = a_0 + \frac{r}{q} = a_0 + \cfrac{1}{\cfrac{q}{r}}. \tag{2}$$

Since $r < q$, this sets up a finite recursive definition for scf, resulting in

$$z = x_1 + \cfrac{1}{x_2 + \cfrac{1}{\cdots + \cfrac{1}{x_n}}}, \tag{3}$$

with $x_n \ge 2$. Alternatively,

$$z = x_1 + \cfrac{1}{x_2 + \cfrac{1}{\cdots + \cfrac{1}{x_n - 1 + \cfrac{1}{1}}}}, \tag{4}$$

and we see that z always has a representation with an *odd* number of denominators.

It is helpful to think of the denominators as indeterminates in finding formulas. For $n \geq 1$, define $p_n(x_1, \ldots, x_n)$ and $q_n(x_1, \ldots, x_n)$ by:

$$\frac{p_n(x_1, \ldots, x_n)}{q_n(x_1, \ldots, x_n)} = x_1 + \cfrac{1}{x_2 + \cfrac{1}{\cdots + \cfrac{1}{x_n}}} \quad , \tag{5}$$

with the convention that $p_1(x_1) = x_1$ and $q_1(x_1) = 1$. There is an immediate relation:

$$\frac{p_n(x_1, \ldots, x_n)}{q_n(x_1, \ldots, x_n)} = x_1 + \frac{1}{\frac{p_{n-1}(x_2, \ldots, x_n)}{q_{n-1}(x_2, \ldots, x_n)}} = \frac{x_1 p_{n-1}(x_2, \ldots, x_n) + q_{n-1}(x_2, \ldots, x_n)}{p_{n-1}(x_2, \ldots, x_n)}. \tag{6}$$

It follows that $q_n(x_1, \ldots, x_n) = p_{n-1}(x_2, \ldots, x_n)$, and so it is natural to define $p_0 = 1$ and say goodbye to q_n. For $n \geq 2$,

$$p_n(x_1, \ldots, x_n) = x_1 p_{n-1}(x_2, \ldots, x_n) + p_{n-2}(x_3, \ldots, x_n) \tag{7}$$

In order to make this recurrence sensible for $n = 1$, it is customary to define $p_{-1} = 0$. The traditional name for p_n is the *continuant*. Here are some of the smaller values.

$$\begin{gathered} p_{-1} = 0, \qquad p_0 = 1, \qquad p_1(x_1) = x_1, \qquad p_2(x_1, x_2) = x_1 x_2 + 1, \\ p_3(x_1, x_2, x_3) = x_1 x_2 x_3 + x_1 + x_3, \\ p_4(x_1, x_2, x_3, x_4) = x_1 x_2 x_3 x_4 + x_1 x_2 + x_1 x_4 + x_3 x_4 + 1. \end{gathered} \tag{8}$$

It is evident from the definition that $p_n(x_1, \ldots, x_n)$ is linear in each of the variables, and so it is natural to wonder which monomials $x_{i_1} \cdots x_{i_r}$ appear. It turns out to be the terms whose *absent* variables appear in disjoint consecutive pairs. We define

$$B_i = B_i(x_i, x_{i+1}) = \frac{1}{x_i x_{i+1}}. \tag{9}$$

and for integers $m < n$, define $\mathcal{I}(m, n)$ to be

$$\emptyset \cup \{i = (i_1, \ldots, i_r) : m \leq i_1,\ i_j + 2 \leq i_{j+1}\ (1 \leq j \leq r - 1),\ i_r \leq n - 1\}. \tag{10}$$

That is, $\mathcal{I}(m, n)$ consists of the *first* elements of all sets of disjoint pairs $(i_j, i_j + 1)$ contained in $\{m, \ldots, n\}$.

Theorem 1. *For all $n \geq 0$,*

$$p_n(x_1, \ldots, x_n) = x_1 \ldots x_n \sum_{i \in \mathcal{I}(1,n)} B_{i_1} \cdots B_{i_r}. \tag{11}$$

Proof. Let

$$\phi_n(x_1, \ldots, x_n) := \frac{p_n(x_1, \ldots, x_n)}{x_1 \cdots x_n}. \tag{12}$$

We see from (8) that $\phi_0 = \phi_1 = 1$ and $\phi_2(x_1, x_2) = 1 + B_1$, so the theorem is valid for $n \le 2$. After division by $x_1 \cdots x_n$, the basic recurrence (7) becomes

$$\phi_n(x_1, \dots, x_n) = \phi_{n-1}(x_2, \dots, x_n) + \frac{\phi_{n-2}(x_3, \cdots, x_n)}{x_1 x_2}. \tag{13}$$

We divide $i \in \mathcal{I}(1, n)$ into two classes. First, if $1 \notin i$, then $i \in \mathcal{I}(2, n)$ (possibly $i = \emptyset$.) Otherwise, $1 \in i$, so $2 \notin i$ and $i = (1, i')$ (as a concatenation), where $i' \in \mathcal{I}(3, n)$ (possibly $i' = \emptyset$.) It follows by induction that

$$\begin{aligned} &\phi_n(x_1, \dots, x_n) \\ &= \sum_{i \in \mathcal{I}(2,n)} B_{i_1} \cdots B_{i_r} + B_1 \cdot \sum_{i \in \mathcal{I}(3,n)} B_{i_1} \cdots B_{i_r} \\ &= \sum_{i \in \mathcal{I}(1,n)} B_{i_1} \cdots B_{i_r}, \end{aligned} \tag{14}$$

as desired. □

The next observation is critical to understanding the Stern sequence.

Theorem 2. *For all n,*

$$p_n(x_1, \dots, x_n) = p_n(x_n, \dots, x_1). \tag{15}$$

Proof. The condition on the missing indices in the terms of the continuant is symmetric under their mirror reflection. To be precise, if $f : x_i \to x_{n+1-i}$, then $x_i x_{i+1} \to x_{n+1-i} x_{n-i}$, so $B_i \to B_{n-i}$ (for $1 \le i \le n-1$), and the condition of the separation of indices by at least two is preserved. □

Corollary 3. *For all n,*

$$p_n(x_1, \dots, x_n) = x_n p_{n-1}(x_1, \dots, x_{n-1}) + p_{n-2}(x_1, \dots, x_{n-2}). \tag{16}$$

We now present some "expansion" formulas for continuants.

Theorem 4. *Using the conventions $p_0 = 1$ and $p_{-1} = 0$, we have, for $m, n \ge 0$*

$$\begin{aligned} p_{m+n}(x_1, \cdots, x_m, y_1, \dots, y_n) &= p_m(x_1, \cdots, x_m) p_n(y_1, \cdots, y_n) \\ &+ p_{m-1}(x_1, \cdots, x_{m-1}) p_{n-1}(y_2, \cdots, y_n). \end{aligned} \tag{17}$$

Proof. Although this can be proved by looking at the "missing terms", it is probably clearest to prove by induction on n for fixed m. For $k \le m$, let $p_k = p_k(x_1, \dots, x_k)$ for short, and let the desired equation for n be expressed as $LHS(n) = RHS(n)$. Then $LHS(0) = RHS(0)$ is $p_m = p_m \cdot 1 + p_{m-1} \cdot 0$, which is trivial, and $LHS(1) = RHS(1)$ is $p_{m+1}(x_1, \cdots, x_m, 1) = p_m \cdot p_1(y_1) + p_{m-1} \cdot 1$, which is the basic recurrence. Since $LHS(n) = y_n LHS(n-1) + LHS(n-2)$ and $RHS(n) = y_n RHS(n-1) + RHS(n-2)$, the result follows by induction. □

It will be useful to have some special values of the continuant. These are to be used in conjunction with the corollary for full effect.

Lemma 5. *(1)* $p_n(x_1,\dots,x_{n-2},x_{n-1},1) = p_{n-1}(x_1,\dots,x_{n-2},x_{n-1}+1)$*;*
(2) $p_n(x_1,\dots,x_{n-2},x_{n-1},0) = p_{n-2}(x_1,\dots,x_{n-2})$*.*

Proof. Once again, let $p_k = p_k(x_1,\dots,x_k)$ for $k \le n-1$. Then

$$\begin{aligned} p_n(x_1,\dots,x_{n-2},x_{n-1},1) &= 1\cdot p_{n-1} + p_{n-2} = (x_{n-1}p_{n-2} + p_{n-3}) + p_{n-2} \\ &= (1+x_{n-1})p_{n-2} + p_{n-3} = p_{n-1}(x_1,\dots,x_{n-2},x_{n-1}+1). \end{aligned} \tag{18}$$

The second equation follows immediately from Corollary 3. □

We remark that informal "proofs" of these equations are:

$$x_{n-2} + \cfrac{1}{x_{n-1} + \cfrac{1}{1}} = x_{n-2} + \frac{1}{x_{n-1}+1}\,, \qquad x_{n-2} + \cfrac{1}{x_{n-1} + \cfrac{1}{0}} = x_{n-2} + \frac{1}{\infty} = x_{n-2}. \tag{19}$$

Theorem 6. *If* $m,n \ge 1$*, then*

$$\begin{aligned} &p_{m+n+1}(x_1,\dots,x_m,0,y_1,\dots,y_n) \\ &= p_{m+n-1}(x_1,\dots,x_m+y_1,\dots,y_n) = \\ &p_{m-1}(x_1,\dots,x_{m-1})p_n(y_1,\dots,y_n) + p_m(x_1,\dots,x_m)p_{n-1}(y_2,\dots,y_n). \end{aligned} \tag{20}$$

Proof. Both equalities can be proved by induction on n for fixed m; the second is actually easier to show directly.

