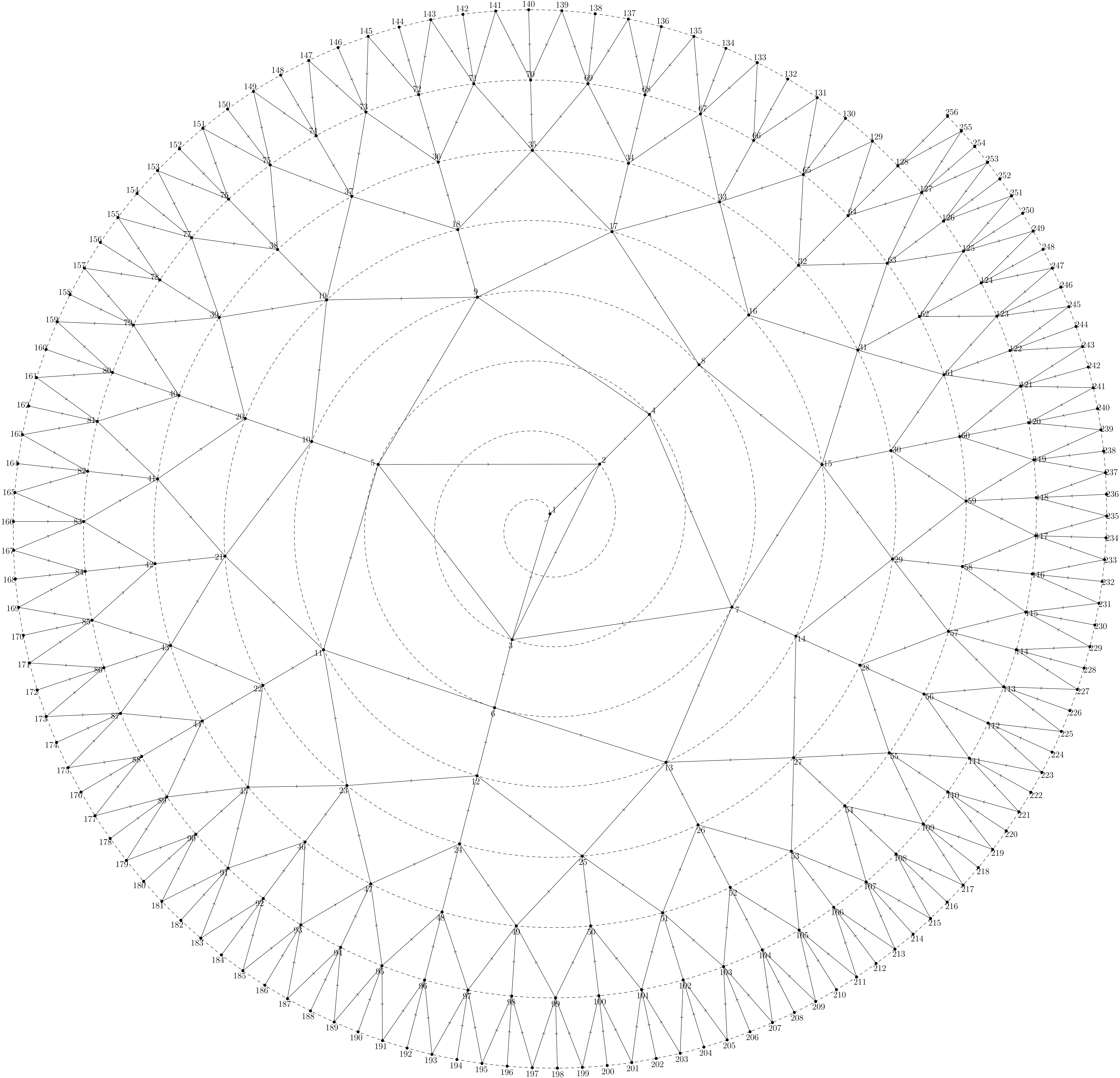
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256

STERN NOTES, MATH 595, SPRING 2012

BRUCE REZNICK, UIUC

1. Introduction and vague overview

1.1. **Organization.** These notes are about the *Stern sequence*, defined by

$$s(0) = 0, s(1) = 1; \qquad s(2n) = s(n), \quad s(2n+1) = s(n) + s(n+1) \text{ for } n \geq 1. \tag{1.1}$$

This first chapter will introduce the subject of the Stern sequence and related mathematical objects, and give some theorems about them which can be proved directly and quickly. Each subsequent chapter will present a useful mathematical perspective, and then apply it to the topics of the previous chapters. Chapter two is about generating functions and formal power series. Chapter three covers constant-coefficient recurrences and their implications. Chapter four presents some of the classical facts about finite simple continued fractions. Chapter five is devoted to Minkowski's ?-function and related questions in real analysis. Chapters six is about digital representation questions. Chapter seven is about the behavior of $s(n) \pmod{d}$. Chapters eight (and possibly nine) will be about applications and special questions, and I'm not completely sure yet what is likely to show up. I'm taking suggestions.

1.2. **The array and the sequence.** Let's start at the beginning. Stern considered the *diatomic array*, which resembles a Pascal's triangle with memory:

$$\begin{array}{ccccccccc}
 & & & & a \quad b & & & & \\
 & & & a & a+b & b & & & \\
 & & a & 2a+b & a+b & a+2b & b & & \\
a & 3a+b & 2a+b & 3a+2b & a+b & 2a+3b & a+2b & a+3b & b \\
 & & & & \cdots & & & &
\end{array} \tag{1.2}$$

$$\boxed{\boxed{Z(a,b)}}$$

In words, we start with "a, b", and repeat each row in the following row, with the sums of consecutive terms inserted between them. The number of terms in the rows is 2, 3, 5, 9, etc. and is evidently one more than a power of two. So the rows in (1.2) are indexed by r, starting with $r = 0$ and having entries indexed from 0 to 2^r.

The formal definition of $Z(r, k; a, b)$ for $r \geq 0$ and $0 \leq k \leq 2^r$ is given by:

$$
\begin{aligned}
Z(0, 0; a, b) &= a, \\
Z(0, 1; a, b) &= b; \\
Z(r, 2k; a, b) &= Z(r-1, k; a, b), \text{ for } r \geq 1; \\
Z(r, 2k+1; a, b) &= Z(r-1, k; a, b) + Z(r-1, k+1; a, b), \text{ for } r \geq 1.
\end{aligned} \tag{1.3}
$$

A few things are evident, and easily proved, from looking at (1.2) and (1.3). First, each entry $Z(r, k; a, b)$ is linear in (a, b) and second, the table has a left-right symmetry if a and b are swapped:

$$Z(r, 2^r - k; a, b) = Z(r, k; b, a). \tag{1.4}$$

Application of these two observations allows us to reduce the study of (1.2) to that of a single diatomic array, $Z(r, k; 0, 1)$:

$$Z(r, k; a, b) = Z(r, k; 1, 0)a + Z(r, k; 0, 1)b = Z(r, 2^r - k; 0, 1)a + Z(r, k; 0, 1)b. \tag{1.5}$$

$$
\begin{array}{c}
0 \quad 1 \\
0 \quad 1 \quad 1 \\
0 \quad 1 \quad 1 \quad 2 \quad 1 \\
0 \quad 1 \quad 1 \quad 2 \quad 1 \quad 3 \quad 2 \quad 3 \quad 1 \\
0 \quad 1 \quad 1 \quad 2 \quad 1 \quad 3 \quad 2 \quad 3 \quad 1 \quad 4 \quad 3 \quad 5 \quad 2 \quad 5 \quad 3 \quad 4 \quad 1 \\
\cdots \\
\boxed{\boxed{Z(0,1)}}
\end{array} \tag{1.6}
$$

Each row of $Z(0, 1)$ appears to be repeated as the first half of the next row. This is easily proved by induction and (1.3).

Theorem 1.1. *If $0 \leq k \leq 2^r$, then $Z(r+1, k; 0, 1) = Z(r, k; 0, 1) = s(k)$.*

Proof. Write $Z(r, k; 0, 1) = Z(r, k)$ for short. By inspection, $Z(1, k) = Z(0, k)$ for $k = 0, 1$. Suppose $Z(r_0, k) = Z(r_0 - 1, k)$ for $0 \leq k \leq 2^{r_0 - 1}$. If $k = 2s$ is even and $2s \leq 2^{r_0}$, then $s \leq 2^{r_0 - 1}$, and by the induction hypothesis and (1.3), $Z(r_0, 2s) = Z(r_0 - 1, s) = Z(r_0, s) = Z(r_0 + 1, 2s)$. A similar argument holds if $k = 2s + 1$ is odd, taking care to check that $2s + 1 \leq 2^{r_0}$ implies that $s + 1 \leq 2^{r_0 - 1}$.

Finally, $Z(r, k) = s(k)$ for small k by inspection, and by (1.1) and (1.3), the two sequences have the same recurrence, so they are always equal. □

(As we shall see in these notes, one must be careful to treat the base cases carefully, since the inductive step is consistent with $Z(r+1, k; a, b) = Z(r, k; a, b)$ for all (a, b), which isn't true.)

It follows from Theorem 1.1 and (1.5) that

$$Z(r,k;a,b) = Z(r,2^r-k;0,1)a + Z(r,k;0,1)b = s(2^r-k)a + s(k)b. \tag{1.7}$$

Another property of $Z(a,b)$ is a kind of self-similarity: any two consecutive elements in a given row can be thought of as the initial conditions of another diatomic array which is generated as a subset of the subsequent rows, contained between the subsequent appearances of the original pair. More formally, if $0 \le r,t$ and $0 \le k \le 2^r, 0 \le n \le 2^t - 1$, then

$$Z(r,k;Z(t,n;a,b),Z(t,n+1;a,b)) = Z(r+t,2^r n+k;a,b). \tag{1.8}$$

(Again, this can be proved by a simple and omitted induction.)

In particular, if $(a,b) = (0,1)$ and $t = \lceil \log_2(n+1) \rceil$ in (1.8), we find that

$$Z(r,k;s(n),s(n+1)) = s(2^r n+k). \tag{1.9}$$

The r-th row below is simply $s(2^r n), s(2^r n+1) \cdots , s(2^r(n+1))$.

$$\begin{array}{ccccccccc} & & & s(n) & & s(n+1) & & & \\ & & s(n) & & s(n)+s(n+1) & & s(n+1) & & \\ s(n) & & 2s(n)+s(n+1) & & s(n)+s(n+1) & & s(n)+2s(n+1) & & s(n+1) \\ & & & & \cdots & & & & \\ & & & & \boxed{\boxed{Z(s(n),s(n+1))}} & & & & \end{array} \tag{1.10}$$

Take $n = 1$ in (1.10) and note that $s(1) = s(2) = 1$ so the r-th row consists of the elements $s(2^r), s(2^r+1), \ldots, s(2^{r+1})$. In this way, we can quickly write $s(n)$ for $1 \le n \le 64$:

$$\begin{array}{c} 1 \quad 1 \\ 1 \quad 2 \quad 1 \\ 1 \quad 3 \quad 2 \quad 3 \quad 1 \\ 1 \quad 4 \quad 3 \quad 5 \quad 2 \quad 5 \quad 3 \quad 4 \quad 1 \\ 1 \quad 5 \quad 4 \quad 7 \quad 3 \quad 8 \quad 5 \quad 7 \quad 2 \quad 7 \quad 5 \quad 8 \quad 3 \quad 7 \quad 4 \quad 5 \quad 1 \\ 1\ 6\ 5\ 9\ 4\ 11\ 7\ 10\ 3\ 11\ 8\ 13\ 5\ 12\ 7\ 9\ 2\ 9\ 7\ 12\ 5\ 13\ 8\ 11\ 3\ 10\ 7\ 11\ 4\ 9\ 5\ 6\ 1 \\ \cdots \\ \boxed{\boxed{Z(1,1)}} \end{array} \tag{1.11}$$

Note for later reference that, even though $s(2^r)$ appears twice above, as the last entry in the $(r-1)$-st row and the first entry in the r-th row, each *pair* of consecutive entries, $(s(n), s(n+1))$, occurs exactly once for $n \geq 1$: (1,1),(1,2),(2,1),(1,3),(3,2), etc.

By linearity, (1.9) and (1.5) imply that

$$s(2^r n + k) = s(2^r - k)s(n) + s(k)s(n+1) \quad \text{for } 0 \leq k \leq 2^r. \tag{1.12}$$

This can also be proved directly by induction on r, by considering the parity of k and iterating (1.3). By making the substitutions $(n, k) \mapsto (n-1, 2^r - k)$ in (1.12), we obtain a convenient generalization.

$$s(2^r n \pm k) = s(2^r - k)s(n) + s(k)s(n \pm 1) \quad \text{for } 0 \leq k \leq 2^r. \tag{1.13}$$

Easy inductions imply that

$$s(2^r) = 1, \qquad s(2^r - 1) = r, \tag{1.14}$$

and so (1.13) implies that

$$s(2^r n \pm 1) = rs(n) + s(n \pm 1). \tag{1.15}$$

For example,

$$\begin{aligned} & s(2012) \\ & = s(1006) \\ & = s(503) \\ & = s(251) + s(252) \\ & = s(125) + 2s(126) \\ & = s(62) + 3s(63) \\ & = 4s(31) + 3s(32) \\ & = 4s(15) + 7s(16) \\ & = 4s(7) + 11s(8) \\ & = 4s(3) + 15s(4) \\ & = 4s(1) + 19s(2) \\ & = 23s(1) \\ & = 23. \end{aligned} \tag{1.16}$$

Notice that the first entry of the argument of s in the r-th row of (1.16) is $\lfloor \frac{2012}{2^r} \rfloor$, and the pattern of whether this argument is even or odd reflects, in reverse order, the binary digits of 2012. See Theorem 1.6 below. Of course, one may stop the computation at any point where $s(n)$ and $s(n+1)$ are known, and use (1.13) to derive these more quickly. For example, $2012 = 64 * 31 + 28$, so $s(2012) = s(31)s(64-28) + s(32)s(28)$, and $s(36) = 4, s(28) = 3$, $s(31) = 5$, $s(32) = 1$ imply that $s(2012) = 5 * 4 + 1 * 3 = 23$.

These recurrences also can be used more abstractly. The Stern sequence seems to have many unexpectedly nice properties. For example,

$$s((2^r-1)^2) = s(2^{2r} - 2^{r+1} + 1) = s(2^{r+1}(2^{r-1}-1)+1) = \\ (r+1)s(2^{r-1}-1) + s(2^{r-1}) = (r+1)(r-1)+1 = r^2 = (s(2^r-1))^2. \tag{1.17}$$

(A check up to 2^{20} shows that $s(n^2) = (s(n))^2$ otherwise for odd n only for $n = 27, 267, 7807$, with no particular pattern evident in the exceptions. This might also be a good place to note that $s(3^n) = 2^n$ for $n = 1, 2, 3, 6$, and for no other $n \le 165$, at least.)

1.3. **Some properties of the rows.** We now turn to some of the properties of the rows of (1.11). Let

$$I_r = \{2^r, 2^r+1, \dots, 2^{r+1}\}, \tag{1.18}$$

so the r-th row of $Z(1,1)$ consists of $s(n)$ for $n \in I_r$. For $n = 2^r + k \in I_r$, let $n^* = 3 \cdot 2^r - n = 2^{r+1} - k$ denote the reflection of n in I_r. By (1.12),

$$s(n) = s(n^*) = s(k) + s(2^r - k), \qquad s(n^*+1) = s(n-1). \tag{1.19}$$

It is clear that $s(n) \in \mathbb{N}$ and that for $n \ge 1$, $s(2n) < s(2n+1) > s(2n+2)$, so the growth of $s(n)$ will be irregular. Let

$$M_r := \max\{s(n) \ : \ n \in I_r\}. \tag{1.20}$$

An inspection of (1.11) shows that

$$\begin{aligned} M_0 &= s(1) = 1, \\ M_1 &= s(3) = 2, \\ M_2 &= s(5) = s(7) = 3, \\ M_3 &= s(11) = s(13) = 5, \\ M_4 &= s(21) = s(27) = 8, \\ M_5 &= s(43) = s(53) = 13. \end{aligned} \tag{1.21}$$

Let (F_m) denote the usual Fibonacci sequence, defined by $F_0 = 0, F_1 = 1$ and $F_m = F_{m-1} + F_{m-2}$, for $m \ge 2$ and let

$$n_r = \frac{2^{r+2} - (-1)^r}{3} = \frac{4}{3} \cdot 2^r - \frac{(-1)^r}{3}; \quad n_r^* = \frac{5}{3} \cdot 2^r + \frac{(-1)^r}{3}. \tag{1.22}$$

These are the integers closest to $\frac{4}{3} \cdot 2^r$ and $\frac{5}{3} \cdot 2^r$, and effectively trisect I_r.

Theorem 1.2.

$$M_r = s(n_r) = s(n_r^*) = F_{r+2}. \tag{1.23}$$

Proof. The theorem is valid by inspection if $r \le 5$. Suppose $n \in I_r$. If $n = 2k$, then $k \in I_{r-1}$, so $s(n) \le M_{r-1}$. If n is odd, then $n = 4k \pm 1$, and $2k \pm 1 \in I_{r-1}$ and $k \in I_{r-2}$. Thus, $s(n) = s(2k) + s(2k \pm 1) = s(k) + s(2k \pm 1)$, so $s(n) \le M_{r-2} + M_{r-1}$. These arguments imply that $M_r \le M_{r-1} + M_{r-2}$, and, based on the initial conditions, that $M_r \le F_{r+2}$ for all r.