For the first, let $p_k = p_k(x_1,\dots,x_k)$ again and for $n \ge 1$, let

$$a_n = p_{m+n+1}(x_1,\dots,x_m,0,y_1,\dots,y_n), \quad b_n = p_{m+n-1}(x_1,\dots,x_m+y_1,\dots,y_n). \tag{21}$$

Then

$$\begin{aligned} a_1 &= y_1 p_{m+1}(x_1,\dots,x_m,0) + p_m = y_1 p_{m-1} + p_m \\ &= y_1 p_{m-1} + (x_m p_{m-1} + p_{m-2}) = (y_1 + x_m)p_{m-1} + p_{m-2} = b_1, \end{aligned} \tag{22}$$

and

$$a_2 = y_2 a_1 + p_{m-1} = y_2 b_1 + p_{m-1} = b_2. \tag{23}$$

Since $a_n = y_n a_{n-1} + a_{n-2}$ and $b_n = y_n b_{n-1} + b_{n-2}$ for $n \ge 3$, the result follows by induction.

For the second identity, it is easier to argue directly, using Theorem 4:

$$\begin{aligned} &p_{m+n+1}(x_1,\dots,x_m,0,y_1,\dots,y_n) \\ &= p_{m+1}(x_1,\dots,x_m,0)p_n(y_1,\dots,y_n) + p_m(x_1,\dots,x_m)p_{n-1}(y_2,\dots,y_n) \\ &= p_{m-1}(x_1,\dots,x_{m-1})p_n(y_1,\dots,y_n) + p_m(x_1,\dots,x_m)p_{n-1}(y_2,\dots,y_n). \end{aligned} \tag{24}$$

□

The informal "proof" of the first of the equations is

$$x_n + \cfrac{1}{0 + \cfrac{1}{y_1 + \cfrac{1}{\cdots}}} = x_n + y_1 + \frac{1}{\cdots}. \tag{25}$$

Another identity of interest combines all of these:

Theorem 7. *If $m, n \geq 1$, then*

$$\begin{aligned} p_{m+n+1}(x_1, \ldots, x_m, z, y_1, \ldots, y_n) = z p_m(x_1, \ldots, x_m) p_n(y_1, \ldots, y_n) \\ + p_{m+n-1}(x_1, \ldots, x_m + y_1, \ldots, y_n). \end{aligned} \tag{26}$$

Proof. We first observe that continuants are multilinear polynomials, and hence

$$\begin{aligned} p_{m+n+1}(x_1, \ldots, x_m, z, y_1, \ldots, y_n) \\ = A(x_1, \ldots, x_m, y_1, \ldots, y_n) \cdot z + p_{m+n+1}(x_1, \ldots, x_m, 0, y_1, \ldots, y_n). \end{aligned} \tag{27}$$

for some function A, from first principles. We evaluate the constant term in (27) by (20), and it suffices to compute the coefficient of z. By (17),

$$\begin{aligned} p_{m+n+1}(x_1, \ldots, x_m, z, y_1, \ldots, y_n) \\ = p_{m+1}(x_1, \ldots, x_m, z) p_n(y_1, \ldots, y_n) + p_m(x_1, \ldots, x_m) p_{n-1}(y_2, \ldots, y_n) \end{aligned} \tag{28}$$

and

$$p_{m+1}(x_1, \ldots, x_m, z) = z \cdot p_m(x_1, \ldots, x_m) + p_{m-1}(x_1, \ldots, x_{m-1}, \tag{29}$$

so the expression for $A(x_1, \ldots, x_m, y_1, \ldots, y_n)$ is established. □

Corollary 8.

$$\frac{\partial p_n}{\partial x_k}(x_1, \ldots, x_n) = p_{k-1}(x_1, \ldots, x_{k-1}) p_{n-k}(x_{k+1}, \ldots, x_n). \tag{30}$$

The final identity has particular significance for the Stern sequence;

Theorem 9. *For all $n \geq 1$,*

$$p_n(x_1, \ldots, x_n) p_{n-2}(x_2, \ldots, x_{n-1}) = p_{n-1}(x_1, \ldots, x_{n-1}) p_{n-1}(x_2, \ldots, x_n) + (-1)^n. \tag{31}$$

Proof. First note that for $n = 1$, this equation is $x_1 \cdot 0 = 1 \cdot 1 + (-1)^1$, and for $n = 2$, it's $(x_1 x_2 + 1) \cdot 1 = x_1 \cdot x_2 + (-1)^2$, and both are true. Let

$$\begin{aligned} h_n(x_1, \ldots, x_n) = \\ p_n(x_1, \ldots, x_n) p_{n-2}(x_2, \ldots, x_{n-1}) - p_{n-1}(x_1, \ldots, x_{n-1}) p_{n-1}(x_2, \ldots, x_n). \end{aligned} \tag{32}$$

Then h_n is linear in x_n, and after expanding by (7), we see that the coefficient of x_n is

$$\begin{aligned} p_{n-1}(x_1, \ldots, x_{n-1}) p_{n-2}(x_2, \ldots, x_{n-1}) \\ - p_{n-1}(x_1, \ldots, x_{n-1}) p_{n-2}(x_2, \ldots, x_{n-1}) = 0 \end{aligned} \tag{33}$$

Thus h_n does not depend on x_n and

$$\begin{aligned} h_n(x_1,\dots,x_n) = h_n(x_1,\dots,x_{n-1},0) = \\ p_n(x_1,\dots,x_{n-1},0)p_{n-2}(x_2,\dots,x_{n-1}) - p_{n-1}(x_1,\dots,x_{n-1})p_{n-1}(x_2,\dots,x_{n-1},0) \\ = p_{n-2}(x_1,\dots,x_{n-2})p_{n-2}(x_2,\dots,x_{n-1}) - \\ p_{n-1}(x_1,\dots,x_{n-1})p_{n-3}(x_2,\dots,x_{n-2}) \\ = -h_{n-1}(x_1,\dots,x_{n-1}). \end{aligned} \tag{34}$$

The result follows by induction. □

We conclude this section with an application to Fibonacci numbers. Let

$$a_n = P_n(1,\dots,1). \tag{35}$$

Then $a_0 = 1$, $a_1 = 1$, $a_2 = 2$ and for $n \ge 2$, (7) gives $a_n = a_{n-1} + a_{n-2}$. It follows that $a_n = F_{n+1}$. That is, assuming there are n denominators below, we have

$$1 + \cfrac{1}{1 + \cfrac{1}{\ddots + \cfrac{\ddots}{1 + \cfrac{1}{1}}}} = \frac{F_{n+1}}{F_n}. \tag{36}$$

Theorem 1 implies that the polynomial $p_n(x_1,\dots,x_n)$ has F_{n+1} terms, which implies that $|\mathcal{I}(1,n)| = F_{n+1}$. By setting $x_i, y_j \equiv 1$ in Theorem 4, we recover the "known" addition formula:

$$F_{n+m+1} = F_{n+1}F_{m+1} + F_nF_m. \tag{37}$$

Under the same conditions, Theorem 6 says that

$$p_{m+n+1}(1,\dots,1,0,1\dots,1) = p_{m+n-1}(1,\dots,2,\dots,1) = F_mF_{n+1} + F_{m+1}F_n. \tag{38}$$

By Theorem 7, with $z = 1$,

$$F_{m+n+2} = F_{m+1}F_{n+1} + p_{m+n-1}(1,\dots,2\dots,1) \tag{39}$$

This is actually nothing new; combining (38) and (39), we find that

$$F_{m+n+2} = F_{m+1}F_{n+1} + F_mF_{n+1} + F_{m+1}F_n = F_{m+2}F_{n+1} + F_{m+1}F_n, \tag{40}$$

which is the addition formula, with $m \to m+1$. But it allows for a nice formula for general z:

$$p_{m+n+1}(1,\dots,1,z,1\dots,1) = (z-1)F_{m+1}F_{n+1} + F_{m+n+2}. \tag{41}$$

Finally, Theorem 9 implies another familiar Fibonacci identity:

$$F_{n+1}F_{n-1} = F_n^2 + (-1)^n. \tag{42}$$

2. The rule of four revisited

Recall that in the first set of notes, we considered an odd number n, $2^r < n < 2^{r+1}$, and wrote $n \sim [a_1, \dots, a_{2v+1}]$ if $[n]_2$, the base 2 representation of n, consists of a_1 1's, followed by a_2 0's, etc, ending with a_{2v+1} 1's. In this case

$$r + 1 = \sum_{j=1}^{2v+1} a_j. \tag{43}$$

We have already proved that

$$\frac{s(n)}{s(n+1)} = a_{2v+1} + \cfrac{1}{a_{2v} + \cfrac{1}{\cdots + \cfrac{1}{a_1}}} = \frac{p_{2v+1}(a_{2v+1}, \dots, a_1)}{q_{2v+1}(a_{2v+1}, \dots, a_1)}, \tag{44}$$

and in view of the last section,

$$s(n) = p_{2v+1}(a_1, \dots, a_{2v+1}), \qquad s(n+1) = p_{2v}(a_1, \dots, a_{2v}). \tag{45}$$

We also defined two related numbers. The first is the image of n in the reflection of the r-th row of the diatomic array:

$$n' = 3 \cdot 2^r - n, \tag{46}$$

so that $n = 2^r + k \implies n' = 2^{r+1} - k$. (This is defined whether or not n is even or odd. The second, $\overleftarrow{n}$, is defined for odd n by the property that $[\overleftarrow{n}]_2$ is the reversal of $[n]_2$; that is,

$$\overleftarrow{n} \sim [a_{2v+1}, \dots, a_1]. \tag{47}$$

The goal in this section is to show that "generically", each value m occurs in the r-th row of the diatomic array four times, and the successors are also related. (The singular cases occur when $[n]_2$ is a palindrome, or near-palindrome.)

Theorem 10. *Suppose n is odd and $s(n) = m$, Then $s(n') = s(\overleftarrow{n}) = s(\overleftarrow{n'}) = m$. Moreover, $\overleftarrow{n'} = (\overleftarrow{n})'$. Let $s(n+1) = a$, and let $\bar{a} \in \{1, m-1\}$ satisfy $a\bar{a} \equiv 1 \pmod m$. Then $s(n'+1) = m - a$, $s(\overleftarrow{n} + 1) = \bar{a}$ and $s(\overleftarrow{n'} + 1) = m - \bar{a}$. Moreover, if $\breve{n} = 2^{r+1} - 1 - n$, then*

$$s(n+1)s(\overleftarrow{n} + 1) = s(n)s(\breve{n}) + 1. \tag{48}$$

Proof. The equation $s(n) = s(n')$ follows either intuitively from the picture of the diatomic array, or from the basic recurrence: if $0 \le k \le 2^r$, then

$$\begin{aligned} s(2^r + k) &= s(2^r \cdot 1 + k) = s(2^r - k)s(1) + s(k)s(1+1) \\ &= s(2^r - k) + s(k) = s(2^r - (2^r - k)) + s(2^r - k) = s(2^r + (2^r - k)). \end{aligned} \tag{49}$$

Since n is odd, $s(n-1) + s(n+1) = s(n)$, hence $s(n-1) = m - a$. And since $2^r \le n - 1 < 2^{r+1}$, $n' + 1 = (n-1)'$ and we have $s(n'+1) = m - a$.

We look more carefully at n'. Suppose

$$n = 2^r + \left(\sum_{k=1}^{r-1} \epsilon_k 2^k\right) + 1, \qquad \epsilon_k \in \{0,1\}. \tag{50}$$

Then it is easy to compute $[n']_2$:

$$n' = 3\cdot 2^r - n = 2^{r+1} + \sum_{k=1}^{r-1} 2^k + 2\cdot 1 - n = 2^r + \sum_{k=1}^{r-1}(1-\epsilon_k)2^k + 1. \tag{51}$$

Informally, $[n]_2$ must begin and end in "1"; $[n']_2$ flips all digits except the first and last. A pattern of $a_1 > 1$ 1's, followed by a_2 0's, etc., turns into one 1, $a_1 - 1$ 0's, a_2 1's, etc, whereas one 1, followed by a_2 0's, a_3 1's, etc become $1 + a_2$ 0's, a_3 1's, etc. The same thing happens at the end (in reverse of course). In short (assuming that $a_1 > 1$ and $a_{2v+1} > 1$ if they appear below):

$$\begin{aligned} n \sim [a_1, \dots, a_{2v+1}] &\implies n' \sim [1, a_1 - 1, \dots, a_{2v+1} - 1, 1], \\ n \sim [1, a_2 \dots, a_{2v}, 1] &\implies n' \sim [a_2 + 1, \dots, a_{2v} + 1], \\ n \sim [a_1, \dots, a_{2v}, 1] &\implies n' \sim [1, a_1 - 1, \dots, a_{2v} + 1], \\ n \sim [1, a_2 \dots, a_{2v}, a_{2v+1}] &\implies n' \sim [a_2 + 1, \dots, a_{2v+1} - 1, 1]. \end{aligned} \tag{52}$$

In each of these four cases, it follows directly from Lemma 5(1) that $s(n) = s(n')$; informally, "1"'s at either end of the argument of a continuant can be absorbed by their nearest neighbor. Also, the symmetries in these equations make it clear that $\overleftarrow{n'} = (\overleftarrow{n})'$.

We turn to the reversals. In view of our earlier remarks,

$$s(n) = p_{2v+1}(a_{2v+1}, \dots, a_1) = p_{2v+1}(a_1, \dots, a_{2v+1}) = s(\overleftarrow{n}) = s(\overleftarrow{n'}). \tag{53}$$

We have also seen from our earlier formulas that

$$s(n+1) = p_{2v}(a_1, \dots, a_{2v}); \qquad s(\overleftarrow{n} + 1) = p_{2v}(a_2, \dots, a_{2v+1}). \tag{54}$$

It follows immediately from Theorem 9 that

$$s(n) p_{2v-1}(a_, \dots, a_{2v}) = s(n+1)s(\overleftarrow{n} + 1) + (-1)^{2v+1}, \tag{55}$$

hence $s(n+1)s(\overleftarrow{n}+1) \equiv 1 \pmod{s(n)}$. Since $s(n+1) = a$ and $1 \le s(\overleftarrow{n}+1) \le m$, we must have $s(\overleftarrow{n}+1) = \bar{a}$.

It is worth taking the time to interpret $p_{2v-1}(a_2, \dots, a_{2v})$. First, we need an alternative expression for n. We claim that $n \sim [a_1, \dots, a_{2v+1}]$ implies that

$$n = 2^{c_1} - 2^{c_2} + - \cdots + 2^{c_{2v+1}} - 1, \tag{56}$$

where

$$c_k = \sum_{i=k}^{2v+1} a_i. \tag{57}$$

The easiest way to prove this is by induction, and was done, I think, in the first set of notes. When $v = 0$, if $[n]_2$ consists of a_1 1's, then clearly $n = 2^{a_1} - 1$. Supposing the formula is valid as given and $[\bar{n}]_2$ consists of $[n]_2$, followed by a_{2v+2} 0's and a_{2v+3} 1's, then

$$\bar{n} = 2^{a_{2v+2}+a_{2v+3}} n + 2^{a_{2v+3}} - 1, \tag{58}$$

which, upon a small amount of reflection, establishes the inductive step.

Recall that $c_1 = r + 1$ and $c_{2v+1} = a_{2v+1}$, so that

$$2^{r+1} - 1 - n = 2^{c_2} - 2^{c_3} + - \cdots 2^{c_{2v+1}} := 2^{a_{2v+1}} \breve{n}. \tag{59}$$

Let $\breve{c}_k = c_k - a_{2v+1}$. Then

$$\breve{n} = 2^{\breve{c}_2} - 2^{\breve{c}_3} + - \cdots + 2^{\breve{c}_{2v}} - 1. \tag{60}$$

Thus, $\breve{n} \sim [a_2, \ldots, a_{2v}]$. That is, the outer blocks of 1's in $[n]_2$ are tossed aside and the other blocks flip parity. We have

$$s(2^{r+1} - 1 - n) = s(2^{a_{2v+1}} \breve{n}) = s(\breve{n}) = p_{2v-1}(a_{2v}, \ldots, a_2) = p_{2v-1}(a_2, \ldots, a_{2v}), \tag{61}$$

and the last formula is established. □

Example. For example, suppose $n = 35$. Then $[n]_2 = [100011]_2$, so $n \sim [1, 3, 2]$, hence $\overleftarrow{n} \sim [2, 3, 1]$ and $[\overleftarrow{n}]_2 = [110001]_2$, so that $\overleftarrow{n} = 49$. We also have $n' = 3 \cdot 32 - n = 61$, $[n']_2 = [111101]$, so $n' \sim [4, 1, 1]_2 = [3+1, 2-1, 1]_2$, and $\overleftarrow{n'} \sim [1, 1, 4]_2$ so that $[\overleftarrow{n'}]_2 = [101111]_2$, hence $\overleftarrow{n'} = 47 (= 3 \cdot 32 - 49)$. As a check, $a = s(n+1) = s(36) = 4$ and $s(\overleftarrow{n}+1) = s(50) = 7$, $s(n'+1) = s(62) = 5 = 9 - 4$ and $s(\overleftarrow{n'}+1) = s(48) = 2 = 9 - 7$. Finally, $2^6 - n - 1 = 28$, so $\breve{n} = 2^{-2} \cdot 28 = 7$ and $s(\breve{n}) = 3$, and indeed, $4 \cdot 7 = 3 \cdot 9 + 1$.

Finally, we report a peculiar result which will become valuable in the discussion of the Minkowski ?-function.