On the other hand, $n_r = 2n_{r-1} - (-1)^r$ and $n_{r-1} - (-1)^r = 2n_{r-2}$, hence

$$s(n_r) = s(n_{r-1}) + s(n_{r-1} - (-1)^r) = s(n_{r-1}) + s(n_{r-2}). \tag{1.24}$$

Since $s(n_r) = F_{r+2}$ for $0 \le r \le 5$, (1.23) follows by induction. As a final remark, if $n \in I_r$ and $s(n) = M_r$, then the argument of the first paragraph implies that $n = 4k \pm 1$, $s(2k \pm 1) = M_{r-1}$ and $s(k) = M_{r-2}$, so these are the only values where the maximum occurs in each row. □

The Binet formula for the Fibonacci numbers states that

$$F_n = \frac{1}{\sqrt{5}}\left(\phi^n - \bar{\phi}^n\right), \qquad \text{where} \quad \phi = \frac{1+\sqrt{5}}{2}, \quad \bar{\phi} = \frac{1-\sqrt{5}}{2}. \tag{1.25}$$

Say that $f(n) = \Theta(g(n))$ if there are positive constants c_j so that $c_1 g(n) \le f(n) \le c_2 g(n)$. If $n \in I_r$, then $\log_2 n - 1 \le r \le \log_2 n$, so $r = \log_2 n + \mathcal{O}(1)$ and

$$\phi^r = e^{r \log \phi} = \Theta(e^{\frac{\log n}{\log 2} \cdot \log \phi}), \tag{1.26}$$

so

$$M_r = \frac{\phi^2}{\sqrt{5}} \cdot \phi^r + o(1) = \Theta(n^\alpha) = o(n), \quad \text{where } \alpha = \frac{\log \phi}{\log 2} \approx .69424. \tag{1.27}$$

By contrast to the irregular growth of $s(n)$, its summatory function is very well-behaved. It is convenient to use a variant notation. For integers $a \le b$, let

$$\begin{aligned} \sum_{n=a}^{b}{}^{*} f(n) &= \sum_{n=a}^{b} f(n) - \tfrac{1}{2}\left(f(a) + f(b)\right) \\ &= \tfrac{1}{2} f(a) + f(a+1) + \cdots + f(b-1) + \tfrac{1}{2} f(b). \end{aligned} \tag{1.28}$$

This is familiar as the trapezoidal estimate to the integral of $f(x)$ from a to b, and has some useful properties

$$\sum_{n=a}^{b}{}^{*} f(n) + \sum_{n=b}^{c}{}^{*} f(n) = \sum_{n=a}^{c}{}^{*} f(n); \qquad \sum_{n=a}^{a}{}^{*} f(n) = 0. \tag{1.29}$$

It will be particularly useful to consider sequences such as

$$\left(\Phi(F(x); r)\right), \quad \text{where } \Phi(F(x); r) := \sum_{n \in I_r}{}^{*} F(s(n)),$$

because $s(2^r) = s(2^{r+1})$ and each $s(2^r)$ would otherwise be counted twice.

Let

$$S(N) := \sum_{n=0}^{N} s(n); \quad S^*(n) := \sum_{n=0}^{N}{}^{*} s(n) = S(N) - \tfrac{1}{2}s(N). \tag{1.30}$$

The following lemma illustrates the utility of this notation.

Lemma 1.3.

$$S^*(2n) = 3S^*(n). \tag{1.31}$$

Proof. We prove a more general result:

$$\begin{aligned}
\sum_{n=2a}^{2b}{}^{*} s(n) &= \tfrac{1}{2}s(2a) + \sum_{k=a+1}^{b-1} s(2k) + \tfrac{1}{2}s(2b) + \sum_{k=a}^{b-1} s(2k+1) \\
&= \tfrac{1}{2}s(a) + \sum_{k=a+1}^{b-1} s(k) + \sum_{k=a}^{b-1}(s(k) + s(k+1)) + \tfrac{1}{2}s(b) \\
&= \tfrac{1}{2}s(a) + \sum_{k=a+1}^{b-1} s(k) + s(a) + \sum_{k=a+1}^{b-1} s(k) + \sum_{k=a+1}^{b-1} s(k) + s(b) + \tfrac{1}{2}s(b) \\
&= 3\left(\tfrac{1}{2}s(a) + \sum_{k=a+1}^{b-1} s(k) + \tfrac{1}{2}s(b)\right) = 3\sum_{n=a}^{b}{}^{*} s(n).
\end{aligned} \tag{1.32}$$

□

Since $S^*(1) = \frac{1}{2}$, it follows that $S^*(2^r) = \frac{1}{2} \cdot 3^r$, $S(2^r) = \frac{1}{2} \cdot (3^r + 1)$ and the sum of the r-th row of (1.11) is $S(2^{r+1}) - S(2^r) + 1 = 3^r + 1$, and:

$$\sum_{n=2^r}^{2^{r+1}-1} s(n) = \sum_{n=2^r+1}^{2^{r+1}} s(n) = \sum_{n=2^r}^{2^{r+1}}{}^{*} s(n) = 3^r. \tag{1.33}$$

This means that the average value of $s(n)$ for $n \in I_r$ is roughly $(\frac{3}{2})^r = (2^r)^\beta$, where

$$\beta = \frac{\log \frac{3}{2}}{\log 2} \approx .58496. \tag{1.34}$$

By comparing (1.34) and (1.27), we see that the ratio of the maximum of a row to the average of a row is unbounded, although the ratio grows slowly. For example the maximum value of $s(n)$ for $n \in I_{20}$ is $M_{20} = F_{22} = 17711$, while $(\frac{3}{2})^{20} \approx 3325$.

We shall show later that there is a continuous, strictly increasing function f, mapping $[0, 1]$ to itself, with the property that, if $N = 2^r(1 + t) \in I_r$, $0 \le t \le 1$, then

$$\frac{1}{3^r} \sum_{n=2^r}^{N}{}^{*} s(n) = f(t). \tag{1.35}$$

That is, the distribution of the "mass" of $s(n), n \in I_r$, is very well-behaved. The function f, which (of course) is differentiable a.e., has the property that $f'(w) = 0$ for every dyadic fraction $w = \frac{p}{2^q}$, but is singular at w when $w = \frac{p}{3 \cdot 2^q}$ is in lowest terms. Note also that by the linearity of the diatomic array and (1.9),

$$(1.36) \quad \sum_{k=0}^{2^r} Z(r,k;a,b) = \frac{3^r+1}{2}(a+b), \quad \sum_{k=0}^{2^r} s(2^r n + k) = \frac{3^r+1}{2} \cdot (s(n)+s(n+1)).$$

As another example of odd behavior of $s(n)$, it is easy to check that $s(F_r)$ is a Fibonacci number for $1 \le r \le 9$, but not for $10 \le r \le 200$ (at least). Even if we allow Lucas numbers ($L_m = \phi^m + \bar{\phi}^m = F_{m-1} + F_{m+1}$), the only "interesting" hits in this range are $s(L_{15}) = F_{10} = 55$ and $s(F_{27}) = L_{11} = 199$.

1.4. **The consecutive pairs** $(s(n), s(n+1))$**.** It is fair to say that some of the most important properties of the Stern sequence are in fact properties of the pairs $(s(n), s(n+1))$. The first one goes back to Stern himself.

Theorem 1.4. *For $n \ge 0$, $\gcd(s(n), s(n+1)) = 1$. If $a, b \ge 1$ and $\gcd(a,b) = 1$, then there is exactly one n so that $s(n) = a$ and $s(n+1) = b$.*

Proof. Inspection of $Z(1,1)$ shows that $\gcd(s(n), s(n+1)) = 1$ for small n. Since $\gcd(a, a+b) = \gcd(a+b, b) = \gcd(a,b)$, it follows that $s(n)$ and $s(n+1)$ are always relatively prime.

Now suppose $\gcd(a,b) = 1$. We induct on $\max(a,b) = m$. If $m = 1$, then the equation $s(n) = s(n+1) = 1$ clearly holds only for $n = 1$. Assume now that $m \ge 2$ and $m = a > b$. Then $s(n) = a$, $s(n+1) = b$ can only happen if $n = 2n' + 1$ is odd, and only if $s(n') = a - b$ and $s(n'+1) = b$. By the inductive hypothesis, this occurs for exactly one n'. A similar proof can be made if $a < b$, or else we can argue by reflection using (1.19). ☐

We will soon give an alternative proof which constructs n based on the continued fraction expansion of $\frac{a}{b}$.

Corollary 1.5. *Suppose $m \ge 1$. There are exactly $\phi(m)$ odd integers n with the property that $s(n) = m$.*

Proof. If n is odd and $s(n) = m$ and $s(n+1) = k$, then $k < m$ and $\gcd(m,k) = 1$. There are exactly $\phi(m)$ such integers k, and by Theorem 1.4, each k corresponds to exactly one n so that $s(n) = m$ and $s(n+1) = k$. ☐

The series $\sum_n s(n)^{-p}$ is never convergent, because $s(2^r) = 1$ for all r. However, Corollary 1.5 and standard results imply that

$$(1.37) \qquad \sum_{n=0}^{\infty} \frac{1}{(s(2n+1))^p} = \sum_{m=1}^{\infty} \frac{\phi(m)}{m^p} = \frac{\zeta(p-1)}{\zeta(p)},$$

provided $Re(p) > 2$.

Let

$$t(n) = \frac{s(n)}{s(n+1)}. \tag{1.38}$$

It follows from Theorem 1.4 that the sequence $(t(n))$ provides an enumeration of the non-negative rationals, and that one can recover $s(n)$ and $s(n+1)$ unambiguously from $t(n)$. (Cantor was a teenager in 1858, so it's understandable that Stern did not explicitly mention that $\mathbb{Q}$ is countable.) Further, for $n \in I_r$, the mirror symmetry implies that

$$t(n^* - 1) = \frac{s(n^*-1)}{s(n^*)} = \frac{s(n+1)}{s(n)} = \frac{1}{t(n)}. \tag{1.39}$$

so reciprocals appear in the same row. We shall see later that

$$\sum_{n=0}^{N} t(n) = \frac{3N}{2} + \mathcal{O}((\log N)^2). \tag{1.40}$$

One can then argue that the "average" positive rational number is $\frac{3}{2}$.

There are several natural ways to express the sequence $(t(n))$. Historically, the first was found by a French watchmaker named Achille Brocot, independently of Stern's work, who was interested in making a practical table of "gear ratios". His table was computed by starting with the fractions $\frac{0}{1}, \frac{1}{0}$, and then, if $\frac{a}{b}, \frac{c}{d}$ are consecutive in the r-th row, they are repeated in the $r+1$-st row, with $\frac{a+c}{b+d}$ inserted between them:

$$\begin{array}{ccccccccc} & & & \frac{0}{1} & & \frac{1}{0} & & & \\ & & & \frac{0}{1} & \frac{1}{1} & \frac{1}{0} & & & \\ & & \frac{0}{1} & \frac{1}{2} & \frac{1}{1} & \frac{2}{1} & \frac{1}{0} & & \\ \frac{0}{1} & \frac{1}{3} & \frac{1}{2} & \frac{2}{3} & \frac{1}{1} & \frac{3}{2} & \frac{2}{1} & \frac{3}{1} & \frac{1}{0} \\ & & & & \cdots & & & & \end{array} \tag{1.41}$$

Brocot array

In Stern terminology, the r-th row of the Borcot array consists of $\frac{s(k)}{s(2^r-k)}$; the numerators are the r-th row of $Z(0,1)$, the denominators are its reversal, or the r-th row of $Z(1,0)$. Each row is increasing from left-to-right. It is not immediately clear how to decode $t(n)$ from this array.

More directly, write the elements of $t(n)$ which appear in each row.

$$
\begin{array}{cccccccc}
 & & & \frac{1}{1} & & & & \\
 & & & \frac{1}{2} & \frac{2}{1} & & & \\
 & & \frac{1}{3} & \frac{3}{2} & \frac{2}{3} & \frac{3}{1} & & \\
\frac{1}{4} & \frac{4}{3} & \frac{3}{5} & \frac{5}{2} & \frac{2}{5} & \frac{5}{3} & \frac{3}{4} & \frac{4}{1} \\
 & & & \cdots & & & &
\end{array} \tag{1.42}
$$

$\boxed{t(n)}$

There is a subtle and confusing correspondence between these two arrays. The odd elements in the r-th row of the Brocot array, $\frac{s(2k+1)}{s(2^r-(2k+1))}$, appear in the $(r-1)$-st row of the $t(n)$ array, in a different order. This depends on reversing binary representations in a way that will be described later. Note that the rows of the $t(n)$ array are *not* constructed as simply as Farey sequences: $\frac{2}{5}$ and $\frac{3}{5}$ appear in the last row, but $\frac{1}{5}$ and $\frac{4}{5}$ do not.

Interestingly, the Stern sequence provides a second enumeration of the rationals. If $\gcd(a,b)=1$, not only is there a unique n so that $\frac{a}{b}=t(n)$, but there is also a unique odd k and r so that $\frac{a}{b}=\frac{s(k)}{s(2^r+k)}$. The array whose r-th row is $\frac{s(k)}{s(2^r+k)}$ has monotone rows, and is essential to Minkow- ski's bizarrely-named ?-function. You may recognize it as a Brocot array starting with $\frac{0}{1},\frac{1}{1}$ instead of $\frac{0}{1},\frac{1}{0}$; the k-th element in the r-th row is $\frac{s(k)}{s(2^r+k)}$:

$$
\begin{array}{ccccccccc}
 & & & & \frac{0}{1} \; \frac{1}{1} & & & & \\
 & & & \frac{0}{1} & \frac{1}{2} & \frac{1}{1} & & & \\
 & & \frac{0}{1} & \frac{1}{3} & \frac{1}{2} & \frac{2}{3} & \frac{1}{1} & & \\
\frac{0}{1} & \frac{1}{4} & \frac{1}{3} & \frac{2}{5} & \frac{1}{2} & \frac{3}{5} & \frac{2}{3} & \frac{3}{4} & \frac{1}{1} \\
 & & & & \cdots & & & &
\end{array} \tag{1.43}
$$

$\boxed{\text{Minkowski ?- array}}$

The *Minkowski ?-function*, which maps $[0,1]$ to itself, is defined recursively by taking the initial conditions $?(0)=0$, $?(1)=1$, and recursively applying the rule:

$$
?\left(\tfrac{a}{b}\right)=\tfrac{k}{2^r},\ ?\left(\tfrac{c}{d}\right)=\tfrac{k+1}{2^r} \implies ?\left(\tfrac{a+c}{b+d}\right)=\tfrac{2k+1}{2^{r+1}}. \tag{1.44}
$$

(We assume above that $\frac{a}{b}, \frac{c}{d}$ are given in lowest terms.) This definition only gives $?(x)$ for rational $x \in [0,1]$: $?(\frac{s(k)}{s(2^r+k)}) = \frac{k}{2^r}$, for $0 \le k \le 2^r$, but we shall see that ? extends uniquely to a continuous function.

1.5. **Continued fractions make their presence known.** There is a simple closed form for $t(n)$, which leads to simple formula for $s(n)$. This formula is useful for many symbolic purposes; however, for any specific numerical n, it is usually easier to use (1.1) directly to compute n; c.f. (1.16).

Suppose n is a positive integer, so $n \in I_r$ for some r, and let $[n]_2$ denote the usual binary representation of n, reading left-to-right, with no leading 0's and first digit 1:

$$n = \sum_{j=0}^{r} \epsilon_j(n) 2^j, \quad \epsilon_j(n) \in \{0,1\} \implies [n]_2 = [\epsilon_r(n), \dots, \epsilon_0(n)]_2 \tag{1.45}$$

It is useful to think of $[n]_2$ in terms of the blocks of consecutive 0's and 1's, and it is also useful to distinguish by the parity of n. If n is odd, then $[n]_2$ consists of an odd number of alternating blocks of 1's and 0's, beginning and ending with 1's. More specifically, suppose $[n]_2$ has a_1 1's, followed by a_2 0's, a_3 1's, etc, ending with a_{2v} 0's and a_{2v+1} 1's. (By convention, assume that $a_j \ge 1$.) We write $n \sim [a_1, \dots, a_{2v+1}]$ and observe that

$$n = 2^{a_1+\cdots+a_{2v+1}} - 2^{a_2+\cdots+a_{2v+1}} + - \cdots + 2^{a_{2v+1}} - 1. \tag{1.46}$$

If n is even, then $[n]_2$ consists of an even number of alternating blocks, beginning with 1's and ending with 0's. If $[n]_2$ has a_1 1's, followed by a_2 0's, a_3 1's, etc, ending with a_{2v-1} 1's and a_{2v} 0's, we write $n \sim [a_1, \dots, a_{2v}]$ and observe that

$$n = 2^{a_1+\cdots+a_{2v}} - 2^{a_2+\cdots+a_{2v}} + - \cdots + 2^{a_{2v-1}+a_{2v}} - 2^{a_{2v}}. \tag{1.47}$$

Suppose $n \sim [a_1, \dots, a_u]$ and $n' \sim [a_1, \dots, a_{u-1}]$. If n is odd, then n' is even, and $n = 2^{a_u} n' + 2^{a_u} - 1$; if n is even, then n' is odd, and $n = 2^{a_u} n'$. In either case, $\sum_j a_j = r + 1$. For example,

$$\begin{aligned} 243 = 2^7 + 2^6 + 2^5 + 2^4 + 2^1 + 2^0 = [11110011]_2 \implies 243 \sim [4,2,2], \\ 140 = 2^7 + 2^3 + 2^2 = [10001100]_2 \implies 140 \sim [1,3,2,2] \end{aligned} \tag{1.48}$$

Consider (1.42) as a binary tree, with the nodes labeled by the consecutive integers.

$$\begin{array}{ccccccc} & & & [1]_2 & & & \\ & & \swarrow & & \searrow & & \\ & [10]_2 & & & & [11]_2 & \\ \swarrow & & \searrow & & \swarrow & & \searrow \\ [100]_2 & & [101]_2 & & [110]_2 & & [111]_2 \\ & & & \cdots & & & \end{array} \tag{1.49}$$

Now write the binary tree with $t(n)$ at the node labeled by $[n]_2$.

$$\begin{array}{ccccccc} & & & \frac{1}{1} & & & \\ & & \swarrow & & \searrow & & \\ & \frac{1}{2} & & & & \frac{2}{1} & \\ \swarrow & & \searrow & & \swarrow & & \searrow \\ \frac{1}{3} & & \frac{3}{2} & & \frac{2}{3} & & \frac{3}{1} \\ & & & \cdots & & & \end{array} \tag{1.50}$$

From the definition,

$$\begin{aligned} t(2n) &= \frac{s(2n)}{s(2n+1)} = \frac{s(n)}{s(n)+s(n+1)} = \frac{1}{1+\frac{1}{t(n)}} = \frac{t(n)}{t(n)+1} \\ t(2n+1) &= \frac{s(2n+1)}{s(2n+2)} = \frac{s(n)+s(n+1)}{s(n+1)} = t(n)+1. \end{aligned} \tag{1.51}$$

Let

$$f_0(x) = \frac{x}{x+1}, \quad f_1(x) = x+1. \tag{1.52}$$

Then we see that, using the terminology of (1.45) and noting that $\epsilon_r(n) = 1$ for all n and $t(0) = 0$, we have

$$t(n) = f_{\epsilon_0(n)}(f_{\epsilon_1(n)}(\cdots(f_{\epsilon_{r-1}(n)}(f_{\epsilon_r(n)}(0)))\cdots)). \tag{1.53}$$

If we let $g^{(k)}$ denote the k-th iterate of a function g, it is routine to check that

$$f_0^{(k)}(x) = \frac{x}{kx+1} = \frac{1}{k+\frac{1}{x}}, \quad f_1^k(x) = x+k. \tag{1.54}$$

The effect of appending k 0's or k 1's to $[n]_2$ implies that

$$t(2^k n) = \frac{1}{k+\frac{1}{t(n)}}, \quad t(2^k n + 2^k - 1) = k + t(n). \tag{1.55}$$

Also note that similar expressions exist if one chooses to view the consecutive pair of Stern-values as a column matrix rather than as a fraction. Let

$$M_0 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \qquad M_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}. \tag{1.56}$$

Then

$$\begin{bmatrix} s(2n) \\ s(2n+1) \end{bmatrix} = M_0 \begin{bmatrix} s(n) \\ s(n+1) \end{bmatrix}, \qquad \begin{bmatrix} s(2n+1) \\ s(2n+2) \end{bmatrix} = M_1 \begin{bmatrix} s(n) \\ s(n+1) \end{bmatrix}. \tag{1.57}$$

It follows that

$$\begin{bmatrix} s(n) \\ s(n+1) \end{bmatrix} = M_{\epsilon_0(n)} M_{\epsilon_1(n)} \cdots M_{\epsilon_{r-1}(n)} M_{\epsilon_r(n)} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{1.58}$$

The powers of these matrices are familiar from the study of continued fractions.

$$M_0^k = \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}, \qquad M_1^k = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}. \tag{1.59}$$

Theorem 1.6.

(i) If n is odd and $n \sim [a_1, \dots, a_{2v+1}]$, then

$$t(n) = \frac{s(n)}{s(n+1)} = a_{2v+1} + \frac{1}{a_{2v} + \frac{1}{\dots + \frac{1}{a_1}}}. \tag{1.60}$$

(ii) If n is even and $n \sim [a_1, \dots, a_{2v}]$, then

$$t(n) = \frac{s(n)}{s(n+1)} = \frac{1}{a_{2v} + \frac{1}{\dots + \frac{1}{a_1}}}. \tag{1.61}$$

Proof. Let $n \sim [a_1, \dots, a_k]$. We argue by induction on k. If $k = 1$, then $n \sim [a_1]$, so $n = 2^{a_1} - 1$ and (1.14) shows that

$$t(2^r - 1) = \frac{s(2^r - 1)}{s(2^r)} = \frac{r}{1} = r. \tag{1.62}$$

If $n > 1$ is odd, then k is odd, $n = 2^{a_{2v+1}} n' + 2^{a_{2v+1}} - 1$, where n' is even, and (1.61) applies to n' by the inductive hypothesis. By (1.55), (1.60) holds for n. Similarly, if $n > 0$ is even, then k is even, $n = 2^{a_{2v+1}} n'$ where n' is odd, and (1.60) applies to n' by the inductive hypothesis. By (1.55), (1.61) holds for n. □

Suppose a rational number $\frac{a}{b} > 1$ is given. One may write it as a simple continued fraction (i.e., as in (1.60)) in two ways, because the final denominator may either be chosen as m (for $m \geq 2$) or $(m-1) + \frac{1}{1}$. Exactly one of these representations will have an odd number of denominators, to which Theorem 1.6 will apply, stating that $s(n) = a, s(n+1) = b$. If $\frac{a}{b} < 1$, either apply Theorem 1.6 to $\frac{b}{a}$ and apply (1.39), or use the representation of $\frac{b}{a}$ with an even number of denominators and apply Theorem 1.6 directly.

Later on, we'll see that Theorem 1.6 implies that the ?-function gives a strictly increasing bijection of the rationals in $[0, 1]$ onto the dyadic rationals in $[0, 1]$. The properties of periodic infinite continued fractions imply that that the ?-function gives a strictly increasing bijection of the algebraic numbers of degree ≤ 2 in $[0, 1]$ onto the rationals.

Further, $h(a, b)$, the sum of the denominators in the continued fraction representation of $\frac{a}{b}$ does not depend on which choice of representation is used, and equals the number of binary digits in n, so $n \in I_{h(a,b)-1}$ and $\frac{a}{b}$ will appear in row $h(a, b) - 1$. For example, $4 = 2 + 1 + 1 = 1 + 2 + 1 = 1 + 1 + 2$ are the ordered partitions of 4 as an odd number of summands and

$$4 = \frac{4}{1}, \quad 2 + \frac{1}{1 + \frac{1}{1}} = \frac{5}{2}, \quad 1 + \frac{1}{2 + \frac{1}{1}} = \frac{4}{3}, \quad 1 + \frac{1}{1 + \frac{1}{2}} = \frac{5}{3}. \tag{1.63}$$

These (together with their reciprocals) are the entries of the fourth row of (1.42).

One somewhat surprising consequence of the properties of simple continued fractions is the following. For odd n, let $\overleftarrow{n}$ denote the integer whose binary expression is the reversal of n; more formally, if $n \sim [a_1, \dots, a_{2v+1}]$, then $\overleftarrow{n} \sim [a_{2v+1}, \dots, a_1]$. It turns out that $s(\overleftarrow{n}) = s(n)$ and $s(n+1)s(\overleftarrow{n}+1) \equiv 1 \pmod{s(n)}$. It also turns out that $[n^*]_2$ can be expressed in terms of $[n]_2$ and that $\overleftarrow{(n^*)} = (\overleftarrow{n})^*$. Indeed, suppose (1.45) holds. Then we first claim that

$$n^* = 1 + \sum_{j=1}^{r-1} (1 - \epsilon_j(n)) 2^j + 2^r, \tag{1.64}$$

because this is equivalent to the assertion that $n + n^* = 2 + \sum_{j=1}^{r-1} 2^j + 2 \cdot 2^r = 3 \cdot 2^r$. Similarly, we obtain

$$\overleftarrow{n} = 1 + \sum_{j=1}^{r-1} \epsilon_{r-j}(n) 2^j + 2^r, \quad \overleftarrow{(n^*)} = (\overleftarrow{n})^* = 1 + \sum_{j=1}^{r-1} (1 - \epsilon_{r-j}(n)) 2^j + 2^r. \tag{1.65}$$

Suppose n is odd and

$$s(n-1) = m - a, \quad s(n) = m, \quad s(n+1) = a. \tag{1.66}$$

Then we have

$$s(n^* - 1) = a, \quad s(n^*) = m, \quad s(n+1) = m - a, \tag{1.67}$$

and if $b < m$ is such that $ab \equiv 1 \pmod{m}$, then (after a later proof),

$$\begin{aligned} s(\overleftarrow{n} - 1) = m - b, \quad s(\overleftarrow{n}) = m, \quad s(\overleftarrow{n} + 1) = b \\ s(\overleftarrow{n}^* - 1) = b, \quad s(\overleftarrow{n}^*) = m, \quad s(\overleftarrow{n}^* + 1) = m - b. \end{aligned} \tag{1.68}$$

Unless $n = 3$, it is always the case that $n \neq n^*$, although $n = \overleftarrow{n}$ or $n = \overleftarrow{n}^*$ occur roughly $2^{r/2}$ times in the r-th row; these correspond to $a^2 \equiv \pm 1 \pmod{m}$, where $m = s(n)$. Thus one can usually expect $s(n) = m$ to occur in groups of four odd n in a row for a given m – for $n, n^*, \overleftarrow{n}, \overleftarrow{n^*}$.

As a numerical example, using the computation $243 \sim [4, 2, 2]$ from (1.48), we have

$$t(243) = \frac{s(243)}{s(244)} = 2 + \frac{1}{2 + \frac{1}{4}} = 2 + \frac{4}{9} = \frac{22}{9}, \tag{1.69}$$

so $s(243) = 22$ and $s(244) = 9$. As a double-check of reflection, $244^* = 140$, $243^* = 141$ and $140 \sim [1, 3, 2, 2]$ so

$$t(140) = \frac{s(140)}{s(141)} = \frac{1}{2 + \frac{1}{2 + \frac{1}{3 + \frac{1}{1}}}} = \frac{1}{2 + \frac{1}{2 + \frac{1}{4}}} = \frac{9}{22}. \tag{1.70}$$

Since $[\overleftarrow{243}]_2 = 11001111$, $\overleftarrow{243} = 207$, and

$$t(207) = \frac{s(207)}{s(208)} = 4 + \frac{1}{2+\frac{1}{2}} = 4 + \frac{2}{5} = \frac{22}{5}. \tag{1.71}$$

Note that $9 \cdot 5 = 22 \cdot 2 + 1$, that $243^* = 3 \cdot 128 - 243 = 141$, $[141]_2 = 10001101$, $[\overleftarrow{141}]_2 = 10110001$, so $\overleftarrow{141} = 177 = 207^*$, and just to finish up,

$$t(141) = \frac{s(141)}{s(142)} = \frac{22}{13}, \quad t(177) = \frac{s(177)}{s(178)} = \frac{22}{17}. \tag{1.72}$$

We can use Theorem 1.6 to determine the solutions to the equation $s(n) = m$ for fixed m. For example, suppose n is odd and $s(n) = 12$. Since $\gcd(12, s(n+1)) = 1$ and $s(n+1) < 12$, we must have $s(n+1) \in \{1, 5, 7, 11\}$. Every positive rational has a finite continued fraction: and, as

$$\frac{12}{1} = 12, \quad \frac{12}{5} = 2 + \frac{1}{2+\frac{1}{2}}, \quad \frac{12}{7} = 1 + \frac{1}{1+\frac{1}{2+\frac{1}{2}}}, \quad \frac{12}{11} = 1 + \frac{1}{11}. \tag{1.73}$$

Since the last two expressions have an even number of denominators, we tweak the innermost denominator to give an odd length:

$$\frac{12}{7} = 1 + \frac{1}{1+\frac{1}{2+\frac{1}{1+\frac{1}{1}}}}, \qquad \frac{12}{11} = 1 + \frac{1}{10+\frac{1}{1}}. \tag{1.74}$$

Thus, we see that $s(n) = 12$ when $[n]_2 \sim [12], [2,2,2], [1,1,2,1,1]$ or $[1,10,1]$, that is, when $[n]_2 = [111111111111]_2 = 2^{12} - 1 = 4095$, $[n]_2 = [110011]_2 = 51$, $[n]_2 = [101101]_2 = 45$ or $[n]_2 = [100000000001]_2 = 2^{11} + 1 = 2049$. Note that $45^* = 51$ and $2049^* = 4093$, and all four binary expressions are palindromes, since $1^2, 5^2, 7^2, 11^2 \equiv 1 \pmod{24}$.

As a symbolic example, $(2^r - 1)^2 = 2^{2r} - 2^{r+1} + 1 = 2^{2r-1} + 2^{2r-2} + \cdots + 2^{r+1} + 1$, so $(2^r - 1)^2 \sim [r-1, r, 1]$ and,

$$t((2^r - 1)^2) = \frac{s((2^r-1)^2)}{s((2^r-1)^2+1)} = 1 + \frac{1}{r + \frac{1}{r-1}} = \frac{r^2}{r^2 - r + 1}. \tag{1.75}$$

1.6. **The Stern sequence mod d and some combinatorial interpretations.** For integers $d \ge 2$ and $0 \le i \le d-1$, let

$$A(d,i) = \{n : s(n) \equiv i \pmod{d}\}, \tag{1.76}$$

and let

$$\begin{aligned} T(N;d,i) &= |\{n \in A(d,i) : 0 \le n < N\}|, \\ U(r;d,i) &= T(2^{r+1};d,i) - T(2^r;d,i). \end{aligned} \tag{1.77}$$

Stern noted that $s(n) \pmod 2$ is periodic with period 3; in this notation, $A(2;0) = 3\mathbb{N}$ and $T(N;2,0) = \lfloor \frac{n+2}{3} \rfloor$, so $U(r;2,0) = \frac{1}{3}(2^r - (-1)^r)$. We shall prove later that each $U(r;d,i)$ satisfies a linear recurrence, and

$$u(d;i) := \lim_{r\to\infty} \frac{U(r;d,i)}{2^r} \tag{1.78}$$

exists as a computable arithmetic function of (d,i). A stronger statement is true:

$$\lim_{N\to\infty} \frac{T(N;d,i)}{N} = u(d,i); \tag{1.79}$$

that is, the limit applies even through elements in the middle of the rows.

Define the arithmetic function $I(m)$ by

$$\frac{I(m)}{m} = \prod_{p\ \mid\ m} \frac{p+1}{p}. \tag{1.80}$$

Then $u(d,0) = \frac{1}{I(d)}$. In some cases, when $I(d_1) = I(d_2)$, an even stronger statement can be made: $I(4) = I(5) = 6$ and $U(r;4;0) = U(r;5,0)$. That is the number of multiples of 4 and the number of multiples of 5 is the same for $\{s(n) : n \in I_r\}$. Further, $I(6) = I(8) = I(9) = I(11) = 12$, and $U(r;6,0) = U(r,9,0) = U(r;11,0)$, but $U(8;r,0)$ is different. The function I has interesting iterative behavior: there exist $a(d), b(d)$ such that, for each integer d and sufficiently large N, $I^{(N)}(d) = 2^{N+a(d)}3^{b(d)}$.

These observations are a consequence of studying the behavior of Stern pairs modulo d. Note that if $(s(n), s(n+1)) \equiv (a,b) \pmod d$, then $\gcd(a,b,d) = 1$. It turns out that the residue classes are always uniformly distributed among these possible pairs. The argument requires a Markov chain model.

In case $d = 3$, stronger information can be presented. Define the set $\mathcal{S}_3 \subset \mathbb{N}$ recursively by:

$$0,5,7 \in \mathcal{S}_3, \qquad 0 < n \in \mathcal{S}_3 \implies 2n, 8n\pm 5, 8n\pm 7 \in \mathcal{S}_3. \tag{1.81}$$

(Thus, the smallest non-negative integers in $\mathcal{S}_3$ are: 0, 5, 7, 10, 14, 20, 28, 33, 35, 40, 45, 47, 49, 51, 56, 61, 63.) This is a member of an interesting family of recursively defined sets, and associated directed graphs on $\mathbb{Z}$.

Theorem 1.7. *$A(3,0) = \mathcal{S}_3$.*

Proof. We first observe that by (1.13),

$$s(2n) = s(n), \quad s(8n\pm5) = 2s(n)+3s(n\pm1), \qquad s(8n\pm7) = s(n)+3s(n\pm1). \tag{1.82}$$

Thus, 3 divides $s(n)$ if and only if 3 divides $s(2n), s(8n\pm5), s(8n\pm7)$. Every $n \in \mathbb{N}$ belongs to exactly one of the congruence classes $0 \pmod 2$, $\pm5 \pmod 8$, $\pm7 \pmod 8$, and if $n \ge 2$, then $n = 2n', 8n' \pm 5$ or $8n' \pm 7$ with $n' < n$. Thus, the inductive construction of $\mathcal{S}_3$ gives all n for which $s(n)$ is a multiple of 3. ☐

We shall also show that $T(N;3,0) = \frac{1}{4}N + \mathcal{O}(N^{1/2})$, with the error bound best possible, and $T(N;3,1) - T(N;3,2) \in \{0,1,2,3\}$.

We turn to digital questions. Let $\mathcal{A} = \{0 = a_0 < a_1 < \cdots a_r\}$ denote a finite subset of $\mathbb{N}$ containing 0, and let $f_{\mathcal{A}}(m)$ denote the number of ways to write m in the form

$$m = \sum_{k=0}^{\infty} \epsilon_k 2^k, \qquad \epsilon_k \in \mathcal{A}. \tag{1.83}$$

For example, the binary representation of m implies that $f_{\{0,1\}}(m) = 1$ for all m.

Theorem 1.8.

$$s(n) = f_{\{0,1,2\}}(n-1). \tag{1.84}$$

Proof. Let $f(m) = f_{\{0,1,2\}}(m)$ for short. Observe that $0 = s(0) = f(-1)$ trivially and $1 = s(1) = f(0)$, since the only way to write 0 as (1.83) is to have $\epsilon_k \equiv 0$ for all k. We now show that

$$f(2n-1) = f(n-1), \qquad f(2n) = f(n) + f(n-1), \tag{1.85}$$

and this will establish the theorem by induction, by comparison with (1.1) and (1.84)

Notice that $m \equiv \epsilon_0 \pmod 2$ in (1.83); moreover,

$$m = \epsilon_0 + 2\sum_{j=0}^{\infty} \epsilon_{j+1} 2^j = \epsilon_0 + 2m', \tag{1.86}$$

and the representation of m' obeys (1.83) as well. It follows that in any representation of $2n-1$ in (1.83), we must have $\epsilon_0 = 1$, with $m' = n-1$, and in any representation of $2n$ in (1.83), we may have $\epsilon_0 = 0$ or 2, with $m = n$ or $n-1$ respectively. □

Theorem 1.8 will become more transparent when we discuss the generating function for the Stern sequence:

$$\sum_{n=0}^{\infty} s(n)x^n = x\prod_{j=0}^{\infty}(1 + x^{2^j} + x^{2^{j+1}}).$$

Another combinatorial interpretation is almost trivial. Let G be a directed graph whose vertices are $\mathbb{N}$ and whose directed edges are precisely those of the the form $(2k, k), (2k+1, k), (2k+1, k+1)$. Then $s(n)$ is the number of paths from n to 1. Somewhat surprisingly, G is planar: to draw it in this way, put the vertices from I_r, $n = 2^r + k$, at the points $(\log(n+1))e^{2\pi i \frac{k}{2^r}}$ on a spiral in the plane.

1.7. **Exercises.** Do some of these. Extensions of Problem 10 will show up in subsequent exercise sets, so it's a good one to do. Let's say a deadline of Fri. Feb. 3, at the beginning of class, at which point I'll pass out solutions.

1. Write today's date as MMDDYYYY $\in [10^7, 10^9)$ and compute $s(n)$. For example, the first day of class was Jan. 18, 2012, and $s(01182012) = 1244$. (Europeans should use DDMMYYYY; $s(18012012) = 15394$.) This problem can be done on several days in a row, especially using a program.

2. Determine n so that $s(n) = 2012$ and $s(n+1) = 595$. Note that $2012 = 2^2 \cdot 503$ and $595 = 5 \cdot 7 \cdot 17$ are relatively prime.

3. Prove that

$$\sum_{n=0}^{N}{}^{*} n = \frac{N^2}{2}, \qquad \sum_{n=0}^{N}{}^{*} n^2 = \frac{N^3}{3} + \frac{N}{6},$$

and compute

$$\sum_{n=0}^{N}{}^{*} n^3.$$

4. Let $\nu_p(n)$ denote the exponent of p in the prime factorization of n. Show that

$$\frac{s(n-1)+s(n+1)}{s(n)} = 1 + 2\nu_2(n).$$

5. Determine, by any correct method, all odd integers n so that $s(n) \in \{10, 11\}$.

6. Using (1.22), compute $[n_r]_2$; there are two slightly different answers, depending on the parity of r.

7. Find and prove a formula relating

$$T(n) := \sum_{k=0}^{\lfloor n/3 \rfloor} s(n-3k)$$

and $S(n)$.