Theorem 11. *Suppose n is odd, $2^{r_0} < n < 2^{r_0} + 1$ and $r \geq r_0 + 1$. Then*

$$\frac{s(2^r + n)}{s(n)} = \frac{s((\overleftarrow{2^r + n})')}{s((\overleftarrow{2^r + n})' + 1)}. \tag{62}$$

Proof. The equality of the numerators is clear from Theorem 10. We unpack the denominator. First observe that $[2^r + n]_2$ consists of one "1", followed by $r - r_0 - 1$ 0's (possibly none) and then $[n]_2$. Thus, $[\overleftarrow{2^r + n}]_2$ consists of $[\overleftarrow{n}]_2$, followed by $r - r_0 - 1$ 0's (possibly none) and followed by one "1", and so $\overleftarrow{2^r + n} - 1 = 2^{r - r_0} \overleftarrow{n}$ and $s(\overleftarrow{2^r + n} - 1) = s(n)$. Finally, we observe once again that, with $m = \overleftarrow{2^r + n}$, since $2^r \leq m - 1 < 2^{r+1}$, we have $(m-1)' = m' + 1$, so the denominators are equal too. □

3. Some specific examples

We can apply the continued fraction formulas from the first section to the Stern sequence. What follows is far from exhaustive, but may serve to inspire you in considering the second homework assignment!

Example. We return to problem 6 on the first homework, restricting the sign. Suppose we are interested in computing $s[2^r n + k]$ for $r \ge r_0 = \lceil \log_2 k \rceil$; that is, $2^{r_0 - 1} < k < 2^{r_0}$. (We might as well assume k is odd and can rule out $k = 1$, because we know the result in this case. Suppose $k \sim [b_1, \dots, b_{2w+1}]$; as we have seen, $r_0 = \sum_j b_j$. Suppose also that $n \sim [a_1, \dots, a_{2v+1}]$. Then,

$$\begin{aligned} N = 2^r n + k = 2^r \cdot \big(2^{a_1 + \cdots + a_{2v+1}} - + \cdots + 2^{a_{2v+1}} - 1\big) \\ + 2^{b_1 + \cdots + b_{2w+1}} - + \cdots + 2^{b_{2w+1}} - 1 \end{aligned} \tag{63}$$

so

$$\begin{aligned} N \sim [a_1, \dots, a_{2v+1}, r - r_0, b_1, \dots, b_{2w+1}] \implies \\ s[N] = p_{2v+2w+3}(a_1, \dots, a_{2v+1}, r - r_0, b_1, \dots, b_{2w+1}). \end{aligned} \tag{64}$$

It follows by Theorem 7 that

$$\begin{aligned} s[N] = (r - r_0) p_{2v+1}(a_1, \dots, a_{2v+1}) p_{2w+1}(b_1, \dots, b_{2w+1}) \\ + p_{2v+2w+1}(a_1, \dots, a_{2v+1} + b_1, \dots, b_{2w+1}). \end{aligned} \tag{65}$$

We already know that $p_{2v+1}(a_1, \dots, a_{2v+1}) = s(n)$ and $p_{2w+1}(b_1, \dots, b_{2w+1}) = s(k)$. We claim that the last expression in (65) is $s(2^{r_0} n + k)$. Indeed, looking at (63) with $r = r_0$, we see that $2^{r_0}(-1)$ cancels $2^{b_1 + \cdots + b_{2w+1}}$, so that $[2^{r_0} n + k] \sim [a_1, \dots, a_{2v+1} + b_1, \dots, b_{2w+1}]$.

We believe that $s(2^r n - k)$ can be handled in a similar, but less interesting way, and omit the details.

For the last example, we adopt exponential notation in a transparently obvious way, so that, for example, (35) becomes $a_n = p_n(1^n)$ [yes, it should be lower case, a typo in the last installment], and

$$p_n(1^n) = F_{n+1} \tag{66}$$

There should be no confusion about expressions such as $p_{n+m}(1^m 2^n)$, etc.

Example. Recall our discussion from Notes, I about

$$n_r = \frac{2^{r+2} - (-1)^r}{3} = \frac{4}{3} \cdot 2^r - \frac{(-1)^r}{3}. \tag{67}$$

We showed that $s(n_r) = F_{r+2}$ and for $2^r < n < 2^{r+1}$, $s(n)$ achieves its maxima at n_r and n'_r. It is worth duplicating the computation of $s(n_r)$ using our current techniques, though we do not address the question of the maximum. First suppose $r = 2t$. Then

$$n_{2t} = \frac{2^{2t+2} - 1}{2^2 - 1} = 2^{2t} + 2^{2t-2} + \cdots + 2^2 + 1 \tag{68}$$

That is, $[n_{2t}]_2 = [1010\cdots 101]_2$, so $n_{2t} \sim [1^{2t+1}]$, and so $s[n_{2t}] = p_{2t+1}(1^{2t+1}) = F_{2t+2}$, as we'd expect. The situation is a little trickier for $r = 2t+1$:

$$n_{2t+1} = \frac{2^{2t+3}+1}{2^2-1} = 2n_{2t}+1 = 2^{2t+1}+2^{2t-1}+\cdots+2^3+2^1+1. \tag{69}$$

That is, $[n_{2t+1}]_2 = [1010\cdots 1011]_2$, so $n_{2t+1} \sim [1^{2t}2]$ and $s[n_{2t+1}] = p_{2t+1}(1^{2t}2)$. By Lemma 5, we can stretch that last "2" into "11", so $s[n_{2t+1}] = p_{2t+2}(1^{2t+2}) = F_{2t+3}$, again, as expected. Notice that $\overleftarrow{n_{2t}} = n_{2t}$ and, somewhat less obviously, $\overleftarrow{n_{2t+1}} = n'_{2t+1}$, which explains why these maxima occur twice, rather than four times.

We now compute $s(n_r \pm 2)$. There are two cases, depending on whether r is even or odd. By staring at the formulas for $[n_r]_2$, we see that

$$\begin{aligned} n_{2t}+2 &= 2^{2t}+2^{2t-2}+\cdots+2^4+2^2+2^1+1,\\ n_{2t}-2 &= 2^{2t}+2^{2t-2}+\cdots+2^4+2^1+1,\\ n_{2t+1}+2 &= 2^{2t+1}+2^{2t-3}+\cdots+2^5+2^3+2^2+1,\\ n_{2t+1}-2 &= 2^{2t+1}+2^{2t-3}+\cdots+2^5+2^3+1. \end{aligned} \tag{70}$$

Thus,

$$\begin{aligned} s(n_{2t}+2) &= p_{2t-1}(1^{2t-2},3),\\ s(n_{2t}-2) &= p_{2t-1}(1^{2t-3},2,2),\\ s(n_{2t+1}+2) &= p_{2t+1}(1^{2t-2},2,1,1) = p_{2t}(1^{2t-2},2,2),\\ s(n_{2t+1}-2) &= p_{2t+1}(1^{2t-1},2,1) = p_{2t}(1^{2t-1},3). \end{aligned} \tag{71}$$

More generally, using Theorem 4 to separate out the 1's, observe that

$$\begin{aligned} p_n(1^{n-1}a) &= a\cdot p_{n-1}(1^{n-1}) + 1\cdot p_{n-2}(1^{n-2}) = aF_n + F_{n-1},\\ p_n(1^{n-2}ab) &= (ab+1)p_{n-2}(1^{n-2}) + b\cdot p_{n-3}(1^{n-3}) = (ab+1)F_{n-1}+bF_{n-2}. \end{aligned} \tag{72}$$

It follows that

$$\begin{aligned} s(n_{2t}+2) &= 3F_{2t-1}+F_{2t-2}, \qquad & s(n_{2t+1}-2) &= 3F_{2t}+F_{2t-1},\\ s(n_{2t}-2) &= 5F_{2t-2}+2F_{2t-3}, \qquad & s(n_{2t+1}+2) &= 5F_{2t-1}+2F_{2t-2}. \end{aligned} \tag{73}$$

By iterating the Fibonacci recurrence, it is easy to see that $F_{n+2} = 3F_{n-1}+2F_{n-2}$, hence:

$$\begin{aligned} F_{n+2}-(3F_{n-1}+F_{n-2}) &= F_{n-2};\\ F_{n+2}-(5F_{n-2}+2F_{n-3}) &= 3(F_{n-1}-F_{n-2})-2F_{n-3} = F_{n-3}. \end{aligned} \tag{74}$$

We summarize this computation.

Theorem 12.

$$s(n_{2t}+2(-1)^t) = F_{2t+2}-F_{2t-2}, \qquad s(n_{2t}-2(-1)^t) = F_{2t+2}-F_{2t-3}. \tag{75}$$

We believe, but have not yet proved, that the second largest value attained by the Stern sequence in $2^r \le n \le 2^{r+1}$ is, in fact, $F_{r+2} - F_{r-3}$, at least for sufficiently large values of r.