8. For $r \ge 1$ and $t \ge 0$, compute $s((2^r+1)^2)$ and $s((2^r-1)(2^{r+t}-1))$.

9. Sometimes the Stern sequence fakes you out. Suppose

$$c_r = \left(\sum_{n \in I_r}{}^{*} 1\right)\left(\sum_{n \in I_r}{}^{*} s(n)^2\right) - \left(\sum_{n \in I_r} s(n)\right)^2.$$

We know that $c_r \ge 0$ by the Cauchy-Schwarz Inequatilty. Show that $c_1 = 1$, $c_2 = 11$ and $c_3 = 111$. Compute the disheartening value of c_4.

10. (The first in a series.) Show that for $k \in \mathbb{N}$, there exist functions $A(k), B(k)$ so that for $r > \log_2 k$,

$$s(2^r - k) = A(k)r + B(k).$$

The most instructive way to do this problem is to see what happens for small values of k first. The recursion is helpful, the continued fraction, less so.

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:

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

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 $\lambda := (\lambda_1, \ldots, \lambda_d) \in \mathbb{C}^d$ 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* $a_n X^n$, 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

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

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

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

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":

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

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:

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

(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

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

(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):

$$\frac{\mathcal{S}(X)}{1-X^3} = \mathcal{S}(X)(1 + X^3 + X^6 + \cdots) = \sum_{n=0}^{\infty} \left(\sum_{j=0}^{\lfloor n/3 \rfloor} s(n-3j) \right) X^n;$$

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

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

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

Thus,

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

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

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

As we have seen,

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

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 $n', n'' \ge 0$; 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 $d \ge 3$; 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 $m \ge 0$, 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 $(1-x)^{-1}$. □

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 $u(n)$.

STERN NOTES, MATH 595, SPRING 2012

BRUCE REZNICK, UIUC

3. Linear recurrences

3.1. **Basics.** In this section, we talk about constant-coefficient linear recurrences; all objects should be viewed as living in $\mathbb{C}$. Recurrences usually arise in one of two forms. The first is the direct one: suppose that $(a(n))$ is a sequence which satisfies the *homogeneous d-th order linear recurrence*:

$$a(n) + \sum_{j=1}^{d} c_j a(n-j) = 0, \qquad n \ge d, \tag{3.1}$$

or, equivalently,

$$a(n+d) = -\sum_{j=1}^{d} c_j a(n+d-j) = 0, \qquad n \ge 0. \tag{3.2}$$

We say that the undetermined values of $a(i)$, $0 \le i \le d-1$, are the *initial conditions*. Associated to the recurrence (3.1) is the *characteristic polynomial*

$$\phi(z) = z^d + \sum_{j=1}^{d} c_j z^{d-j}. \tag{3.3}$$

If $\phi(t_0) = 0$, then $a(n) = t_0^n$ is easily seen to satisfy (3.1). We also make the fairly obvious point that if (3.1) holds for $(a(n))$, then so does

$$a(n+k) + \sum_{j=1}^{d} c_j a(n+k-j) = 0, \qquad n \ge d, \tag{3.4}$$

for $k \ge 1$, and by adding together identities such as (3.4), it is not hard to prove that, if $\zeta(z) = \phi(z)\eta(z)$, then $(a(n))$ also satisfies the recurrence whose characteristic polynomial is ζ.

The second way recurrences arise is as a *matrix system*: d sequences $(a_j(n))$, $1 \le j \le d$, which are related by:

$$a_j(n+1) = \sum_{k=1}^{d} m_{jk} a_k(n), \qquad n \ge 0, \quad 1 \le j \le d,$$
$$\begin{pmatrix} a_1(n+1) \\ \cdots \\ a_d(n+1) \end{pmatrix} = \begin{pmatrix} m_{11} & \cdots & m_{1d} \\ \cdots & \cdots & \cdots \\ m_{d1} & \cdots & m_{dd} \end{pmatrix} \begin{pmatrix} a_1(n) \\ \cdots \\ a_d(n) \end{pmatrix}. \tag{3.5}$$

More formally, let $A(n) = (a_1(n) \cdots a_d(n))^T$ be the column vector of sequences and let $M = [m_{jk}]$ be the matrix of coefficients. Then (3.5) becomes

$$A(n+1) = MA(n) \implies A(n) = M^n A(0), \quad n \ge 0. \tag{3.6}$$

Here, $A(0)$ provides the initial condition.

Although the methods of solutions for these two kinds of recurrences are different, they can each be transformed to the other. The Cayley-Hamilton Theorem states that if $\phi(\lambda) = \det(\lambda I_d - M)$ is the characteristic polynomial of M, then $\phi(M) = 0$, where "0" is construed as the $d \times d$ matrix of 0's. Supposing ϕ is given by (3.3) for convenience, we then have

$$\begin{aligned} \phi(t) = t^d + c_1 t^{d-1} + \cdots + c_d &\implies M^d + c_1 M^{d-1} + \cdots + c_d I_d = 0 \\ &\implies M^{n+d} + c_1 M^{n+d-1} + \cdots + c_d M^n = 0 \\ &\implies A(n+d) + c_1 A(n+d-1) + \cdots + c_d A(d) = 0, \end{aligned} \tag{3.7}$$

where the final "0" is the zero column vector. The last equation in (3.7) is simply the assertion that each $a_j(n)$ satisfies (3.1).

On the other hand, (3.2) can be simple-mindedly rewritten as:

$$\begin{pmatrix} a(n+1) \\ a(n+2) \\ \cdots \\ a(n+d-1) \\ a(n+d) \end{pmatrix} = \begin{pmatrix} 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & 1 \\ -c_d & -c_{d-1} & \cdots & -c_2 & -c_1 \end{pmatrix} \begin{pmatrix} a(n) \\ a(n+1) \\ \cdots \\ a(n+d-2) \\ a(n+d-1) \end{pmatrix}. \tag{3.8}$$

The matrix in (3.8) (or its transpose) is sometimes called the *companion matrix* to the polynomial ϕ, and has characteristic polynomial $(-1)^d \phi$.

3.2. **Solving recurrences.** For completeness' sake, we include a self-contained proof of the method of Partial Fractions, which requires the Fundamental Theorem of Algebra to factor the denominator, but no explicit theory of complex variables.

Suppose

$$F(z) = \frac{p(z)}{q(z)} = \frac{b_n z^n + b_{n-1} z^{n-1} + \cdots + b_0}{c_m z^m + c_{m-1} z^{m-1} + \cdots + c_0} \tag{3.9}$$

is a rational function, where $b_k, c_k \in \mathbb{C}$, $b_n c_m \neq 0$ and $m > n$. Assume that F is presented in lowest terms and $c_m = 1$. Suppose further that

$$q(z) = \prod_{j=1}^{r} (z - z_j)^{m_j}, \tag{3.10}$$

where $z_j \in \mathbb{C}$, the z_j's are distinct, $m_j \in \mathbb{N}$ with $\sum m_j = m$, and $p(z_j) \neq 0$. Then as we tell our calculus students:

Theorem 3.1. *Suppose F is given by* (3.9)*, and* (3.10) *holds. Then there exist $r_{j\ell} \in \mathbb{C}$ so that*

$$F(z) = \sum_{j=1}^{r} \left(\frac{r_{j1}}{z - z_j} + \cdots + \frac{r_{jm_j}}{(z - z_j)^{m_j}} \right). \tag{3.11}$$

Conversely, if F is given by (3.11)*, then F is a rational function of shape* (3.9) *for which $n < m$, and* (3.10) *hold.*

Proof. We induct on $m = \deg(q)$. If $m = 1$, then $n = 0$ and there is nothing to prove. Suppose the theorem is valid for q with $\deg q \leq m - 1$ and suppose $q(z) = (z - z_1)^{m_1} \bar{q}(z)$, with $\bar{q}(z_1) \neq 0$. We do not rule out the possibility that $\bar{q}(z) \equiv 1$. Consider the expression

$$F(z) - \frac{\alpha}{(z - z_1)^{m_1}} = \frac{p(z) - \alpha \bar{q}(z)}{q(z)}.$$

Let $p_\alpha(z) = p(z) - \alpha \bar{q}(z)$; if $\alpha_0 = \frac{p(z_1)}{\bar{q}(z_1)}$, then $p_{\alpha_0}(z_1) = 0$, hence $p_{\alpha_0}(z) = (z - z_1)\hat{p}(z)$. It follows that

$$F(z) = \frac{\alpha_0}{(z - z_1)^{m_1}} + \frac{\hat{p}(z)}{(z - z_1)^{m_1 - 1} \bar{q}(z)},$$

and a partial fraction expression for $F(z) - \frac{\alpha_0}{(z - z_1)^{m_1}}$ exists by the inductive hypthesis.

For the converse, if F is given by (3.11), then multiplication by $q(z)$ yields a polynomial on the right-hand side, with degree at most $m - 1$. □

We now return to (3.1) and add a subtle new hypothesis:

$$c_d \neq 0. \tag{3.12}$$

(This retriction actually offers no practical limitations. Suppose (3.12) is not satisfied. If $c_j = 0$ for all j, then (3.1) has only the zero solution. Otherwise, suppose $c_e \neq 0$ and $c_j = 0$ for $e + 1 \leq j \leq d$. Let $\tilde{a}(n) = a(n + d - e), n \geq 0$. Then (3.1) becomes

$$\tilde{a}(n) + \sum_{j=1}^{e} c_j \tilde{a}(n - j) = 0, \qquad n \geq e,$$

with no constraint involving $a(k)$ for $0 \leq k < d - e$. In this case, first solve for $\tilde{a}$ by the algorithm described below, and then write $a(n) = \tilde{a}(n - (d - e))$ for $n \geq d - e$.)

It is now convenient to define

$$\psi(z) = 1 + \sum_{j=1}^{d} c_j z^j = z^d \phi(z^{-1}),$$

as a polynomial with true degree d (by (3.12)) and note that $z_j \neq 0$ in (3.10) and

$$\phi(z) = \prod_{j=1}^{r} (z - z_j)^{m_j} \iff \psi(z) = \prod_{j=1}^{r} (1 - z z_j)^{m_j}.$$

Let $M = \max(1, \sum_j |c_j|) \geq 1$. If $|a(i)| \leq T$ for $i = 0, \dots, d-1$, then

$$|a(d)| \leq \sum_{j=1}^{d} |c_j||a(d-j)| \leq T \sum_{j=1}^{d} |c_j| \leq TM,$$

and so $|a(i)| \leq MT$ for $i = 1, \dots, d$. An easy induction implies that $|a(n)| \leq M^{n+1-d}T$ for each $n \geq d$, and so the generating function for $(a(n))$ will be an analytic function with radius of convergence $\geq M^{-1}$. Let

$$f(z) := \sum_{n=0}^{\infty} a(n) z^n.$$

Then, as we saw in (2.4),

$$\psi(z) f(z) = \sum_{n=0}^{d-1} \left(a(n) + \sum_{j=1}^{n} c_j a(n-j) \right) z^n := p(z), \tag{3.13}$$

so $f(z)$ is a rational function:

$$f(z) = \frac{p(z)}{\psi(z)} = \frac{p(z)}{\prod_{j=1}^{r} (1 - z_j z)^{m_j}}, \tag{3.14}$$

where $\deg(p) \leq d - 1 < d$.

It follows from Theorem 3.1 that there exist $r_{j\ell} \in \mathbb{C}$ such that

$$\sum_{n=0}^{\infty} a_n z^n = \sum_{j=1}^{r} \left(\frac{r_{j1}}{1 - z_j z} + \cdots + \frac{r_{jm_j}}{(1 - z_j z)^{m_j}} \right). \tag{3.15}$$

The power series for the right-hand side was already computed in (2.18):

$$\frac{1}{1 - \lambda z} = \sum_{n=0}^{\infty} \lambda^n z^n, \quad \frac{1}{(1 - \lambda z)^{r+1}} = \sum_{n=0}^{\infty} \binom{n+r}{r} \lambda^n z^n, \ r \geq 1.$$

Thus, the coefficient of z^n in the j-th summand in (3.15) can be expressed as follows:

$$\left(r_{j1} + \sum_{\ell=2}^{m_j} r_{j\ell} \cdot \frac{(n+1) \cdots (n+\ell-1)}{(\ell-1)!} \right) z_j^n = p_j(n) z_j^n, \tag{3.16}$$

where p_j is a polynomial with degree $\le m_j - 1$. The coefficients of p_j depend on the r_ℓ's, which depend on the initial conditions of the recurrence. We have therefore proved the main theorem about linear recurrences.

Theorem 3.2. *If $(a(n))$ is a sequence satisfying (3.1) with $c_d \neq 0$, and if*

$$\phi(z) = z^d + \sum_{j=1}^{k} c_j z^{d-j} = \prod_{j=1}^{r} (z - z_j)^{m_j}, \tag{3.17}$$

then there exist polynomials p_j so that for $n \ge 0$,

$$a(n) = \sum_{j=1}^{r} p_j(n) z_j^n, \quad \deg(p_j) \le m_j - 1. \tag{3.18}$$

Conversely, any sequence $(a(n))$ defined by (3.18) satisfies the recurrence (3.1).

The proof of the last assertion is that (3.18) implies that f is given by (3.15), and so ψf is a polynomial. In practice, "most" polynomials have distinct roots, so the polynomials p_j are, in fact, "usually" constants.

As noted in Chapter 2, the set of sequences satisfying a recurrence such as (3.1) forms a d-dimensional vector space. One natural basis follows from (3.18), namely

$$\left\{ n^i z_j^n : \quad 1 \le j \le r,\ 0 \le i \le m_j - 1 \right\}.$$

A somewhat more natural basis comes from (3.14) by considering those sequences whose generating functions are given by $\frac{z^i}{\psi(z)}$ for $0 \le i \le d-1$. Define $b_0(n)$ to be the sequence satisfying (3.1) with initial conditions

$$b_0(0) = \cdots = b_0(d-2) = 0, \quad b_0(d-1) = 1.$$

It follows from (3.13) that

$$\sum_{n=0}^{\infty} b_0(n) z^n = \frac{z^{d-1}}{\psi(z)}.$$

Upon dividing (3.2) by z^k, $1 \le k \le d-1$, we see that

$$\sum_{n=0}^{\infty} b_0(n+k) z^n = \frac{z^{d-1-k}}{\psi(z)}.$$

In other words, the sequences $\{(b_0(n)), (b_0(n+1)), \ldots, (b_0(n+d-1))\}$ form a basis for this subspace. This is familiar in the Fibonacci sequence setting.

There isn't too much to say about solving (3.6) directly. The standard methodology is to put the matrix in Jordan canonical form:

$$M = C^{-1} D C \implies M^n = C^{-1} D^n C. \tag{3.19}$$

If the characteristic polynomial of M has distinct roots, then D is diagonal and D^n is easy to calculate; otherwise, D is a block diagonal matrix.

3.3. **Standard examples.** Here is a simple example of a recurrence with a repeated root. Let

$$a(n) = 4a(n-1) - 4a(n-2); \quad a(0) = r,\ a(1) = 2s; \quad \phi(z) = (z-2)^2. \tag{3.20}$$

Then $a(2) = 4(2s-r), a(3) = 4(8s-4r) - 4(2s) = 24s - 16r = 8(3s-2r)$, and

$$\begin{aligned} (1-4z+4z^2)\sum_{n=0}^{\infty} a(n)z^n &= a(0) + (a(1) - 4a(0))z \\ + \sum_{n=2}^{\infty}(a(n) - 4a(n-1) + 4a(n-2))z^n &= r + (2s-4r)z. \end{aligned} \tag{3.21}$$

Thus,

$$\begin{aligned} \sum_{n=0}^{\infty} a(n)z^n &= \frac{r+(2s-4r)z}{(1-2z)^2} = \frac{2r-s}{1-2z} + \frac{-r+s}{(1-2z)^2} \\ \implies a(n) &= ((2r-s) + (s-r)(n+1))\, 2^n = 2^n(ns - (n-1)r), \end{aligned} \tag{3.22}$$

as can verified by the first few terms of the series given above.

Now the inevitable Fibonacci example. The Fibonacci numbers are defined by $F_0 = 0, F_1 = 1, F_n = F_{n-1} + F_{n-2}$, or $F_n - F_{n-1} - F_{n-2} = 0$ for $n \ge 2$. Following the previous procedure, we see that

$$\sum_{n=0}^{\infty} F_n z^n = \frac{F_0 + (F_1 - F_0)z}{1-z-z^2} = \frac{z}{1-z-z^2}.$$

Since $z^2 - z - 1 = (z-\phi)(z-\bar{\phi})$, where $\phi = \frac{1+\sqrt{5}}{2} \approx 1.618$ and $\bar{\phi} = \frac{1-\sqrt{5}}{2} \approx -.618$, there exist constants c_j so that $F_n = c_1\phi^n + c_2\bar{\phi}^n$. The initial conditions imply $c_1 + c_2 = 0$ and $c_1\phi + c_2\bar{\phi} = 1$, yielding the *Binet formula*:

$$F_n = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right). \tag{3.23}$$

Since $|\bar{\phi}| < 1$, we see that $F_n \approx \frac{\phi^n}{\sqrt{5}}$; in fact, F_n is the closest integer to $\frac{\phi^n}{\sqrt{5}}$ for $n \ge 0$. It is also easy to extend the definition of the Fibonacci sequence to negative n; the equation $\phi\bar{\phi} = -1$ is instrumental in showing that $F_{-n} = (-1)^{n-1}F_n$.

It follows from the geometric series that

$$\begin{aligned} \sum_{n=0}^{\infty} F_n z^n = \frac{z}{1-z-z^2} &= z\sum_{m=0}^{\infty}(z+z^2)^m = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\binom{i+j}{j} z\cdot z^i(z^2)^j \\ \implies F_n &= \sum_{j\ge 0}\binom{n-1-j}{j}. \end{aligned} \tag{3.24}$$

(The binomial coefficient shuts off the final sum at $j = \lfloor \frac{n-1}{2} \rfloor$.) This formula lets you find the Fibonacci numbers by summing along a slope of Pascal's triangle, and will show up in the next chapter.

Closely related are the *Lucas* numbers, defined by $L_0 = 2, L_1 = 1$, and $L_n = L_{n-1} + L_{n-2}$ for $n \geq 2$. Since (L_n), (F_n) and (F_{n+1}) are three sequences with the same second-order recurrence, they are linearly dependent. Thus, there exist constants c_i so that $c_1 L_n + c_2 F_n + c_3 F_{n+1} = 0$, and (F_n) and (F_{n+1}) are not proportional, so we may take $c_1 = 1$. Putting $n = 0, 1$, we see that

$$\begin{aligned} 2 + c_3 = 1 + c_2 + c_3 = 0 &\implies c_2 = 1, c_3 = -2 \\ &\implies L_n = 2F_{n+1} - F_n = F_{n-1} + F_{n+1}. \end{aligned}$$

Similarly, (F_n), (L_n) and (L_{n+1}) are linearly dependent, and $F_n = \frac{1}{5}(2L_{n+1} - L_n)$. The similarity of coefficients is not accidental; it's an exercise that for fixed α, β, m,

$$L_{n+m} = \alpha F_{n+1} + \beta F_n \iff F_{n+m} = \frac{1}{5}\left(\alpha L_{n+1} + \beta L_n\right).$$

For a fixed positive integer m, there must be a dependence among (F_n), (F_{n+1}) and (F_{n+m}). Taking $n = 0, 1$ in the equation $F_{n+m} = \alpha F_n + \beta F_{n+1}$ implies that $F_m = \beta, F_{m+1} = \alpha + \beta$, so it's easy to derive the *Fibonacci addition formula*:

$$F_{n+m} = (F_{m+1} - F_m)F_n + F_m F_{n+1} = F_{m+1}F_n + F_m F_{n-1}. \tag{3.25}$$

It is also easy to show that $L_n = \phi^n + \bar{\phi}^n$ and $F_n L_n = F_{2n}$. Finally, observe that

$$\begin{aligned} (x+y)^n + (x-y)^n &= 2\sum_{k \geq 0} \binom{n}{2k} x^{n-2k} y^{2k}; \\ (x+y)^n - (x-y)^n &= 2\sum_{k \geq 0} \binom{n}{2k+1} x^{n-2k-1} y^{2k+1}. \end{aligned} \tag{3.26}$$

Taking $x = \frac{1}{2}$ and $y = \frac{\sqrt{5}}{2}$, we find that $x + y = \phi$ and $x - y = \bar{\phi}$, and in view of the formulas for F_n and L_n, it follows from the binomial theorem that

$$L_n = \frac{1}{2^{n-1}} \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} 5^k, \qquad F_n = \frac{1}{2^{n-1}} \sum_{k=0}^{\lfloor (n-1)/2 \rfloor} \binom{n}{2k+1} 5^k. \tag{3.27}$$

It does not seem to be intuitively obvious that the sums should be divisible by 2^{n-1}, nor that the ratio of the sums should approach $\sqrt{5}$ as $n \to \infty$. It takes nothing away from the romance of Fibonacci numbers to observe that many of their properties are similar to those satisfied by *any* sequence satisfying the second order linear recurrence; $a(n) = \alpha a(n-1) + \beta a(n-2)$. This is especially true when $\beta = \pm 1$ (so the roots of the characteristic equation are reciprocals or nearly so), and when $a(0) = 0, a(1) = 1$.

3.4. **Two basic Stern recurrences.** In this section, we make a first pass at answering two basic questions about the Stern sequence: how many n have the property that $3 \mid s(n)$, and what is the behavior of $\sum s(n)^2$?

Recalling (1.76) and (1.77), $A(3,0)$ denotes the set of n for which $3 \mid s(n)$, $U(r;3,0)$ is the number of elements of $A(3,0)$ in I_r and $T(n;3,0)$ is the number of elements of $A(3,0)$ which are $\le n$. We showed in Theorem 1.7 that $0 < n \in A(3,0)$ if and only if $2n, 8n \pm 5, 8n \pm 7 \in A(3,0)$. We set $a_r = U(r;3,0)$ for short.

We can easily compute the first few values of $a_r := U(r;3,0)$; namely, $a_0 = a_1 = 0, a_2 = a_3 = a_4 = 2, a_5 = 10, a_6 = 18$.

Theorem 3.3. *If $r \ge 3$, then*

$$a_r = a_{r-1} + 4a_{r-3}. \tag{3.28}$$

Proof. Since $s(2^r) = 1$, no power of 2 is in $A(3,0)$. By (1.81), the number of n in $A(3,0)$ in the five "covering congruences" of 0 (mod 2), 5 (mod 8), -5 (mod 8), 7 (mod 8), -7 (mod 8) is equal to $a_{r-1}, a_{r-3}, a_{r-3}, a_{r-3}, a_{r-3}$ respectively. □

The characteristic equation of (3.28) is $z^3 - z^2 - 4 = (z-2)(z^2+z+2) = (z-2)(z-\mu)(z-\bar{\mu})$, where

$$\mu = \frac{-1+\sqrt{7}i}{2} \approx -.5 + 1.323i, \quad \bar{\mu} = \frac{-1-\sqrt{7}i}{2} \approx -.5 - 1.323i. \tag{3.29}$$

It follows that $a_r = c_1 2^r + c_2\mu^r + c_3\bar{\mu}^r$, where

$$\begin{aligned} 0 &= c_1 + c_2 + c_3 \\ 0 &= 2c_1 + \mu c_2 + \bar{\mu} c_3 \\ 2 &= 4c_1 + \mu^2 c_2 + \bar{\mu}^2 c_3. \end{aligned} \tag{3.30}$$

Since $|\mu| = |\bar{\mu}| = \sqrt{2}$, the asymptotic growth of a_r is determined by c_1. Although (3.30) is easy to solve, there is a trick to computing c_1 directly. Observe that μ and $\bar{\mu}$ are the roots of $z^2 + z + 2 = 0$. Thus, if we multiply the rows above by 2, 1, and 1, successively and add, we find $2 = 8c_1$, so $c_1 = \frac{1}{4}$. A routine computation, the rest of which we omit, shows that

$$a_r = \frac{1}{4}\cdot 2^r + \left(\frac{-7+5\sqrt{7}i}{56}\right)\mu^r + \left(\frac{-7-5\sqrt{7}i}{56}\right)\bar{\mu}^r. \tag{3.31}$$

It follows by a routine (but easy-to-get-wrong) computation using $\mu\bar{\mu} = 2$ that:

$$\begin{aligned} T(2^r;3,0) = \sum_{k=0}^{r-1} a_k &= \frac{2^r}{4} + \left(\frac{i}{4\sqrt{7}}\right)\cdot(\bar{\mu}\mu^r - \mu\bar{\mu}^r) + \frac{1}{2} \\ &= \frac{2^r}{4} + \left(\frac{i}{2\sqrt{7}}\right)\cdot(\mu^{r-1} - \bar{\mu}^{r-1}) + \frac{1}{2}. \end{aligned} \tag{3.32}$$

As $|\mu| = |\bar{\mu}| = \sqrt{2}$, it follows that $|T(2^r; 3, 0) - \frac{2^r}{4}| = \mathcal{O}(2^{r/2})$. In Theorem 3.12 we prove the stronger estimate that $T(n; 3, 0) = \frac{n}{4} + \mathcal{O}(n^{1/2})$.

Another way to look at these equations is to write

$$(\sqrt{2}) \cdot \frac{-1 \pm \sqrt{7}i}{2\sqrt{2}} = \sqrt{2}e^{\pm i\alpha}, \qquad \frac{-7 \pm 5\sqrt{7}i}{56} = \frac{1}{\sqrt{14}} \cdot e^{\pm i\beta}. \tag{3.33}$$

Then (3.31) becomes

$$a_r = \frac{1}{4} \cdot 2^r + \frac{2^{r/2}}{\sqrt{14}} \cdot \left(e^{i(r\alpha+\beta)} + e^{-i(r\alpha+\beta)}\right) = \frac{1}{4} \cdot 2^r + 2^{r/2}\sqrt{2/7}\cos(r\alpha + \beta). \tag{3.34}$$

Niven's Theorem states that if $\frac{\theta}{\pi}$ and $\cos(\theta)$ are both rational, then $2\cos(\theta) \in \mathbb{N}$. It follows from (3.33) that $\cos(2\alpha) = \frac{3}{4}$, hence $\frac{\alpha}{2\pi}$ is irrational. It follows that the values of the sequence $(\cos(r\alpha + \beta))$ are dense in $[-1, 1]$; the coefficient of the error term in (3.34) thus gets arbitrarily close to $\sqrt{2/7} \approx .5345$.

We turn to the second question. We saw in (1.33) that $\sum_{n=2^r}^{2^{r+1}-1} s(n) = 3^r$. What can one say about $\sum_{n=2^r}^{2^{r+1}-1} s(n)^2$? It is helpful to make a more general definition. For integers $u, v \geq 0$, let

$$m_{u,v}(r) := \sum_{n=2^r}^{2^{r+1}-1} s(n)^u s(n+1)^v. \tag{3.35}$$

A quick lemma uses the row-reflection property (1.19):

Lemma 3.4. *For all u, v, r, $m_{u,v}(r) = m_{v,u}(r)$.*

Proof. A reparameterization via $m + n = 3 \cdot 2^r - 1 = 2^{r+1} + 2^r - 1$ shows that

$$m_{u,v}(r) = \sum_{n=2^r}^{2^{r+1}-1} s(n)^u s(n+1)^v = \sum_{m=2^{r+1}-1}^{2^r} s(3 \cdot 2^r - m - 1)^u s(3 \cdot 2^r - m)^v$$
$$= \sum_{m=2^r}^{2^{r+1}-1} s((m+1)^*)^u s(m^*)^v = \sum_{m=2^r}^{2^{r+1}-1} s((m+1))^u s(m)^v = m_{v,u}(r).$$

□

Lemma 3.5. *The following family of recurrences hold for $r \geq 0$:*

$$m_{u,v}(r+1) = \sum_{k=0}^{v} \binom{v}{k} m_{u+k,v-k}(r) + \sum_{\ell=0}^{u} \binom{u}{\ell} m_{u-\ell,v+\ell}(r). \tag{3.36}$$

Proof. The general reindexing

$$\sum_{n=2a}^{2b-1} f(n) = \sum_{n=a}^{b-1} f(2n) + \sum_{n=a}^{b-1} f(2n+1), \tag{3.37}$$

applied to $a = 2^r, b = 2^{r+1}$, implies that

$$m_{u,v}(r+1) = \sum_{n=2^r}^{2^{r+1}-1} s(2n)^u s(2n+1)^v + \sum_{n=2^r}^{2^{r+1}-1} s(2n+1)^u s(2n+2)^v$$
$$= \sum_{n=2^r}^{2^{r+1}-1} s(n)^u (s(n)+s(n+1))^v + \sum_{n=2^r}^{2^{r+1}-1} (s(n)+s(n+1))^u s(n+1)^v$$

Expansion by the binomial theorem leads to (3.36). □

Theorem 3.6. *For each integer p, the sequences* $(m_{i,p-i}(r))$, $0 \le i \le p$, *satisfy the same linear recurrence of order* $\lfloor \frac{p}{2} \rfloor + 1$.

Proof. There are $p+1$ sequences $(m_{i,p-i}(r))$, and (3.36) shows that they satisfy a matrix linear recurrence. But Lemma 3.4 shows that we can rewrite (3.36) so as to limit our attention to the $\lfloor \frac{p}{2} \rfloor + 1$ sequences $(m_{i,p-i}(r))$ with $i \ge p/2$. □

In particular, Lemmas 3.4 and 3.5 imply that $m_{2,0}(r) = m_{0,2}(r)$ and

$$\begin{gathered} m_{2,0}(r+1) = m_{2,0}(r) + m_{2,0}(r) + 2m_{1,1}(r) + m_{0,2}(r), \\ m_{1,1}(r+1) = m_{2,0}(r) + m_{1,1}(r) + m_{1,1}(r) + m_{0,2}(r) \\ \implies \begin{pmatrix} m_{2,0}(r+1) \\ m_{1,1}(r+1) \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} m_{2,0}(r) \\ m_{1,1}(r) \end{pmatrix}. \end{gathered} \tag{3.38}$$

The characteristic polynomial of $\left[\begin{smallmatrix} 3 & 2 \\ 2 & 2 \end{smallmatrix}\right]$ is $\lambda^2 - 5\lambda + 2$, which has roots

$$\nu = \frac{5+\sqrt{17}}{2} \approx 4.562, \qquad \bar{\nu} = \frac{5-\sqrt{17}}{2} \approx .438. \tag{3.39}$$

The initial conditions are: $m_{2,0}(0) = s(1)^2 = 1$, $m_{2,0}(1) = s(2)^2 + s(3)^2 = 5$. After some computations, $m_{2,0}(r)$ simplifies to:

Theorem 3.7.

$$m_{2,0}(r) = \frac{1}{\sqrt{17}} \cdot (\nu^{r+1} - \bar{\nu}^{r+1}). \tag{3.40}$$

Since $\bar{\nu} \in (0,1)$, it follows that $m_{2,0}(r) = \lfloor \nu^{r+1}/\sqrt{17} \rfloor$. The formula for $m_{1,1}(r)$ can be found from the relation, $m_{1,1}(r+1) = m_{2,0}(r+1) - m_{2,0}(r)$. For later reference, we observe that $s(2^r) = s(2^{r+1}) = 1$ implies that

$$m_{u,0}(r) = m_{0,u}(r) = \sum_{n \in I_r}^{*} s(n)^u. \tag{3.41}$$

This is not true for $m_{u,v}(r)$ when $uv > 0$ because $s(2^r - 1) \ne s(2^{r+1} + 1)$.

3.5. **A summation technique.** Recall the "trapezoidal sum" from (1.28):

$$\sum_{n=a}^{b}{}^{*} f(n) = \sum_{n=a}^{b} f(n) - \frac{f(a)+f(b)}{2} = \sum_{n=a}^{b-1} \frac{f(n)+f(n+1)}{2};$$

these expressions are additive; c.f. (1.29).

For $2^r m \le n \le 2^r(m+1)$, $s(n)$ is easily expressible in terms of $s(m)$ and $s(m+1)$. For this reason, we define the sequence $S(f;m) = (S(f;m,r))$ by

$$S(f;m,r) := \sum_{n=2^r m}^{2^r(m+1)}{}^{*} f(n). \tag{3.42}$$

Lemma 3.8. *Suppose that there is a finite set of sequences $\{(a_j(r)) : 1 \le j \le e\}$ with the property that, for given f and each $m \in \mathbb{N}$, $S(f;m,r) = a_{j_m}(r)$ for some $j_m, 1 \le j_m \le e$. Then there exists $d \le e$ and c_ℓ, $1 \le \ell \le d$, so that, for all m:*

$$S(f;m,d) + \sum_{\ell=1}^{d} c_j S(f;m,d-\ell) = 0. \tag{3.43}$$

Proof. The $e+1$ e-tuples $v(r) := (a_1(r), \dots a_e(r))$, $0 \le r \le e$, are linearly dependent, so $\sum_{r=0}^{e} \lambda_r v(r) = 0$ for some non-zero λ_r. Let d be the largest r so that $\lambda_r \ne 0$ and then let $c_j = \lambda_{d-j}/\lambda_d$ for $0 \le j \le d$. □

When there is no ambiguity, for a sequence $(a(r))$, we define

$$Y(a;r) := a(r) + \sum_{\ell=1}^{d} c_j a(r-\ell) = 0, \tag{3.44}$$

so that (3.43) is simply $Y(S(f;m);d) = 0$.

Lemma 3.9. *For all (f,m,r), we have*

$$S(f;m,r+1) = S(f;2m,r) + S(f;2m+1,r). \tag{3.45}$$

Proof. This is an application of (1.29); simply reinterpret

$$\sum_{n=2^{r+1}m}^{2^{r+1}(m+1)}{}^{*} f(n) = \sum_{n=2^r(2m)}^{2^r(2m+1)}{}^{*} f(n) + \sum_{n=2^r(2m+1)}^{2^r(2m+2)}{}^{*} f(n). \tag{3.46}$$

□

Lemma 3.10. *If $Y(S(f,m);d) = 0$ for all m, then for all $r \ge d$,*

$$Y(S(f;m);r) = 0. \tag{3.47}$$

Proof. We prove (3.47) by induction on r; the base case is (3.43). Assuming that (3.47) holds, we find that

$$Y(S(f;m);r+1) = Y(S(f;2m);r) + Y(S(f;2m+1);r) \tag{3.48}$$

by repeated application of Lemma 3.9, and this establishes the inductive step. □

Theorem 3.11. *Suppose $Y(S(f,m);d) = 0$ for all m and for $t \in \mathbb{N}$, let*

$$A_t(r) = \sum_{n=0}^{2^r t}{}^{*} f(n). \tag{3.49}$$

Then for $r \geq d$, the sequence $(A_t(r))$ satisfies

$$Y(A_t; r) = 0 \tag{3.50}$$

Proof. Suppose $t = 2^{r_1} + 2^{r_2} + \cdots + 2^{r_k}$, with $r_1 > r_2 > \cdots > r_k$, and let $N_0 = 0$ and $N_j = 2^{r_1} + \cdots + 2^{r_j} = N_{j-1} + 2^{r_j}$ for $j = 1, \ldots, k$, so that $t = N_k$. Further, for $1 \leq j \leq k$, let $M_j = 2^{-r_j} N_{j-1}$, so that $N_{j-1} = 2^{r_j} M_j$ and $N_j = 2^{r_j}(M_j + 1)$. Then

$$\begin{aligned} A_t(r) &= \sum_{n=0}^{2^r t}{}^{*} f(n) = \sum_{j=1}^{k} \sum_{n=2^r N_{j-1}}^{2^r N_j}{}^{*} f(n) \\ &= \sum_{j=1}^{k} \left(\sum_{n=2^{r+r_j} M_j}^{2^{r+r_j}(M_j+1)}{}^{*} f(n) \right) = \sum_{j=1}^{k} S(f; M_j, r + r_j). \end{aligned} \tag{3.51}$$

is a sum of sequences, each of which satisfies (3.50) by Lemma 3.10. □

The hypotheses of Theorem 3.11 might seem to be formidably strong, and they usually are, unless the function f being summed relates to the Stern sequence. For example, suppose $f(n) = s(n)$ itself, and write $f(m) = a$ and $s(m+1) = b$. Then

$$\begin{aligned} S(f; m, 0) &:= \frac{s(m)}{2} + \frac{s(m+1)}{2} = \frac{a+b}{2}, \\ S(f; m, 1) &:= \frac{s(2m)}{2} + s(2m+1) + \frac{s(2m+2)}{2} \\ &= \frac{a + 2(a+b) + b}{2} = \frac{3a+3b}{2}. \end{aligned} \tag{3.52}$$

That is, $S(f; m, 1) = 3S(f; m, 0)$, and so, as we've already seen in Lemma 1.3,

$$\sum_{n=0}^{2N}{}^{*} s(n) = 3 \sum_{n=0}^{N}{}^{*} s(n). \tag{3.53}$$

We will now apply Theorem 3.11 to the two situations of §3.4.