Example. One final example was found by computer exploration. Let

$$m_r = \frac{(2^r-1)(2^{r+1}-1)}{3} = \frac{2^{2r+1}+1}{3} - 2^r = n_{2r-1} - 2^r. \tag{76}$$

(Even without the other expression, $m+r$ has to be integral because one of $\{r, r+1\}$ is even, making one of the factors in the numerator a multiple of 3. We wish to show that

$$s(m_{2t}) = 3F_{2t}^2, \qquad s(m_{2t+1}) = F_{2t+2}^2. \tag{77}$$

The proof is just a computation. Recall that $p_n(1^n) = F_{n+1}$ and $p_n(1^{n-1}2) = p_{n+1}(1^{n+1}) = F_{n+2}$.

First, if $r = 2t$ is even, then

$$\begin{aligned} m_{2t} &= n_{4t-1} - 2^{2t} \\ &= 2^{4t-1} + \cdots + 2^{2t+3} + 2^{2t+1} - 2^{2t} + 2^{2t-1} + \cdots + 2^3 + 2^1 + 1 \\ &= 2^{4t-1} + \cdots + +2^{2t+3} + 2^{2t} + 2^{2t-1} + \cdots + 2^3 + 2^1 + 1, \end{aligned} \tag{78}$$

so $[m_{2t}] \sim [1^{2t-3}2^2 1^{2t-3}2]$ and so

$$\begin{aligned} s[m_{2t}] &= p_{4t-3}(1^{2t-3}2^2 1^{2t-3}2) = p_{4t-2}(1^{2t-3}2^2 1^{2t-1}) \\ &= p_{2t-2}(1^{2t-3}2)p_{2t}(21^{2t-1}) + p_{2t-3}(1^{2t-3})p_{2t-1}(1^{2t-1}) \\ &= F_{2t}F_{2t+2} + F_{2t}F_{2t-2} = F_2t(3F_{2t}) = 3F_{2t}^2. \end{aligned} \tag{79}$$

The other case is similar and will be written up in the next batch of notes.

4. Corrections and typos

The first thing I'd like to do is clarify a nasty little point that I tried to avoid earlier. Recall that we were talking about linear recurrences and we assumed that

$$a_n + \sum_{j=1}^{d} c_j a_{n-j} = 0, \qquad n \ge d, \tag{80}$$

where $c_d \ne 0$.

What happens if $c_d = 0$? In the presentation we gave, it messes things up, because the characteristic polynomial $\phi(t)$ has a root at 0, and this means that the reciprocal polynomial $\psi(t)$ fails to have constant term 1. But mathematical presentations are just a subset of mathematical reality!

Suppose

$$a_n + \sum_{j=1}^{d} c_j a_{n-j} = 0, \qquad n \ge d, \tag{81}$$

and $c_k \neq 0$, with $c_j = 0$ for $k+1 \leq j \leq d$. Then, as far as the actual equations go, we have

$$a_n + \sum_{j=1}^{k} c_j a_{n-j} = 0, \qquad n \geq d, \tag{82}$$

In other words, no equation involves a_i for $i < d-k$. To be tedious, if we let $b_n = a_{n+(d-k)}$, then it *is* true that

$$b_n + \sum_{j=1}^{k} c_j b_{n-j} = 0, \qquad n \geq k, \tag{83}$$

and the usual method gives us

$$\begin{aligned} b_n &= \sum_{j=1}^{r} p_j(n) z_j^n \qquad (\text{for } n \geq 0) \\ \implies a_n &= \sum_{j=1}^{r} p_j(n-(d-k)) z_j^{n-d-k} \qquad (\text{for } n \geq d-k), \end{aligned} \tag{84}$$

and a_n is arbitrary for $n < d-k$. Since there exist polynomials $\bar{p}_j$ so that

$$\bar{p}_j(n) = p_j(n-(d-k)) z_j^{-(d-k)}, \tag{85}$$

we are justified in saying that the closed formula is "valid" for $n \geq d-k$. Nineteenth century mathematicians saw that 0^n could be construed as having the value 1 at $n=0$ and 0 for $n>0$. However, $n^k 0^n$ does not take a non-zero value only at $n=k$, so they invented some ugly notations to take care of it. I think it's easier to say that we have a formula for a_n if $n \geq n_0$.

Why does this matter? As was pointed out to me, in the Notes, III (Second supplement), we studied an important sequence $(A_t(r))$ for which $A_t(r+1) = A_t(r)$ for $r \geq 1$, I glossed over this issue in my original discussion, and if you'll forgive some mixed notations, what's really going on is that $(A_t(r))$ satisfies the linear recurrence:

$$a_n + (-1)\cdot a_{n-1} + 0 \cdot a_{n-2} = 0, \qquad n \geq 2, \tag{86}$$

What this means is simply that $A_t(1) = A_t(2) = A_t(3) = \cdots$, with no information about $A_t(0)$. That's all, and that's why the proof of Theorem 2 on p.3 is so awkward. We can't go from $\Delta(m)$ to $\Delta(2m)$, but we *can* say that $\Delta(2m) = \Delta(4m)$.

Finally, a few egregious errors from the second installment of Notes, IV. (What I get from trying to write things up an hour before class.) I won't bother with un-closed parentheses and the like, which are annoying but don't affect the meaning.

- p.9 In the statement of Theorem 11, the condition should be $n < 2^{r_0+1}$ not $n < 2^{r_0} + 1$.

• p.11 Theorem 12 is somewhat garbled: Equation (75) should read (forgive the labels)

$$s(n_r + 2(-1)^r) = F_{r+2} - F_{r-2} \qquad s(n_r - 2(-1)^r) = F_{r+2} - F_{r-3}. \tag{87}$$

5. Some specific examples, continued and corrected

We now present the final example in one whole.

Example. One final example was found by computer exploration. Let

$$m_r = \frac{(2^r - 1)(2^{r+1} - 1)}{3} = \frac{2^{2r+1} + 1}{3} - 2^r = n_{2r-1} - 2^r. \tag{88}$$

(Even without the other expression, m_r has to be integral because one of $\{r, r+1\}$ is even, making one of the factors in the numerator a multiple of 3.) We wish to show that

$$s(m_{2t}) = 3F_{2t}^2, \qquad s(m_{2t+1}) = F_{2t+2}^2. \tag{89}$$

The proof is just a computation. Recall from earlier notes that $p_n(1^n) = F_{n+1}$, $p_n(1^{n-1}2) = p_{n+1}(1^{n+1}) = F_{n+2}$, $p_n(1^{n-1}3) = 3F_n + F_{n-1}$, and $F_{n+2} + F_{n-2} = 2F_n + F_{n-1} + F_{n-2} = 3F_n$.

First, if $r = 2t$ is even, then

$$\begin{aligned} m_{2t} &= n_{4t-1} - 2^{2t} \\ &= 2^{4t-1} + \cdots + 2^{2t+3} + 2^{2t+1} - 2^{2t} + 2^{2t-1} + \cdots + 2^3 + 2^1 + 1 \\ &= 2^{4t-1} + \cdots + +2^{2t+3} + 2^{2t} + 2^{2t-1} + \cdots + 2^3 + 2^1 + 1, \end{aligned} \tag{90}$$

so $[m_{2t}] \sim [1^{2t-3}2^2 1^{2t-3}2]$ (binary "10101" becomes "10011") and so

$$\begin{aligned} s[m_{2t}] &= p_{4t-3}(1^{2t-3}2^2 1^{2t-3}2) = p_{4t-2}(1^{2t-3}2^2 1^{2t-1}) \\ &= p_{2t-2}(1^{2t-3}2)p_{2t}(21^{2t-1}) + p_{2t-3}(1^{2t-3})p_{2t-1}(1^{2t-1}) \\ &= F_{2t}F_{2t+2} + F_{2t}F_{2t-2} = F_{2t}(3F_{2t}) = 3F_{2t}^2. \end{aligned} \tag{91}$$

If $r = 2t + 1$ is even, then

$$\begin{aligned} m_{2t+1} &= n_{4t+1} - 2^{2t+1} \\ &= 2^{4t+1} + \cdots + 2^{2t+3} + 2^{2t-1} + \cdots + 2^3 + 2^1 + 1, \end{aligned} \tag{92}$$

so $[m_{2t+1}] \sim [1^{2t-1}31^{2t-2}2]$ (binary "10101" becomes "10001"), and so

$$\begin{aligned} s[m_{2t+1}] &= p_{4t-1}(1^{2t-1}31^{2t-2}2) = p_{4t}(1^{2t-1}31^{2t}) \\ &= p_{2t}(1^{2t-1}3)p_{2t}(1^{2t}) + p_{2t-1}(1^{2t-1})^2 = (3F_{2t} + F_{2t-1})F_{2t+1} + F_{2t}^2 \\ &= (F_{2t+2} + F_{2t})(F_{2t+2} - F_{2t}) + F_{2t}^2 = F_{2t+2}^2. \end{aligned} \tag{93}$$

Less interesting computations give $s(n_r \pm 2^j)$ for other values of j.