3.6. **The Stern sequence mod 3.** Suppose $f = \chi_{A(d,i)}$, so for $T \subseteq \mathbb{N}$, the expression $\sum_{n \in T} f(n)$ counts the number of n in T for which $s(n) \equiv i \pmod{d}$. By (1.12),

$$s(2^r m + k) = s(2^r - k)s(m) + s(k)s(m+1), \quad 0 \leq k \leq 2^r.$$

If $(s(m_1), s(m_1+1)) \equiv (s(m_2), s(m_2+1)) \pmod{d}$, then it follows that $s(2^r m_1 + k) \equiv s(2^r m_2 + k) \pmod{d}$ for $0 \leq k \leq 2^r$, so that

$$S(\chi_{A(d,i)}; m_1, r) = S(\chi_{A(d,i)}; m_2, r). \tag{3.54}$$

Since $s(m) \equiv s(m+1) \equiv 0 \pmod{d}$ is impossible, this means there are at most d^2-1 different sequences $(S(\chi_{A(d,i)}; m, r))$, parameterized by $(i \pmod{d}, j \pmod{d})$. The hypothesis of Lemma 3.8 is satisfied and so Theorem 3.11 applies.

We now apply this in detail to the case $(d,i) = (3,0)$, $f = \chi = \chi_{A_{3,0}}$ and keep this notation for the rest of this section; we return to the general case in a later chapter.

There are, conceivably eight different cases, depending on the congruence classes mod 3; however, two observations reduce the number of different sequences: first, if $(s(m'), s(m'+1)) \equiv (s(m+1), s(m)) \pmod{3}$, then $S(\chi; m, r) = S(\chi; m', r)$ by mirror symmetry; second, if $(s(m'), s(m'+1)) \equiv -(s(m), s(m+1)) \pmod{3}$, then $s(2^r m' + k) \equiv -s(2^r m + k) \pmod{3}$, so $\chi(2^r m' + k) = \chi(2^r m + k)$ and again, $S(\chi; m, r) = S(\chi; m', r)$.

Let $S_{ij}(r) = S(\chi; m, r)$ in the case that $(s(m), s(m+1)) \equiv (i,j) \pmod{3}$. Thus, $S_{01}(r) = S_{02}(r) = S_{10}(r) = S_{20}(r)$, $S_{11}(r) = S_{22}(r)$ and $S_{12}(r) = S_{21}(r)$, so there are only three different sequences. A look at the initial conditions

$$
\begin{array}{ll}
S_{01}(r): & \begin{array}{ccccccccc} 0 & & & & & & & & 1 \\ 0 & & & & 1 & & & & 1 \\ 0 & & 1 & & 1 & & 2 & & 1 \\ 0 & 1 & 1 & 2 & 1 & 3 & 2 & 3 & 1 \end{array} \\
S_{11}(r): & \begin{array}{ccccccccc} 1 & & & & & & & & 1 \\ 1 & & & & 2 & & & & 1 \\ 1 & & 3 & & 2 & & 3 & & 1 \\ 1 & 4 & 3 & 5 & 2 & 5 & 3 & 4 & 1 \end{array} \\
S_{12}(r): & \begin{array}{ccccccccc} 1 & & & & & & & & 2 \\ 1 & & & & 3 & & & & 2 \\ 1 & & 4 & & 3 & & 5 & & 2 \\ 1 & 5 & 4 & 7 & 3 & 8 & 5 & 7 & 2 \end{array}
\end{array}
\tag{3.55}
$$

shows that

$$
\begin{aligned}
(S_{01}(0), S_{11}(0), S_{12}(0)) &= (\tfrac{1}{2}, 0, 0), \\
(S_{01}(1), S_{11}(1), S_{12}(1)) &= (\tfrac{1}{2}, 0, 1), \\
(S_{01}(2), S_{11}(2), S_{12}(2)) &= (\tfrac{1}{2}, 2, 1), \\
(S_{01}(3), S_{11}(3), S_{12}(3)) &= (\tfrac{5}{2}, 2, 1).
\end{aligned}
\tag{3.56}
$$

It follows, unsurprisingly in view of Theorem 3.3, that

$$
S_{ij}(3) = S_{ij}(1) + 4S_{ij}(0). \tag{3.57}
$$

This derivation of the recurrence uses no combinatorial information; instead, its existence is forced by dimensional considerations. An alternative approach to this method would apply Lemma 3.9 to observe that $S_{ij}(r+1) = S_{ik}(r) + S_{kj}(r)$, where $i + j \equiv k \pmod{3}$. (This matrix recurrence is the technique we will use for general bases in a later chapter.)

In the notation of §3.4, we have $S_{11}(r) = a_r$, $S_{12}(r) = \frac{1}{2}a_{r+1}$ and $S_{01}(r) = T(2^r;3,0) = \frac{1}{4}a_{r+2}$, and further

$$\begin{aligned} \left|S_{11}(r) - \tfrac{1}{4}\cdot 2^r\right| &\le \sqrt{\tfrac{2}{7}}\cdot(\sqrt{2})^r, \\ \left|S_{12}(r) - \tfrac{1}{4}\cdot 2^r\right| &\le \sqrt{\tfrac{1}{7}}\cdot(\sqrt{2})^r, \\ \left|S_{01}(r) - \tfrac{1}{4}\cdot 2^r\right| &\le \sqrt{\tfrac{1}{14}}\cdot(\sqrt{2})^r. \end{aligned} \tag{3.58}$$

Thus, it follows that for all (m, r),

$$\left|S(\chi;m,r) - \frac{1}{4}\cdot 2^r\right| \le \sqrt{\tfrac{2}{7}}\cdot(\sqrt{2})^r. \tag{3.59}$$

Theorem 3.12.

$$\left|T(N;3,0) - \frac{N}{4}\right| = \mathcal{O}(N^{1/2}). \tag{3.60}$$

Proof. First observe that

$$T(N;3,0) = \sum_{n=0}^{2^N}{}^{*}\chi(n) + \frac{\chi(0)+\chi(N)}{2},$$

so the difference between the two is either $\frac{1}{2}$ or 1, depending on whether $N \in A(3,0)$. This will not affect the asymptotics. Next, suppose that $N = 2^{r_1} + \cdots + 2^{r_k}$. Then

$$\sum_{n=0}^{N}{}^{*}\chi(n) = \sum_{j=1}^{k} S(\chi;M_j,r_j). \tag{3.61}$$

Therefore,

$$\begin{aligned} \left|T(3,0;N) - \frac{N}{4}\right| &\le 1 + \left|\sum_{n=0}^{N}{}^{*}\chi(n) - \frac{N}{4}\right| \\ \le 1 + \left|\sum_{j=1}^{k} S(\chi;M_j,r_j) - \frac{1}{4}\sum_{j=1}^{k} 2^{r_j}\right| &\le 1 + \sum_{j=1}^{k}\left|S(\chi;M_j,r_j) - \frac{1}{4}\cdot 2^{r_j}\right| \\ \le 1 + \sum_{j=1}^{k}\sqrt{\tfrac{2}{7}}\cdot(\sqrt{2})^{r_j} \le 1 + \sqrt{\tfrac{2}{7}}\left(\sum_{\ell=0}^{r_1}(\sqrt{2})^{\ell}\right) &= 1 + \sqrt{\tfrac{2}{7}}\cdot\frac{(\sqrt{2})^{r_1+1}-1}{\sqrt{2}-1} \\ < 1 + \sqrt{\tfrac{2}{7}}(\sqrt{2}+1)(\sqrt{2})(N^{1/2}-1) &= \frac{2\sqrt{2}+2}{\sqrt{7}}\cdot N^{1/2} < 2N^{1/2}. \end{aligned} \tag{3.62}$$

□

3.7. **How fast does the Stern sequence grow?** We return briefly to Theorem 3.7. Using the general methods, we can obtain the recurrence more easily. Writing $s(m) = a$ and $s(m+1) = b$, with $f(n) = s(n)^2$, we have

$$\begin{aligned}
S(f;m,0) &= \frac{f(m)}{2} + \frac{f(m+1)}{2} = \frac{a^2+b^2}{2},\\
S(f;m,1) &= \frac{f(2m)}{2} + f(2m+1) + \frac{f(2m+2)}{2}\\
&= \frac{a^2+2(a+b)^2+b^2}{2} = \frac{3a^2+4ab+3b^2}{2},\\
S(f;m,2) &= \frac{f(4m)}{2} + f(4m+1) + f(4m+2) + f(4m+3) + \frac{f(4m+4)}{2}\\
&= \frac{3a^2+4ab+3b^2}{2} + (2a+b)^2 + (a+2b)^2 = \frac{13a^2+20ab+13b^2}{2}.
\end{aligned} \tag{3.63}$$

It may be readily verified that

$$S(f;m,2) - 5S(f;m,1) + 2S(f,m,0) = 0. \tag{3.64}$$

Since $S(f;1,r) = m_{2,0}(r)$, this is a much faster derivation of Theorem 3.7; however, you don't get the recurrence for $g(N)$, which is *not* a sum of this kind.

We now discuss, without tremendous detail,

$$m_{k,0}(r) := \sum_{n=2^r}^{2^{r+1}}{}^{*} s(n)^k = S(s(n)^k;1,r). \tag{3.65}$$

As above, we suppose that $s(m) = a$ and $s(m+1) = b$, so that

$$\begin{aligned}
S(f;m,0) &= \frac{1}{2}\left(a^k + b^k\right);\\
S(f;m,1) &= S(f;m,0) + (a+b)^k;\\
S(f;m,2) &= S(f;m,1) + (2a+b)^k + (a+2b)^k.
\end{aligned} \tag{3.66}$$

As we have already seen, when $k = 1$, $S(f;m,1) = 3S(f;m,0)$ and $m_{1,0}(r) = 3^r$. When $k = 3$, there is an unusually pleasant recurrence:

$$S(f;m,2) - 7S(f;m,1) = (2a+b)^3 + (a+2b)^3 - 6(a+b)^3 - 3(a^3+b^3) = 0.$$

It follows that $m_{3,0}(r+2) = 7m_{3,0}(r+1)$ for $r \ge 0$, and that

$$m_{3,0}(0) = 1, \qquad m_{3,0}(r) = 9 \cdot 7^{r-1}, \qquad r \ge 1. \tag{3.67}$$

Note that this second-order recurrence has one root equal to zero and another way to express (3.67) would be as $m_{3,0}(r) = \frac{9}{7} \cdot 7^r - \frac{2}{7} \cdot 0^r$. Alternatively, Lemma 3.5 implies that $m_{3,0}(r) = m_{0,3}(r)$, $m_{2,1}(r) = m_{1,2}(r)$ and

$$\begin{aligned} m_{3,0}(r+1) &= 2m_{3,0}(r) + 3m_{2,1}(r) + 3m_{1,2}(r) + m_{0,3}(r) \\ m_{2,1}(r+1) &= m_{3,0}(r) + 2m_{2,1}(r) + 2m_{1,2}(r) + m_{0,3}(r) \\ \implies \begin{pmatrix} m_{3,0}(r+1) \\ m_{2,1}(r+1) \end{pmatrix} &= \begin{pmatrix} 3 & 6 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} m_{3,0}(r) \\ m_{2,1}(r) \end{pmatrix}. \end{aligned} \tag{3.68}$$

The characteristic polynomial of $\begin{smallmatrix} 3 & 6 \\ 2 & 4 \end{smallmatrix}$ is $\lambda^2 - 7\lambda$. The cases for $k = 4, 5$ are deferred to the solutions to the second set of exercises.

Is there a point to this? We might ask the vague question: how large is $s(n)$? Certainly individual values vary quite a bit: if $2^r \le n \le 2^{r+1}$, $1 \le s(n) \le F_{r+2}$, as we have seen. However, the results on sums of powers suggest that there might be some regular behavior. Define the *t-th power mean* for $2^r \le n \le 2^{r+1}$:

$$M(r;t) := \left(\frac{1}{2^r} \sum_{n=2^r}^{2^{r+1}}{}^{*} s(n)^t \right)^{1/t}. \tag{3.69}$$

We have already seen that $M(r,1) = \left(\frac{3}{2}\right)^r$, which suggests the definition

$$L(r;t) := M(r,1)^{1/r}. \tag{3.70}$$

In this notation, $L(r;1) = \frac{3}{2}$. This subject ties in with some traditional analysis, and a classical slick inequality proof.

Theorem 3.13. *If $x_k > 0$ for $1 \le k \le n$, then the function*

$$M(t) := \left(\frac{1}{n} \sum_{k=1}^{n} x_k^t \right)^{1/t} \tag{3.71}$$

is increasing for $t \ge 0$, and $\lim_{t \to \infty} M(t) = \max x_k$.

Proof. We first consider the auxiliary function

$$\Phi(t) := \log \left(\sum_{k=1}^{n} x_k^t \right), \tag{3.72}$$

and commit calculus, finding by routine computation that

$$\Phi''(t) = \frac{\left(\sum_{k=1}^{n} x_k^t \right) \left(\sum_{k=1}^{n} x_k^t (\log x_k)^2 \right) - \left(\sum_{k=1}^{n} x_k^t \log x_k \right)^2}{\left(\sum_{k=1}^{n} x_k^t \right)^2}. \tag{3.73}$$

The numerator of (3.73) is non-negative by Cauchy-Schwartz, so Φ is convex for all t. This implies that

$$\Psi(t) := \frac{\Phi(t) - \Phi(0)}{t}$$

is an increasing function for $t \geq 0$, as is $e^{\Psi(t)}$. But

$$\Psi(t) = \frac{\log\left(\sum_{k=1}^{n} x_k^t\right) - \log\left(\sum_{k=1}^{n} x_k^0\right)}{t} = \frac{1}{t}\log\left(\frac{1}{n}\sum_{k=1}^{n} x_k^t\right) \implies e^{\Psi(t)} = M(t). \tag{3.74}$$

If $M = \max x_k$, then $M^t \leq \sum_{k=1}^{n} x_k^t \leq nM^t$, so $n^{-1/t}M \leq M(t) \leq M$. □

For example,

$$\lim_{t\to\infty} L(r,t) = F_{r+2}^{1/r}.$$

Lemma 3.14. *Suppose*

$$a_r = c\lambda^r + \sum_{j=1}^{t} g_j(r)\lambda_j^r, \tag{3.75}$$

where $c > 0$*,* $g_j(r)$ *is a polynomial in* r *and* $\lambda > |\lambda_j|$ *is a positive real. Then*

$$\lim_{r\to\infty} a_r^{1/r} = \lambda. \tag{3.76}$$

Proof. Since $a_r = \lambda^r b_r$, where $\lim b_r = c > 0$, and since $c^{1/r} \to 1$, the proof is immediate. □

It follows from Lemma 3.14 that

$$\lim_{r\to\infty}\left(\lim_{t\to\infty} L(r,t)\right) = \lim_{r\to\infty} F_{r+2}^{1/r} = \Phi = \frac{1+\sqrt{5}}{2} \approx 1.6180. \tag{3.77}$$

A more interesting, and open, question, is the computation of

$$\lim_{t\to\infty}\left(\lim_{r\to\infty} L(r,t)\right). \tag{3.78}$$

which, based on the lemma, is closely related to the behavior of the recurrences that we've found. In fact, taking into account earlier equations and the numerical values of the roots not presented here explicitly (for $t = 4, 5, 6, 7$, and providing an uninteresting hint to the exercises), we have

$$\begin{gathered} \lim_{r\to\infty} L(r,1) = \tfrac{3}{2} = 1.5, \quad \lim_{r\to\infty} L(r,2) = \left(\tfrac{5+\sqrt{17}}{4}\right)^{1/2} \approx 1.5102, \\ \lim_{r\to\infty} L(r,3) = \left(\tfrac{7}{2}\right)^{1/3} \approx 1.5183, \quad \lim_{r\to\infty} L(r,4) \approx 1.5249, \\ \lim_{r\to\infty} L(r,5) \approx 1.5305, \lim_{r\to\infty} L(r,6) \approx 1.5354, \lim_{r\to\infty} L(r,7) \approx 1.5396. \end{gathered} \tag{3.79}$$

The limiting behavior is far from clear.

3.8. **Sylvester's Theorem.** The material in this section is mostly adapted from the papers *On the length of binary forms*, to appear in Quadratic and Higher Degree Forms, (K. Alladi, M. Bhargava, D. Savitt, P. Tiep, eds.), Developments in Math. Springer, New York, http://arxiv.org/pdf/1007.5485.pdf. and *Sums of even powers of real linear forms*, Mem. Amer. Math. Soc., Volume 96, Number 463, March, 1992 (MR 93h.11043), both of which can be downloaded from my website.

We give an application of linear recurrences with no direct ties to Stern sequences, but which resonates with various number theoretic questions: namely, the representation of binary forms as a sum of powers of linear forms. The main result was proved by J. J. Sylvester in 1851.

The representation of quadratic forms as a sum of squares of linear forms is well understood and a standard part of linear algebra. This is less so for higher powers of linear forms, even though the simplest case (two variables) has been understood for more than 150 years.

Consider also the classical question of quadrature. Suppose $S \subseteq \mathbb{R}^n$ and measure μ are given. A *quadrature formula of strength* m *for* $(S, d\mu)$ is an exact formula

$$\int_S f\ d\mu = \sum_{k=1}^{r} \lambda_k f(t_k), \quad f \in \mathbb{R}[x_1, \dots, x_n], \quad \deg f \le m. \tag{3.80}$$

In this section we take the case of $S = [-1, 1] \subseteq \mathbb{R}^1$ and Lebesgue measure: (3.80) is a quadrature formula of strength m provided it holds for $f(t) = t^i$, $0 \le i \le m$, so

$$\frac{1 - (-1)^{m+1}}{m+1} = \int_{-1}^{1} t^i\ dt = \sum_{k=1}^{r} \lambda_k t_k^i, \qquad 0 \le i \le m. \tag{3.81}$$

If $r < m$, then (3.81) says that the successive "moments" must satisfy an r-th order linear recurrence. We can turn (3.81) into a single equation by constructing the appropriate generating function:

$$\begin{aligned} \int_{-1}^{1} (x + ty)^m\ dt &= \sum_{i=0}^{m} \binom{m}{i} \left(\int_{-1}^{1} t^i\ dt \right) x^{m-i} y^i \\ &= \sum_{i=0}^{m} \binom{m}{i} \left(\sum_{k=1}^{r} \lambda_k t_k^i \right) x^{m-i} y^i = \sum_{k=1}^{r} \lambda_k (x + t_k y)^m. \end{aligned} \tag{3.82}$$

In other words, a quadrature formula on an interval in $\mathbb{R}$ is the same as a representation of a binary forms of degree m as a sum of m-th powers of linear forms.

We consider binary *m-ic* forms with complex coefficients; that is, homogeneous polynomials of degree m in two variables, written as:

$$f(x, y) = \sum_{i=0}^{m} \binom{m}{i} a_i x^{m-i} y^i. \tag{3.83}$$

(It is both customary and convenient to factor the binomial coefficient out in the expression.) The *length* or *rank* of f is the smallest integer r with the property that there exists an expression:

$$f(x,y) = \sum_{k=1}^{r} \gamma_k(\alpha_k x + \beta_k y)^m. \tag{3.84}$$

One quick remark: if r in (3.84) is minimal, then the linear forms $\{\alpha_k x + \beta_k y\}$ will be pairwise non-proportional; otherwise, two could be combined. Under this restriction, we say that (3.84) is an *honest* representation.