Since $2^{2r-1} < m_r, n_{2r-1} < 2^{2r}$, it is interesting to compare $s[m_r]$ with $s[n_{2r-1}]$, the maximum value of $s[n]$ in that range. A routine computation, which we omit, shows that

$$\begin{aligned} \lim_{r\to\infty} \frac{s[m_{2t}]}{s[n_{4t-1}]} &= \lim_{r\to\infty} \frac{3F_{2t}^2}{F_{4t+1}} = \frac{3(5-\sqrt{5})}{10}; \\ \lim_{r\to\infty} \frac{s[m_{2t+1}]}{s[n_{4t+1}]} &= \lim_{r\to\infty} \frac{F_{2t+2}^2}{F_{4t+3}} = \frac{(5+\sqrt{5})}{10}. \end{aligned} \tag{94}$$

6. The rule of four revisited, revisited

We complete our discussion with a look at a few more continued fractions. Recall that

$$\begin{aligned} \frac{s(n)}{s(n+1)} &= a_{2v+1} + \cfrac{1}{a_{2v} + \cfrac{1}{\cdots + \cfrac{1}{a_1}}} = \frac{p_{2v+1}(a_1,\ldots,a_{2v+1})}{p_{2v}(a_1,\ldots,a_{2v})}\ , \\ \frac{s(\overleftarrow{n})}{s(\overleftarrow{n}+1)} &= a_1 + \cfrac{1}{a_2 + \cfrac{1}{\cdots + \cfrac{1}{a_{2v+1}}}} = \frac{p_{2v+1}(a_1,\ldots,a_{2v+1})}{p_{2v}(a_2,\ldots,a_{2v+1})}\ . \end{aligned} \tag{95}$$

For completeness sake, we consider the other two fractions. As before, suppose that $s(n) = m$ and $s(n+1) = a$, then $s(n') = m$ and $s(n'+1) = m - a$ and since

$$\frac{m}{m-a} = 1 + \frac{a}{m-a} = 1 + \cfrac{1}{\cfrac{m-a}{a}} = 1 + \cfrac{1}{\cfrac{m}{a} - 1}, \tag{96}$$

we have, formally,

$$\frac{s(n')}{s(n'+1)} = 1 + \cfrac{1}{a_{2v+1} - 1 + \cfrac{1}{a_{2v} + \cfrac{1}{\cdots + \cfrac{1}{a_1}}}}\ . \tag{97}$$

This is very familiar, but we must be alert to two cases. If $a_{2v+1} > 1$ (that is, iff $\frac{m}{a} \geq 2$ iff $a < m - a$), then this is a genuine continued fraction representation. However, if

$a_{2v+1} = 1$, then the expression simplifies:

$$\frac{s(n')}{s(n'+1)} = 1 + \cfrac{1}{0 + \cfrac{1}{a_{2v} + \cfrac{1}{\cdots + \cfrac{1}{a_1}}}} = 1 + a_{2v} + \cfrac{1}{a_{2v-1} + \cfrac{1}{\cdots + \cfrac{1}{a_1}}}. \tag{98}$$

This should look familiar. The same sort of thing happens for $s(\overleftarrow{n'})$ and $s(\overleftarrow{n'}+1)$, and because of the importance (see Thm. 11 in Notes, IV), we write it out: if $a_1 > 1$, then

$$\frac{s(\overleftarrow{n'})}{s(\overleftarrow{n'}+1)} = 1 + \cfrac{1}{a_1 - 1 + \cfrac{1}{a_2 + \cfrac{1}{\cdots + \cfrac{1}{a_{2v+1}}}}}, \tag{99}$$

and if $a_1 = 1$, then

$$\frac{s(\overleftarrow{n'})}{s(\overleftarrow{n'}+1)} = 1 + a_2 + \cfrac{1}{a_{23} + \cfrac{1}{\cdots + \cfrac{1}{a_{2v+1}}}}. \tag{100}$$

Finally, we spend a second talking about an obvious unanswered question: what about

$$\frac{p_{2v}(a_1, \dots, a_{2v})}{p_{2v-1}(a_1, \dots, a_{2v-1})} \ ? \tag{101}$$

It's convienient, if inconsistent, to write $\hat{n} \sim [a_1, \dots, a_{2v-1}]$, so that, as before, $n = 2^{a_{2v}+a_{2v+1}}\bar{n} + 2^{2v+1} - 1$. Then,

$$p_{2v}(a_1, \dots, a_{2v}) = s(n+1) = s(2^{a_{2v}+a_{2v+1}}\bar{n} + 2^{2v+1}) = s(2^{a_{2v}}\bar{n} + 1), \tag{102}$$

and since $s(\bar{n}) = p_{2v-1}(a_1, \dots, a_{2v-1})$, we have our answer:

$$\frac{p_{2v}(a_1, \dots, a_{2v})}{p_{2v-1}(a_1, \dots, a_{2v-1})} = \frac{s(2^{a_{2v}}\bar{n} + 1)}{s(2^{a_{2v}}\bar{n})} = \frac{s((2^{a_{2v}}\bar{n} + 1)')}{s(2^{a_{2v}}\bar{n} + 1)' + 1)} \tag{103}$$

Some questions may be better left unasked. Actually, it is an interesting exercise to calculate $[(2^{a_{2v}}\bar{n} + 1)']_2$, which (in this draft at least) we shall leave to the reader. There are, as one might expect, four cases, depending on whether $a_1 = 1$ or $a_1 > 1$ and whether $a_{2v} = 1$ or $a_{2v} > 1$.

7. Polynomials mapping to $\mathbb{Z}$

There is a decided contrast between what we've seen about the growth of $s(f(2^r))$, where $f \in \mathbb{Z}[x]$ (basically polynomial) and the basically exponential growth in $s(n_r)$, even though n_r is basically linear in 2^r. (It is more accurate to say that, because of the "$(-1)^r$" in the definition, n_{2r} and n_{2r+1} are separately quadratic in r, as is m_r.) Of course there are denominators in these case. The intuitive reason is that the base 2 representations of $f(2^r)$ are all the same for sufficiently large r, except for certain blocks whose size is linear in r. The polynomial comes from plugging linear entries into a continuant of fixed index. On the other hand, the base 2 representation of n_r and m_r contain of blocks of fixed size, so that bounded entries are put into a continuant of linearly increasing index.

In order to see what's going on, we make a detour into some very classical results on polynomials. Let $\mathcal{P}_d$ denote the set of real polynomials of degree $\le d$ and let

$$\mathcal{P}_{d,\mathbb{Z}} = \{f \in \mathcal{P}_d : f : \mathbb{Z} \to \mathbb{Z}\}. \tag{104}$$

Certainly, $\mathcal{P}_d \cap \mathbb{Z}[x] \subseteq \mathcal{P}_{d,\mathbb{Z}}$, but the inclusion is not strict; e.g., $\frac{x(x-1)}{2} \in \mathcal{P}_{2,\mathbb{Z}}$. More generally, define $x^{(k)}$ recursively by

$$x^{(0)} = 1, \qquad x^{(k)} = x \cdot (x-1)^{(k-1)} = x^{(k-1)}(x-(k-1)), \quad k \ge 1. \tag{105}$$

It is customary and natural to write

$$\binom{x}{k} = \frac{x^{(k)}}{k!}, \tag{106}$$

since, when $k \le x \in \mathbb{N}$, we recover the usual binomial coefficient:

$$\binom{x}{k} = \frac{x^{(k)}}{k!} = \frac{x(x-1)\cdots(x-(k-1))}{k!} = \frac{x!/(x-k)!}{k!}. \tag{107}$$

It follows from the definition that $\binom{x}{k} = 0$ for $x = 0, \dots, k-1$ and if $x = -y < 0$, then

$$\begin{aligned} \binom{x}{k} = \binom{-y}{k} &= \frac{(-y)(-y-1)\cdots(-y-(k-1))}{k!} \\ = (-1)^k \cdot \frac{(y+k-1)\cdots(y+1)y}{k!} &= (-1)^k \binom{y+k-1}{k}. \end{aligned} \tag{108}$$

It follows that $x \in \mathbb{Z} \implies \binom{x}{k} \in \mathbb{Z}$ and so $\binom{x}{k} \in \mathcal{P}_{d,\mathbb{Z}}$ for $0 \le k \le d$. Indeed, $\{\binom{x}{k}, 0 \le k \le d\}$ is easily seem to be an "upper diagonal" basis for $\mathcal{P}_d$.

One of the standard approaches to understanding polynomials from their values is *Lagrange interpolation.* Fix $x_0 < x_1 \cdots < x_d$. Observe that if $f, g \in \mathcal{P}_d$ and $f(x) = g(x)$ for $x = x_j$, $0 \le j \le d$, then there exists $h \in \mathbb{R}[x]$ so that

$$f(x) - g(x) = h(x) \cdot \left(\prod_{j=0}^{d} (x - x_j) \right); \tag{109}$$

degree considerations imply that $h = 0$, so $f = g$. We now define

$$\phi_j(x_0, \dots, x_d; x) = \phi_j(x) := \prod_{i \neq j} \frac{x - x_i}{x_j - x_i} \in \mathcal{P}_d. \tag{110}$$

Then $\phi_j(x_i) = 0$ for $i \neq j$ and $\phi_j(x_j) = 1$. It follows that

$$L_{d,f}(x) := \sum_{i=0}^{d} f(x_i)\phi_i(x) \in \mathcal{P}_d \tag{111}$$

has the property that $L_{d,f}(x_j) = f(x_j)$ and so, in fact, $f = L_{d,f}$. That is, a polynomial in 1 variable of degree d is completely determined by its value at $d+1$ distinct points. There is no such "clean" criterion for polynomials in more than 1 variable, unfortunately.

Observe that, if we take $x_i = i$, then

$$\begin{aligned} \phi_j(x) &:= \prod_{i \neq j} \frac{x - i}{j - i} = \prod_{i=0}^{j-1} \left(\frac{x - i}{j - i} \right) \prod_{i=j+1}^{d} \left(\frac{x - i}{j - i} \right) \\ &= (-1)^{d-j} \binom{x}{j} \binom{x - (j+1)}{d - j} \in \mathcal{P}_{d,\mathbb{Z}}. \end{aligned} \tag{112}$$

Thus, if $f \in \mathcal{P}_{d,\mathbb{Z}}$, then since $f(i) \in \mathbb{Z}$, we have that f is a $\mathbb{Z}$-linear combination of $\{\phi_0, \dots, \phi_d\}$, and conversely, any such linear combination is in $\mathcal{P}_{d,\mathbb{Z}}$. This characterization is somewhat unsatisfactory, however. If $\deg(f) = k$, then $f \in \mathcal{P}_d$ for every $d \geq k$, yet the representations in terms of Lagrange interpolation are different for each such d, because every polynomial $\phi_j(x_0, \dots, x_d; x$ has exact degree d. If $f \in \mathcal{P}_{d,\mathbb{Z}}$ has actual degree k, we'd prefer that it only be represented in terms of polynomials with degree $\leq k$, so that the representations don't change as we increase d.

One way to deal with this is to look for a different basis. Define the operator Δ by

$$\Delta f(x) := f(x+1) - f(x). \tag{113}$$

Since

$$\Delta 9 x^k) = (x+1)^k - x^k = \sum_{i=0}^{k-1} \binom{k}{i} x^i, \tag{114}$$

$\Delta : \mathcal{P}_d \to \mathcal{P}_{d-1}$. The telescoping sum

$$f(n) - f(m) = \sum_{x=m}^{n-1} f(x+1) - f(x) = \sum_{x=m}^{n-1} \Delta f(x) \tag{115}$$

for $m < n$ implies that $f \in \mathcal{P}_{d,\mathbb{Z}}$ if and only if $f(0) \in \mathbb{Z}$ and $\Delta p \in \mathcal{P}_{d-1,\mathbb{Z}}$. If we define $\Delta^k(f) = \Delta(\Delta^{k-1}(f))$ as usual, we can iterate this result to say that

$$f \in \mathcal{P}_{d,\mathbb{Z}} \iff f(0),\ (\Delta f)(0),\ (\Delta^2 f)(0), \dots, (\Delta^d f)(0) \in \mathbb{Z}. \tag{116}$$

This is most easily visualized by writing down a *difference table* in which the first row is $f(n), n \geq 0$, the second row is $(\Delta f)(n), n \geq 0$, etc, until all 0's appear in the $(d+1)$-st row. For example, if $f(x) = x^3$, the difference table is

$$
\begin{array}{ccccccccccccccc}
\ldots & & 0 & & 1 & & 8 & & 27 & & 64 & & 125 & & \ldots \\
& \ldots & & 1 & & 7 & & 19 & & 37 & & 61 & & \ldots & \\
& & \ldots & & 6 & & 12 & & 18 & & 24 & & \ldots & & \\
& & & \ldots & & 6 & & 6 & & 6 & & \ldots & & & \\
& & & & \ldots & & 0 & & 0 & & \ldots & & & &
\end{array} \tag{117}
$$

We now make a 17th century set of observations. Note that for $k \geq 1$,

$$
\begin{aligned}
\Delta(x^{(k)}) = (x+1)^{(k)} - x^{(k)} = ((x+1) - (x-(k-1)))x^{(k-1)} = kx^{(k-1)} \\
\Longrightarrow \Delta\binom{x}{k} = \binom{x}{k-1}.
\end{aligned} \tag{118}
$$

and $\Delta(x^{(0)}) = 1 - 1 = 0$. Thus,

$$
\begin{aligned}
f(x) = \sum_{k=0}^{d} a_k \binom{x}{k} \Longrightarrow \Delta f(x) = \sum_{k=1}^{d} a_k \binom{x}{k-1} = \sum_{k=0}^{d-1} a_{k+1} \binom{x}{k} \\
\left(\Longrightarrow \Delta^j f(x) = \sum_{k=0}^{d-j} a_{k+j} \binom{x}{k}. \right)
\end{aligned} \tag{119}
$$

Conveniently enough,

$$
x^{(k)}\big|_{x=0} = \begin{cases} 1 & \text{if } k = 0, \\ 0 & \text{if } k > 0. \end{cases} \tag{120}
$$

Thus, if f is given as above, then

$$
a_k = (\Delta^k f)(0), \tag{121}
$$

or, to write it as the traditional *Newton's formula*, if $f \in \mathcal{P}_d$, then

$$
f(x) = \sum_{k=0}^{d} \frac{(\Delta^k f)(0)}{k!} x^{(k)}. \tag{122}
$$

Using x^3 as an example, from the difference table given above, we see that

$$
x^3 = 0 \cdot \binom{x}{0} + 1 \cdot \binom{x}{1} + 6 \cdot \binom{x}{2} + 6 \cdot \binom{x}{3} = x + 3x(x-1) + x(x-1)(x-2), \tag{123}
$$

as may be easily verified. The similarity to Taylor's formula is no accident of course, the differentiation operator D has the same matrix on $\mathcal{P}_d$ with respect to the basis $\{\frac{1}{k!} \cdot x^k : 0 \leq k \leq d\}$ as does Δ on $\mathcal{P}_d$ with respect to the basis $\{\binom{x}{k} : 0 \leq k \leq d\}$.

In particular, we see that $f \in \mathcal{P}_{d,\mathbb{Z}}$ if and only if f is a $\mathbb{Z}$-linear combination of $\{\binom{x}{0},\dots\binom{x}{d}\}$, and if $\deg f = k < d$, then only $\{\binom{x}{0},\dots\binom{x}{k}\}$ is needed, so the representations do not depend on d as it increases.

It is almost obligatory here to observe that since Δ and $\sum$ are inverse operations, we obtain summation formulas at virtually no cost. For $k \ge 0$,

$$\sum_{x=0}^{n}\binom{x}{k} = \sum_{x=0}^{n}\left(\binom{x+1}{k+1} - \binom{x}{k+1}\right) = \binom{n+1}{k+1} - \binom{0}{k+1} = \binom{n+1}{k+1}. \tag{124}$$

In a method that goes back to Bernoulli, this allows us to sum any polynomial, after merely writing down its difference table. For example,

$$\begin{aligned}\sum_{x=0}^{n} x^3 &= 1\cdot\binom{n+1}{2} + 6\cdot\binom{n+1}{3} + 6\cdot\binom{n+1}{4}\\ &= \frac{(n+1)n}{2} + (n+1)n(n-1) + \frac{(n+1)n(n-1)(n-2)}{4}\\ &= \frac{n(n+1)}{4}\left(2+4(n-1)+(n-1)(n-2)\right) = \frac{n^2(n+1)^2}{4}.\end{aligned} \tag{125}$$

But this isn't really what we are interested in! Let

$$\mathcal{P}_{d,2^{\mathbb{N}}} = \{f \in \mathcal{P}_d : f(2^r) \in \mathbb{Z}, \quad \text{for sufficiently large } r\} \tag{126}$$

Clearly, if $f \in \mathcal{P}_{d,\mathbb{Z}}$, then $f \in \mathcal{P}_{d,2^{\mathbb{N}}}$, and if $f \in \mathcal{P}_{d,2^{\mathbb{N}}}$, then $2^{-m}f \in \mathcal{P}_{d,2^{\mathbb{N}}}$. Further, we can always replace $f(x)$ by $f(2^{r_0}x)$ to assume, without loss of generality, that $f(2^r) \in \mathbb{Z}$ for $r \ge 0$. Also notice that $\mathcal{P}_{d,2^{\mathbb{N}}}$ is closed under addition and multiplication (when degrees are adjusted.)

In the rest of this section, we assume that $f \in \mathcal{P}_d$ and $Mf \in \mathbb{Z}[x]$ for an *odd* denominator M, and M is minimal with this property, so that the gcd of the coefficients of Mf is a power of 2; without loss of generality, we can take this gcd to be 1.

First, suppose $d = 1$, $f \in \mathcal{P}_{1,2^{\mathbb{N}}}$ and $Mf(x) = g(x) = a_1x + a_0$. We derive a contradiction from $M > 1$. Suppose otherwise, and let p be a prime factor of M. Then M (and so p) will divide both $g(2) = Mf(2) = 2a_1 + a_0$ and $g(1) = Mf(1) = a_1 + a_0$, and so p divides both

$$2a_1 + a_0 - (a_1 + a_0) = a_1 \qquad \text{and} \qquad -(2a_1 + a_0) + 2(a_1 + a_0) = a_0. \tag{127}$$

That is, $(M/p)f \in \mathbb{Z}[x]$. This contradicts the minimality of M, so $\mathcal{P}_{1,2^{\mathbb{N}}}$ consists of the linear polynomials in $\mathbb{Z}[x]$.