If (3.84) is honest, then $\alpha_k = 0$ for at most one zero; hence after writing $c_k = \gamma_k \alpha_k^m$ and $\lambda_k = \beta_k/\alpha_k$,

$$f(x,y) = \sum_{k=1}^{r} c_k(x + \lambda_k y)^m, \quad \text{or} \quad f(x,y) = \sum_{k=1}^{r-1} c_k(x + \lambda_k y)^m + c_r y^m. \tag{3.85}$$

A comparison with (3.83) shows that (3.85) is equivalent to:

$$\begin{gathered} a_i = \sum_{k=1}^{r} c_k \lambda_k^i, \quad 0 \le i \le m \qquad \text{or} \\ a_i = \sum_{k=1}^{r-1} c_k \lambda_k^i \ \ (0 \le i \le m-1), \quad a_m = \sum_{k=1}^{r-1} c_k \lambda_k^m + c_r \end{gathered} \tag{3.86}$$

Theorem 3.15 (Sylvester). *Suppose f is given by (3.83) and suppose*

$$h(x,y) = \sum_{t=0}^{r} c_t x^{r-t} y^t = \prod_{j=1}^{r} (-\beta_j x + \alpha_j y) \tag{3.87}$$

is a given product of pairwise distinct linear factors. Then there exist $\lambda_k \in \mathbb{C}$ so that (3.84) holds if and only if

$$\begin{pmatrix} a_0 & a_1 & \cdots & a_r \\ a_1 & a_2 & \cdots & a_{r+1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m-r} & a_{m-r+1} & \cdots & a_m \end{pmatrix} \cdot \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_r \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}; \tag{3.88}$$

that is, if and only if

$$\sum_{t=0}^{r} a_{\ell+t} c_t = 0, \qquad \ell = 0, 1, \ldots, m-r. \tag{3.89}$$

If (3.88) holds, we say that h, as defined by (3.87) is a *Sylvester form* for p. Sylvester's Theorem in practice involes writing the successive "Hankel" matrices in (3.88) with increasing r until there exists one with a non-zero null vector $(c_0, \ldots, c_r)^t$ so that the resulting Sylvester form has distinct factors. Afterwards, the computation of the λ_k is routine.

Proof. First suppose that (3.84) holds. Then for $0 \le i \le m$,

$$a_i = \sum_{k=1}^{r} \lambda_k \alpha_k^{m-i} \beta_k^i \implies \sum_{t=0}^{r} a_{\ell+t} c_t = \sum_{k=1}^{r} \sum_{t=0}^{r} \lambda_k \alpha_k^{m-\ell-t} \beta_k^{\ell+t} c_t$$
$$= \sum_{k=1}^{r} \lambda_k \alpha_k^{m-\ell-r} \beta_k^{\ell} \sum_{t=0}^{r} \alpha_k^{r-t} \beta_k^t c_t = \sum_{k=1}^{r} \lambda_k \alpha_k^{m-\ell-r} \beta_k^{\ell}\ h(\alpha_k, \beta_k) = 0.$$

Now suppose that (3.85) holds and suppose first that $c_r \ne 0$. We may assume without loss of generality that $c_r = 1$ and that $\alpha_j = 1$ in (3.87), so that the β_j's are distinct. Define the *infinite* sequence $(\tilde{a}_i)$, $i \ge 0$, by:

$$\tilde{a}_i = a_i \quad \text{if} \quad 0 \le i \le r-1; \qquad \tilde{a}_{r+\ell} = -\sum_{t=0}^{r-1} \tilde{a}_{t+\ell} c_t \quad \text{for} \quad \ell \ge 0; \tag{3.90}$$

so that $(\tilde{a}_i)$ satisfies (3.86) and extends the finite sequence $(a_0, \dots, a_m)$:

$$\tilde{a}_i = a_i \quad \text{for} \quad i \le m. \tag{3.91}$$

Theorem 3.2 now implies that there exist λ_k so that for all i,

$$\tilde{a}_i = \sum_{k=1}^{r} \lambda_k \beta_k^i. \tag{3.92}$$

In particular,

$$\begin{gathered} f(x,y) = \sum_{i=0}^{m} \binom{m}{i} a_j x^{m-i} y^i = \\ \sum_{k=1}^{r} \lambda_k \sum_{i=0}^{m} \binom{m}{i} \beta_k^i x^{m-i} y^i = \sum_{k=1}^{r} \lambda_k (x + \beta_k y)^m, \end{gathered} \tag{3.93}$$

as claimed in (3.84).

If $c_r = 0$, then $c_{r-1} \ne 0$, because h has distinct factors. We may proceed as before, replacing r by $r-1$ and taking $c_{r-1} = 1$, so that (3.87) becomes

$$h(x,y) = \sum_{t=0}^{r-1} c_t x^{r-t} y^t = x \prod_{j=1}^{r-1} (y - \beta_j x). \tag{3.94}$$

Since $c_r = 0$, the system (3.88) can be rewritten as

$$\begin{pmatrix} a_0 & a_1 & \cdots & a_{r-1} \\ a_1 & a_2 & \cdots & a_r \\ \vdots & \vdots & \ddots & \vdots \\ a_{d-r} & a_{m-r+1} & \cdots & a_{m-1} \end{pmatrix} \cdot \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{r-1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

We may now argue as before, except that (3.91) becomes

$$\tilde{a}_i = a_i \quad \text{for} \quad i \le m-1, \qquad a_m = \tilde{a}_m + \lambda_m \tag{3.95}$$

for some λ_m, and (3.93) becomes

$$f(x,y) = \sum_{i=0}^{m} \binom{m}{i} a_j x^{m-i} y^i =$$

(3.96)
$$\lambda_r y^m + \sum_{k=1}^{r-1} \lambda_k \sum_{i=0}^{m} \binom{m}{i} \beta_k^i x^{m-i} y^i = \lambda_r y^m + \sum_{k=1}^{r-1} \lambda_k (x + \beta_k y)^m,$$

By (3.94), (3.96) meets the description of (3.84), completing the proof. □

In 1886, Gundelfinger studied the case where the Sylvester form h has repeated factors. The factor $(-\beta x + \alpha y)^\ell$ of h corresponds to a summand $q(x,y)(\alpha x + \beta y)^{d+1-\ell}$ in f, where q is an arbitrary form of degree $\ell - 1$.

The classical application of Sylvester's Theorem is to "canonical forms". If $m = 2s-1$ and $r = s$, then the matrix in (3.88) has dimensions $s \times (s+1)$ and so has a non-trivial null-vector; for a "general" f, the resulting form h has distinct factors, and so a general binary form of degree $2s-1$ has a representation as a sum of s $2s-1$-st powers of linear forms, which is unique up to a permutation of the summands. If $d = 2s$ and $r = s$, then the matrix in (3.88) is square, and since its determinant is not 0 in general, there is no corresponding h. However, for general f, for any α, there there exists $\lambda_0 = \lambda_0(\alpha)$ so that the Hankel matrix for $f - \lambda_0(x + \alpha y)^{2s}$ *does* have a non-trivial null vector and is a sum of s $2s$-th powers, hence general f is a sum of $s+1$ $2s$-th powers in infinitely many ways.

Without going into details, one can also define the K-*length* of a form f for any subfield $K \subseteq \mathbb{C}$ which contains the coefficients of f, and Sylvester's Theorem can be adapted to that case as well. Here is one example:

$$H(x,y) = 3x^5 - 20x^3y^2 + 10xy^4 = \binom{5}{0} \cdot 3\ x^5 + \binom{5}{1} \cdot 0\ x^4 y$$
$$+ \binom{5}{2} \cdot (-2)\ x^3y^2 + \binom{5}{3} \cdot 0\ x^2y^3 + \binom{5}{4} \cdot 2\ xy^4 + \binom{5}{5} \cdot 0\ y^5;$$

$$\begin{pmatrix} 3 & 0 & -2 & 0 \\ 0 & -2 & 0 & 2 \\ -2 & 0 & 2 & 0 \end{pmatrix} \cdot \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff (c_0, c_1, c_2, c_3) = r(0,1,0,1).$$

Thus, H has a *unique* Sylvester form of degree 3: $h(x,y) = y(x^2 + y^2)$, which factors as $y(y - ix)(y + ix)$. Accordingly, there exist $\lambda_k \in \mathbb{C}$ so that

$$H(x,y) = \lambda_1 x^5 + \lambda_2 (x + iy)^5 + \lambda_3 (x - iy)^5.$$

It may be checked that $\lambda_1 = \lambda_2 = \lambda_3 = 1$; this is how H was constructed.

To find representations for H of length 4, we consider (3.85) for H with $r = 4$:

$$H_4(H) \cdot (c_0, c_1, c_2, c_3, c_4)^t = (0,0)^t \iff 3c_0 - 2c_2 + 2c_4 = -2c_1 + 2c_3 = 0$$
$$\iff (c_0, c_1, c_2, c_3, c_4) = r_1(2,0,3,0,0) + r_2(0,1,0,1,0) + r_3(0,0,1,0,1),$$

hence $h(x,y) = r_1x^2(2x^2+3y^2) + y(x^2+y^2)(r_2x+r_3y)$. Given a field K, it is unclear whether there exist $\{r_\ell\}$ so that h splits into distinct factors over K. We have found such $\{r_\ell\}$ for small imaginary quadratic fields. For example, the choice $(r_1,r_2,r_3) = (1,0,2)$ gives $h(x,y) = (2x^2+y^2)(x^2+2y^2)$ and

$$24H(x,y) = 4(x+\sqrt{-2}y)^5 + 4(x-\sqrt{-2}y)^5 + (2x+\sqrt{-2}y)^5 + (2x-\sqrt{-2}y)^5.$$

Similarly, $(r_1,r_2,r_3) = (2,0,9)$ and $(2,0,-5)$ give $h(x,y) = (x^2+3y^2)(4x^2+3y^2)$ and $(x^2-y^2)(4x^2+5y^2)$, leading to representations for H of length 4 over $\mathbb{Q}(\sqrt{-3})$ and $\mathbb{Q}(\sqrt{-5})$. The simplest such representation we have found for $\mathbb{Q}(\sqrt{-6})$ uses $(r_1,r_2,r_3) = (8450,0,-104544)$ and

$$h(x,y) = (5x+12y)(5x-12y)(6\cdot 13^2x^2+33^2y^2).$$

We believe, but have not proved, that examples such as these exist for every imaginary quadratic field. A different theorem of Sylvester, from 1864, implies that H has *no* representation as a sum of fewer than five *real* fifth powers of linear forms.

As another example with number theory applications, consider the representations of $(xy)^k$ as a sum of $2k$-th powers of linear forms. The square Hankel matrix for $f(x,y) = \binom{2k}{k}x^ky^k$ has 1's on the NE-SW diagonal, and so is non-singular. Thus there is no representation of f as a sum of k $2k$-th powers of linear forms.

Let $\zeta_m = e^{2\pi i/m}$. It is easy to see that $\sum_{j=0}^{m-1}\zeta_m^{rj} = 0$ unless $m \mid r$, in which case it equals m. It turns out that the full set of minimal representations of $f(x,y) = \binom{2k}{k}x^ky^k$ as a sum of $(k+1)$ $2k$-th powers is

$$(k+1)\binom{2k}{k}x^ky^k = \sum_{j=0}^{k}(\zeta_{2k+2}^{j}wx + \zeta_{2k+2}^{-j}w^{-1}y)^{2k}, \qquad 0 \neq w \in \mathbb{C}. \tag{3.97}$$

Evaluate the right-hand side of (3.97) by expanding the powers:

$$\begin{aligned}\sum_{j=0}^{k}(\zeta_{2k+2}^{j}wx + \zeta_{2k+2}^{-j}w^{-1}y)^{2k} &= \sum_{j=0}^{k}\sum_{t=0}^{2k}\binom{2k}{t}\zeta_{2k+2}^{j(2k-t)-jt}w^{(2k-t)-t}x^{2k-t}y^t \\ &= \sum_{t=0}^{2k}\binom{2k}{t}w^{2k-2t}x^{2k-t}y^t\left(\sum_{j=0}^{k}\zeta_{k+1}^{j(k-t)}\right).\end{aligned} \tag{3.98}$$

Since the only multiple of $k+1$ in the set $\{k-t : 0 \le t \le 2k\}$ occurs for $t = k$, (3.98) reduces to the left-hand side of (3.97). The representations in (3.97) arise because the null-vectors of the resulting $(k-1)\times(k+1)$ Hankel matrix can only be $(c_0,0,\dots,0,c_{k+1})^t$ and $c_0x^{k+1}+c_{k+1}y^{k+1}$ is a Sylvester form when $c_0c_{k+1}\neq 0$. We state without proof that by making the change of variables in $(x,y)\mapsto(x-iy,x+iy)$ in (3.98) we obtain the expressions:

$$\binom{2k}{k}(x^2+y^2)^k = \frac{1}{k+1}\sum_{j=0}^{k}\left(\cos(\tfrac{j\pi}{k+1}+\theta)x + \sin(\tfrac{j\pi}{k+1}+\theta)y\right)^{2k}, \quad \theta\in\mathbb{C}. \tag{3.99}$$

When θ is real, these are related to the "Hilbert identities" used in solving the Waring problem. Even for real θ, the earliest instance of (3.99) in the literature seems to be by Friedman from 1957. Representations of $(x_1^2 + \cdots + x_n^2)^k$ as a real sum of $2k$-th powers and can be identified with quadrature formulas of strength $2k+1$ on the $S^{n-1} \subseteq \mathbb{R}^n$. In this sense, (3.99) can be traced back to work of Mehler from 1864.

We finish by revisiting quadrature formulas on $[-1,1]$. For strength 3, (3.82) becomes:

$$2x^3 + 3 \cdot \tfrac{2}{3} xy^2 = \sum_{k=1}^{r} \lambda_k (x + t_k y)^3.$$

Sylvester's Theorem with $r = 2$ yields

$$\begin{pmatrix} 2 & 0 & 2/3 \\ 0 & 2/3 & 0 \end{pmatrix} \cdot \begin{pmatrix} c_0 \\ c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff (c_0, c_1, c_2) = r(1, 0, -3).$$

Thus, $h(x,y) = x^2 - 3y^2$ and there exist λ_k so that $2x^3 + 3 \cdot \frac{2}{3} xy^2 = \lambda_1(\sqrt{3}x + y)^3 + \lambda_2(\sqrt{3}x - y)^3$. Cleaning up, we find the Gaussian quadrature formula of strength 3:

$$2x^3 + 2xy^2 = (x + \gamma y)^3 + (x - \gamma y)^3 \iff \int_{-1}^{1} f(t)\, dt = f(\gamma) + f(-\gamma); \quad \gamma = \sqrt{\tfrac{1}{3}}.$$

For strength 4, the matrix

$$\begin{pmatrix} 2 & 0 & 2/3 \\ 0 & 2/3 & 0 \\ 2/3 & 0 & 2/5 \end{pmatrix}$$

is non-singular, so there are no 2-point quadrature formulas of strength 3. But

$$\begin{pmatrix} 2 & 0 & 2/3 & 0 \\ 0 & 2/3 & 0 & 2/5 \end{pmatrix} \cdot \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

leads to $h(x,y) = rx(3x^2 - 2y^2) + sy(5x^2 - 3y^2)$. Thus, there are actually infinitely many 3-point quadrature formulas of strength 4, and we can choose (r,s) to ensure that any particular point is included. In this way, we see that there is no guarantee that the points $\{t_k\}$ need to be in the interval of integration, although practical people also want to choose points to minimize error. There will be an exercise or two on this, as well as one showing that there is exactly one choice of (r,s) for a strength 4 quadrature formula which actually has strength 5.

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 $n \ge 1$, define $p_n(x_1,\dots,x_n)$ and $q_n(x_1,\dots,x_n)$ by:

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

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

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

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

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

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.

$$\begin{gathered} p_{-1} = 0, \qquad p_0 = 1, \qquad p_1(x_1) = x_1, \qquad p_2(x_1,x_2) = x_1x_2 + 1, \\ p_3(x_1,x_2,x_3) = x_1x_2x_3 + x_1 + x_3, \\ p_4(x_1,x_2,x_3,x_4) = x_1x_2x_3x_4 + x_1x_2 + x_1x_4 + x_3x_4 + 1. \end{gathered} \tag{8}$$

It is evident from the definition that $p_n(x_1,\dots,x_n)$ 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

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

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

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

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

Proof. Let

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

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

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

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

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

as desired. □

The next observation is critical to understanding the Stern sequence.

Theorem 2. *For all n,*

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

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,*

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

We now present some "expansion" formulas for continuants.

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

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

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 $p_k = p_k(x_1, \ldots, x_k)$ 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*

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

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

$$\begin{aligned} a_1 &= y_1p_{m+1}(x_1,\dots,x_m,0) + p_m = y_1p_{m-1} + p_m \\ &= y_1p_{m-1} + (x_mp_{m-1}+p_{m-2}) = (y_1+x_m)p_{m-1} + p_{m-2} = b_1, \end{aligned} \tag{22}$$

and

$$a_2 = y_2a_1 + p_{m-1} = y_2b_1 + p_{m-1} = b_2. \tag{23}$$

Since $a_n = y_na_{n-1} + a_{n-2}$ and $b_n = y_nb_{n-1} + b_{n-2}$ 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

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

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),

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

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

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

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

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

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 $r \geq r_0 = \lceil \log_2 k \rceil$; 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,

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

so

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

It follows by Theorem 7 that

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

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, $[n_{2t}]_2 = [1010\cdots101]_2$, 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, $[n_{2t+1}]_2 = [1010\cdots1011]_2$, 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

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

$$\lim_{r\to\infty} \frac{s[m_{2t+1}]}{s[n_{4t+1}]} = \lim_{r\to\infty} \frac{F_{2t+2}^2}{F_{4t+3}} = \frac{(5+\sqrt{5})}{10}.$$

6. The rule of four revisited, revisited

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

$$\frac{s(n)}{s(n+1)} = a_{2v+1} + \cfrac{1}{a_{2v} + \cfrac{1}{\cdots + \cfrac{1}{a_1}}} = \frac{p_{2v+1}(a_1,\ldots,a_{2v+1})}{p_{2v}(a_1,\ldots,a_{2v})}\ , \qquad (95)$$

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

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

$$\frac{m}{m-a} = 1 + \frac{a}{m-a} = 1 + \cfrac{1}{\cfrac{m-a}{a}} = 1 + \cfrac{1}{\cfrac{m}{a} - 1}, \qquad (96)$$

we have, formally,

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

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

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

$\Delta : \mathcal{P}_d \to \mathcal{P}_{d-1}$. 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 $k \geq 1$,

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

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

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

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 $Mf(x) = g(x) = a_1 x + a_0$. 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 $g(x) = Mf(x) = a_2 x^2 + a_1 x + a_0$, 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 $k \geq 2$, 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 $ord_p(2) \leq d$, 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 $\nu_p(M) \leq \nu_p(M_d)$ 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

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

□

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 $r = 0, 1, \ldots, d-1$ 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.)