We have seen that $f(x) = \binom{x}{k} \in \mathcal{P}_{d,\mathbb{Z}} \subseteq \mathcal{P}_{d,2^{\mathbb{N}}}$, but the odd part of the denominator may not be large. In fact, up to powers of 2, $\binom{x}{2} \in \mathbb{Z}[x]$. On the other hand, we have already seen that $\frac{x^2-1}{3}, \frac{2x^2+1}{3} \in \mathcal{P}_{2,2^{\mathbb{N}}}$. We now show that these are essentially the only cases.

Suppose $f \in \mathcal{P}_{2,2^{\mathbb{N}}}$, and for $r \ge 0$, we have $f(2^r) \in \mathbb{Z}$. Suppose $g(x) = Mf(x) = a_2x^2+a_1x+a_0$, where M is again minimal and suppose p^k is a prime power factor of M.

We first show that $p = 3$ and then that $k = 1$. By hypothesis, $f(1), f(2), f(4) \in \mathbb{Z}$, so that

$$g(x) := a_2x^2 + a_1x + a_0 \equiv 0 \pmod{p}, \qquad \text{for } x = 1, 2, 4. \tag{128}$$

Recall that $\mathbb{Z}/p\mathbb{Z}$ is a field, and so g is identically zero if it has three distinct zeros. Thus, if $p > 3$, then each a_i is a multiple of p, contradicting the minimality of M. Now suppose $M = 3^k$ with $k \ge 2$, so $3^2 \mid M$. We have to be a bit careful, because $\mathbb{Z}/9\mathbb{Z}$ is not a field. (Indeed, the polynomial $3(x^2-1)$ has 6 zeros in $\mathbb{Z}/9\mathbb{Z}$.) However, assuming

$$g(x) = a_2x^2 + a_1x + a_0 \equiv 0 \pmod{9}, \qquad \text{for } x = 1, 2, 4. \tag{129}$$

we note that

$$\begin{aligned} h(4) - 3h(2) + 2h(1) &= 6a_2 \equiv 0 \pmod{9}, \\ -h(4) + 5h(2) - 4h(1) &= 2a_1 \equiv 0 \pmod{9}, \\ h(4) - 6h(2) + 8h(1) &= 3a_0 \equiv 0 \pmod{9}. \end{aligned} \tag{130}$$

Thus a_0, a_1 and a_2 are all multiples of 3, violating minimality once again. In the end, $M = 3$, and we have $a_2x^2 + a_1x + a_0 \equiv a_2(x-1)(x-2) \pmod{3}$; that is, $a_1 \equiv 0$ and $a_0 \equiv -a_2$. It follows that

$$f(x) = a_2 \cdot \frac{x^2-1}{3} + q(x), \qquad q \in \mathbb{Z}[x]. \tag{131}$$

Thus, $\mathcal{P}_{2,2^{\mathbb{N}}}/(\mathbb{Z}[x]) = \{0, \frac{x^2-1}{3}, \frac{2x^2+1}{3}\}$.

This situation will clearly get messier as d increases. For example, if $q \in \mathbb{Z}[x] \cap \mathcal{P}_{d-2}$, then $\frac{x^2-1}{3}q \in \mathcal{P}_{d,2^{\mathbb{N}}}$. On the other hand, a similar argument to that given above shows that if prime $p \mid M$ for $p \in \mathcal{P}_{d,2^{\mathbb{N}}}$, then $ord_p(2) \le d$, hence

$$p \mid M(d) := \prod_{k=1}^{d} (2^k - 1). \tag{132}$$

We now present $\Lambda_d \in \mathcal{P}_{d,2^{\mathbb{N}}}$ which has $M(d)$ as its denominator. We do not yet claim that $\nu_p(M) \le \nu_p(M_d)$ for all $f \in \mathcal{P}_{d,2^{\mathbb{N}}}$, but that is a reasonable conjecture. We need a lemma of independent interest, which is surely known in cases where the variable is "q", rather than "x".

Lemma 13. *For $d, r \in \mathbb{N}$, let*

$$F_{r,d}(x) = \frac{\prod_{i=1}^{d} (x^{r+i} - 1)}{\prod_{i=1}^{d} (x^i - 1)} \in \mathbb{Z}[x]. \tag{133}$$

Then $F_{r,d}(x) \in \mathbb{Z}[x]$.

Proof. Recall that for $e, n \in \mathbb{N}$,

$$\left\lfloor \frac{n}{e} \right\rfloor - \left\lfloor \frac{n-1}{e} \right\rfloor = \begin{cases} 1 & \text{if } e \mid n, \\ 0 & \text{otherwise.} \end{cases} \tag{134}$$

Recall also that the cyclotomic polynomials $\Phi_e(x) \in \mathbb{Z}[x]$ are irreducible and have the property that for every $n \in \mathbb{N}$,

$$x^n - 1 = \prod_{e \mid n} \Phi_e(x) = \prod_{e=1}^{\infty} \Phi_e^{\lfloor \frac{n}{e} \rfloor - \lfloor \frac{n-1}{e} \rfloor}(x) \tag{135}$$

The only factors in the denominator of $F_{r,d}(x)$ are powers of $\Phi_e(x)$ for $e \le d$, and the net exponent in the quotient is

$$\begin{aligned} &\sum_{i=1}^{d} \left(\left\lfloor \frac{r+i}{e} \right\rfloor - \left\lfloor \frac{r+i-1}{e} \right\rfloor \right) - \sum_{i=1}^{d} \left(\left\lfloor \frac{i}{e} \right\rfloor - \left\lfloor \frac{i-1}{e} \right\rfloor \right) \\ &= \left\lfloor \frac{r+d}{e} \right\rfloor - \left\lfloor \frac{r}{e} \right\rfloor - \left\lfloor \frac{d}{e} \right\rfloor \ge 0. \end{aligned} \tag{136}$$

□

It follows from this lemma that

$$F_{r,d}(2) = \frac{\prod_{i=1}^{d} (2^{r+i} - 1)}{\prod_{i=1}^{d} (2^i - 1)} \in \mathbb{Z}. \tag{137}$$

for all d and $r \ge 0$, and so

$$\Lambda_d(x) := \prod_{i=1}^{d} \frac{2^i x - 1}{2^i - 1} \in \mathcal{P}_{d,2^{\mathbb{N}}}. \tag{138}$$

Observe that the denominator in Λ_d is our friend $M(d)$. Let

$$\bar{\Lambda}_d(x) = \prod_{i=1}^{d} \frac{x - 2^{i-1}}{2^i - 1} \in \mathcal{P}_{d,2^{\mathbb{N}}}. \tag{139}$$

Observe that $\bar{\Lambda}_d(2^r) = 0$ for $r = 0, 1, \dots, d-1$ and that

$$\bar{\Lambda}_d(2^r) = \prod_{i=1}^{d} \frac{2^r - 2^{i-1}}{2^i - 1} = \prod_{i=1}^{d} \frac{2^{i-1}(2^{r-i+1} - 1)}{2^i - 1} = 2^{\binom{d}{2}} \Lambda_d(2^{r-d}). \tag{140}$$

Thus, $\bar{\Lambda}_d \in \mathcal{P}_{d,2^{\mathbb{N}}}$ as well, and $s(\bar{\Lambda}_d(2^r)) = s(\Lambda_d(2^{r-d})$. Finally, note that

$$\bar{\Lambda}_2(2^r) = \frac{(2^r - 1)(2^{r-1} - 1)}{(2-1)(4-1)} = m_{r-1}. \tag{141}$$

Although there is a nice pattern to $s(m_r)$, if we let

$$\ell_r = \bar{\Lambda}_3(2^r) = \frac{(2^r-1)(2^{r-1}-1)(2^{r-2}-1}{(2-1)(4-1)(8-1)}, \tag{142}$$

then the pattern for $a_r = s(\ell_r)$ is less clear, though it begins promisingly enough:

$$a_3 = 1, \quad a_4 = 4, \quad a_5 = 27, \quad a_6 - 100, \quad a_7 = 256, \quad a_8 = 484. \tag{143}$$

That is, 1, followed by four squares and one cube. Then $a_9 = 1157$, which doesn't look like much, but factors as $13 \cdot 89$; both are Fibonacci numbers. A check for $r \leq 30$ yields nothing else of interest, except that 2^5 divides a_r for $r = 12, 13, 14, 24, 25, 26, 29$, which seems to be rather frequently. A few interesting factorizations are:

$$\begin{aligned} a_{12} &= 2^6 \cdot 13 \cdot 167, \quad a_{13} = 2^5 \cdot 5^2 \cdot 11 \cdot 29, \\ a_{14} &= 2^6 \cdot 11 \cdot 31^2, \quad a_{19} = 2^4 \cdot 3^4 \cdot 11^2 \cdot 1579. \end{aligned} \tag{144}$$

(If 1579 sounds familiar, it may be that $F_{17} = 1597$. D-oh.)