STERN NOTES, CHAPTER 5 (VERSION 1)

BRUCE REZNICK, UIUC

1. The Brocot array

The Brocot array, which goes back to 1861, is historically at least as prominent as the Stern sequence. It can be construed as a double diatomic array. The basic idea was discussed at the beginning of these notes: if $\frac{a}{b}$ and $\frac{c}{d}$ are consecutive in the r-th row, then the *mediant*, $\frac{a+c}{b+d}$, is inserted beween them in the $r+1$-st row. The full array starts with $(\frac{0}{1}, \frac{1}{0})$:

$$\begin{array}{ccccccccc} & & & & \frac{0}{1} \ \frac{1}{0} & & & & \\ & & & \frac{0}{1} & \frac{1}{1} & \frac{1}{0} & & & \\ & & \frac{0}{1} & \frac{1}{2} & \frac{1}{1} & \frac{2}{1} & \frac{1}{0} & & \\ \frac{0}{1} & \frac{1}{3} & \frac{1}{2} & \frac{2}{3} & \frac{1}{1} & \frac{3}{2} & \frac{2}{1} & \frac{3}{1} & \frac{1}{0} \\ & & & & \cdots & & & & \end{array} \tag{1}$$

It is clear, using the earlier notation, that the r-th row (starting with $r = 0$) has 2^r elements, and that the k-th element (starting with $k = 0$) is

$$\frac{Z(r,k;0,1)}{Z(r,k;0,1)} = \frac{s(k)}{s(2^r - k)}. \tag{2}$$

Considering the symmetry of the array and the dubiousness of "$\frac{1}{0}$", it is customary to write only the first half of this picture, thus starting in effect with $(\frac{0}{1}, \frac{1}{1})$, and call it the *Brocot array*:

$$\begin{array}{ccccccccc} & & & & \frac{0}{1} \ \frac{1}{1} & & & & \\ & & & \frac{0}{1} & \frac{1}{2} & \frac{1}{1} & & & \\ & & \frac{0}{1} & \frac{1}{3} & \frac{1}{2} & \frac{2}{3} & \frac{1}{1} & & \\ \frac{0}{1} & \frac{1}{4} & \frac{1}{3} & \frac{2}{5} & \frac{1}{2} & \frac{3}{5} & \frac{2}{3} & \frac{3}{4} & \frac{1}{1} \\ & & & & \cdots & & & & \end{array} \tag{3}$$

So, in the same labeling, the k-th entry in the r-th row is

$$\frac{Z(r,k;0,1)}{Z(r,k;1,1)} = \frac{s(k)}{s(2^{r+1}-k)} = \frac{s(k)}{s(2^r+k)}. \tag{4}$$

(This last expression is possible for $0 < k < 2^r$, because then $(2^{r+1}-k)' = 2^r + k$; it is true for $k = 0, 2^r$ because $s(2^r) = s(2^{r+1})$.)

Lemma 1. *The entries of the r-th row of the Brocot array are increasing.*

Proof. We prove something more, that

$$\frac{s(k+1)}{s(2^r+k+1)} - \frac{s(k)}{s(2^r+k)} = \frac{1}{s(2^r+k)s(2^r+k+1)}. \tag{5}$$

Indeed, this is true in the first row; $\frac{1}{1} - \frac{0}{1} = \frac{1}{1}$. Assuming it is true in the r-th row, with entries $\frac{a}{b}, \frac{c}{d}$, we note that

$$\begin{aligned} \frac{a+c}{b+d} - \frac{a}{b} &= \frac{b(a+c)-a(b+d)}{b(b+d)} = \frac{bc-ad}{b(b+d)} = \frac{1}{b(b+d)}, \\ \frac{c}{d} - \frac{a+c}{b+d} &= \frac{c(b+d)-d(a+c)}{d(b+d)} = \frac{bc-ad}{d(b+d)} = \frac{1}{d(b+d)}. \end{aligned} \tag{6}$$

Thus the result holds by induction. □

(We also note that $s(k+1)s(2^r+k) - s(k)s(2^r+k+1) = 1$, which suggests that the next problem set will contain an examination of $s(m)s(n+1) - s(m+1)s(m)$.)

This property can be used to show that every rational in $[0,1]$ appears in the Brocot array, but we can also use Theorem 11 from last week: If k is odd, $2^{r_0} < k < 2^{r_0+1}$ and $r \ge r_0 + 1$, then

$$\frac{s(2^r+k)}{s(k)} = \frac{s((\overleftarrow{2^r+k})')}{s((\overleftarrow{2^r+k})'+1)}. \tag{7}$$

Let's work backwards. Suppose $0 < \frac{p}{q} < 1$ and suppose that the quotient in (7) is $\frac{q}{p}$. We already know that there exists a unique odd n so that $s(n) = q$ and $s(n+1) = p$. Suppose $2^r < n < 2^{r+1}$, and write $n = 2^r + \ell$, so that $n' = 2^{r+1} - \ell$. We have

$$n = (\overleftarrow{2^r+k})' \iff 2^{r+1} - \ell = \overleftarrow{2^r+k} \iff \overleftarrow{2^{r+1}-\ell} = 2^r + k. \tag{8}$$

Thus, if we keep the same r, and define k by

$$k = (\overleftarrow{2^{r+1}-\ell}) - 2^r, \tag{9}$$

then we have

$$\frac{p}{q} = \frac{s(k)}{s(2^r+k)} = \frac{s(n+1)}{s(n)}. \tag{10}$$

It is possible to derive a more explicit formula for k from this information, but it is actually faster and more instructive to do it from scratch.

The *Minkowski ?-function* is defined on $[0, 1]$, but for now we define it on the entries of the Brocot array by:

$$? \left(\frac{s(k)}{s(2^r + k)} \right) := \frac{k}{2^r}. \tag{11}$$

We first note that $?(x)$ is well-defined, because $s(2k) = s(k)$, $s(2^{r+1}+2k) = s(2^r+k)$ and $\frac{2k}{2^{r+1}} = \frac{k}{2^r}$. It is also strictly increasing on each row of the Brocot array, and so on the rationals. It will be helpful to also consider the inverse function $?^{-1}(x)$

$$?^{-1} \left(\frac{k}{2^r} \right) = \frac{s(k)}{s(2^r + k)}. \tag{12}$$

In order to develop the closed form, we consider the continued fractions of $?^{-1}(\frac{k}{2^r})$ for small values of r.

$$\frac{1}{2} \to \frac{1}{2^1};$$

$$\frac{1}{3} \to \frac{1}{2^2}, \quad \cfrac{1}{1 + \cfrac{1}{2}} \to \frac{3}{2^2}; \tag{13}$$

$$\frac{1}{4} \to \frac{1}{2^3}, \quad \cfrac{1}{2 + \cfrac{1}{2}} \to \frac{3}{2^3}, \quad \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{2}}} \to \frac{5}{2^3}, \quad \cfrac{1}{1 + \cfrac{1}{3}} \to \frac{7}{2^3}.$$

There are several patterns implicit in this table, which we now prove. The first lemma may well appear earlier in the notes.

Lemma 2. *If $0 < k < 2^r$ is odd, $s(k) = p$ and $s(2^r + k) = q$, then $s(2^{r+1} + k) = p + q$.*

Proof. We have seen *ad nauseum* that $s(2^r n + k)$ is linear in r for $r \geq \lceil \log_2 k \rceil$, with coefficient $s(n)s(k)$. Just take $n = 1$. □

Lemma 3.

$$? \left(\frac{p}{q} \right) = \frac{k}{2^r} \implies ? \left(\frac{p}{p+q} \right) = \frac{k}{2^{r+1}} \quad \textit{and} \quad ? \left(\frac{q}{p+q} \right) = \frac{2^{r+1} - k}{2^{r+1}}. \tag{14}$$

Proof. We assume that $s(k) = p$ and $s(2^r + k) = q$. As noted above, $s(2^{r+1} + k) = p + q$. Since $2^r < 2^r + k < 2^{r+1}$, $(2^r + k)' = 2^{r+1} - k$ and $s(2^{r+1} - k) = q$; since $2^{r+1} < 2^{r+1} + k < 2^{r+2}$, $(2^{r+1} + k)' = 2^{r+2} - k$ and $s(2^{r+2} - k) = p + q$. Thus, we have

$$\frac{s(k)}{s(2^{r+1} + k)} = \frac{p}{p+q}, \qquad \frac{s(2^{r+1} - k)}{s(2^{r+1} + (2^{r+1} - k))} = \frac{q}{p+q}. \tag{15}$$

□

On taking $x = \frac{p}{q}$ and noting that $?(\frac{1}{2}) = \frac{1}{2}$, it follows by induction from this lemma that for $x \in (0,1) \cap \mathbb{Q}$,

$$?(x)+?(1-x) = 1, \qquad ?\left(\frac{x}{x+1}\right) = \frac{?(x)}{2}, \qquad ?\left(\frac{1}{x+1}\right) = 1 - \frac{?(x)}{2}. \tag{16}$$

It is now convenient to introduce some notation. Let $a = (a_1, \dots, a_w) \in \mathbb{N}^w$ for some $w \ge 1$ and let

$$[a] = [0, a_1, \dots, a_w] := 0 + \cfrac{1}{a_1 + \cfrac{1}{\cdots + \cfrac{1}{a_w}}} \ . \tag{17}$$

We also define

$$\|a\| = \sum_{j=1}^{w} a_j, \tag{18}$$

and observe that if $\|b\| = r + 1$, then either $(b_1, \dots, b_w) = (1 + a_1, \dots, a_w)$ or $(b_1, \dots, b_w) = (1, a_2, \dots, a_w)$, where $\|a\| = r$, depending whether or not $a_1 > 1$.

Lemma 4.

$$[0, a_1, \dots, a_w] = \frac{p}{q} \implies [0, 1{+}a_1, \dots, a_w] = \frac{p}{p+q}, \quad [0, 1, a_1, \dots, a_w] = \frac{q}{p+q} \ . \tag{19}$$

Proof. Since

$$\frac{q}{p} = a_1 + \cfrac{1}{\cdots + \cfrac{1}{a_w}} \ , \tag{20}$$

we have immediately that

$$[0, 1 + a_1, \dots, a_w] = \cfrac{1}{1 + \cfrac{q}{p}}, \quad \text{and} \quad [0, 1, a_1, \dots, a_w] = \cfrac{1}{1 + \cfrac{p}{q}} \ . \tag{21}$$

□

Theorem 5.

$$?([0, a_1, \dots, a_w]) = \frac{1}{2^{a_1 - 1}} - \frac{1}{2^{a_1 + a_2 - 1}} + - \cdots + \frac{(-1)^{w+1}}{2^{a_1 + a_2 + \cdots + a_w - 1}} \ . \tag{22}$$

Proof. We first observe that there is no ambiguity in the definition arising from

$$[0, a_1, \dots, a_{w-1}, a_w, 1] = [0, a_1, \dots, a_{w-1}, a_w + 1], \tag{23}$$

inasmuch as

$$\frac{(-1)^{w+1}}{2^{a_1 + a_2 + \cdots + a_w - 1}} + \frac{(-1)^{w+2}}{2^{a_1 + a_2 + \cdots + a_w + 1 - 1}} = \frac{(-1)^{w+1}}{2^{a_1 + a_2 + \cdots + a_w + 1 - 1}}. \tag{24}$$

To make the right-hand side look more familiar, observe that

$$\frac{1}{2^{a_1-1}} - \frac{1}{2^{a_1+a_2-1}} + - \cdots + \frac{(-1)^{w+1}}{2^{a_1+a_2+\cdots+a_w-1}} = \frac{k}{2^{\|a\|-1}} \tag{25}$$

for some odd k. We induct on $\|a\|$. The cases in which $\|a\| \le 3$ are [1], [2] = [1, 1], [3] = [2, 1] and [1, 2] = [1, 1, 1]. Referring to the previous table, we do indeed have

$$\begin{gathered} ?([1]) = \frac{1}{2^{1-1}}; \qquad ?([2]) = ?\left(\frac{1}{2}\right) = \frac{1}{2^{2-1}}; \\ ?([3]) = ?\left(\frac{1}{3}\right) = \frac{1}{2^{3-1}}, \qquad ?([1,2]) = ?\left(\frac{2}{3}\right) = \frac{1}{2^{1-1}} - \frac{1}{2^{1+2-1}} = \frac{3}{4}. \end{gathered} \tag{26}$$

Suppose $\|b\| = r + 1$. There are two cases. If $b_1 > 1$, then $b = (1 + a_1, \ldots, a_w)$ with $\|a\| = r$. Write $[0, a_1, \ldots, a_w] = \frac{p}{q}$ and $?(\frac{p}{q}) = \frac{k}{2^r}$. By Lemmas 3 and 4 and the inductive hypothesis,

$$\begin{gathered} ?([0, 1 + a_1, \ldots, a_w]) = ?\left(\frac{p}{p+q}\right) = \frac{k}{2^{r+1}} = \\ \frac{1}{2} \cdot \left(\frac{1}{2^{a_1-1}} - \frac{1}{2^{a_1+a_2-1}} + - \cdots + \frac{(-1)^{w+1}}{2^{a_1+a_2+\cdots+a_w-1}}\right) \\ = \frac{1}{2^{1+a_1-1}} - \frac{1}{2^{1+a_1+a_2-1}} + - \cdots + \frac{(-1)^{w+1}}{2^{1+a_1+a_2+\cdots+a_w-1}}. \end{gathered} \tag{27}$$

If $b_1 = 1$, then $b = (1, a_1, \ldots, a_w)$ with $\|a\| = r$. Using the same notation and arguments as above,

$$\begin{gathered} ?([0, 1, a_1, \ldots, a_w]) = ?\left(\frac{q}{p+q}\right) = \frac{2^{r+1} - k}{2^{r+1}} = 1 - \frac{1}{2} \cdot \frac{k}{2^r} = \\ \frac{1}{2^{1-1}} - \frac{1}{2} \cdot \left(\frac{1}{2^{a_1-1}} - \frac{1}{2^{a_1+a_2-1}} + - \cdots + \frac{(-1)^{w+1}}{2^{a_1+a_2+\cdots+a_w-1}}\right) \\ = \frac{1}{2^{1-1}} - \frac{1}{2^{1+a_1-1}} + - \cdots + \frac{(-1)^{w+2}}{2^{1+a_1+a_2+\cdots+a_w-1}}. \end{gathered} \tag{28}$$

This completes the proof. □

It is worth noting that this proof essentially establishes the formula for $?^{-1}$, and that it does *not* try to compute the continued fraction expansion of the mediant from its component parts, which would have been the first, natural step. We could have derived (32) from the previous formula for the Stern sequence, although the appearance of n' means that there would be cases depending on whether the initial and terminal denominators are 1. (This case breakdown does not occur in the above derivation, but see below.) One can also use this approach to derive a closed formula for the Stern sequence, but it seems longer than what we did.

We can always specify that w is even. In this case, the binary expression of $\frac{k}{2^r}$ might depends on whether $a_1 > 1$ or not. If $a_1 > 1$, then this expression contains

$a_1 - 1$ 0's, followed by a_2 1's, a_3 0's, ending up with a_w 1's. If $a_1 = 1$, then we have a_2 1's, a_3 0's, ending up with a_w 1's.

If $a^{(r,1)}, \dots a^{(r,2^{r-1})}$ are the continued fraction denominators which appear corresponding to $\frac{s(k)}{s(2^r+k)}$ for odd k on the r-th row, then $\|a^{(r,j)}\| = r+1$, and the corresponding denominators on the $(r+1)$-st row are, first, the ones from the r-th row, in order, with a_1 incremented by one, and then, the ones from the r-th row, in reverse order, with 1 appended to the left. Thus, for $r \le 3$, we have:

$$\begin{gathered} [1] \\ [2] \quad [1,1] \\ [3] \quad [2,1] \quad [1,1,1] \quad [1,2] \\ [4] \quad [3,1] \quad [2,1,1] \quad [2,2], \quad [1,1,2], \quad [1,1,1,1], \quad [1,2,1], \quad [1,3] \\ \dots \end{gathered} \tag{29}$$

2. Some finite computations

It is worth noting that the computation of continued fractions with repeating denominators is closely linked to linear recurrences. For example, suppose we consider the continued fraction with n identical denominators equal to a.

$$a + \cfrac{1}{a + \cfrac{1}{\dots + \cfrac{1}{a}}} = \frac{p_n(a,\dots,a)}{p_{n-1}(a,\dots,a)} \, . \tag{30}$$

Let $x_n(a) = p_n(a,\dots,a)$. Then:

$$x_0(a) = 1, \quad x_1(a) = a, \quad x_n(a) = a x_{n-1}(a) + x_{n-2}(a). \tag{31}$$

The characteristic equation and its roots are

$$\lambda^2 - a\lambda - 1 = 0 \implies \lambda_+ = \frac{a + \sqrt{a^2+4}}{2}, \qquad \lambda_- = \frac{a - \sqrt{a^2+4}}{2} \, . \tag{32}$$

We know that $x_n(a) = c_1\lambda_+^n + c_2\lambda_-^n$, and since $x_{-1}(a) = 0$, at least consistently with the initial conditions, it is not hard to show that

$$x_n(a) = \frac{\lambda_+^{n+1} - \lambda_-^{n+1}}{\lambda_+ - \lambda_-} = \frac{\lambda_+^{n+1} - \lambda_-^{n+1}}{\sqrt{a^2+4}}. \tag{33}$$

This reminds one that $x_n(1) = F_{n+1}$. If $a \in \mathbb{N}$, then $\lambda_+ > 1$ and $\lambda_+\lambda_- = -1$ imply that $\lambda_-^m \to 0$ as $m \to \infty$ so that as $n \to \infty$,

$$x_n(a) = a + \cfrac{1}{a + \cfrac{1}{\dots + \cfrac{1}{a}}} = \frac{x_n(a)}{x_{n-1}(a)} \to \lambda_+ = \frac{a + \sqrt{a^2+4}}{2} \, . \tag{34}$$

Inasmuch as

$$x_n(a) = a + \frac{1}{x_{n-1}(a)}, \tag{35}$$

if we knew *a priori* that $\lim x_n(a) = \theta$, for some θ, we would have $\theta = a + \frac{1}{\theta}$, which implies that $\theta^2 - a\theta - 1 = 0$. This is true, as part of a more general calculation in the next section.

In terms of the Minkowski ?-function, we have

$$\begin{aligned} ?\left(\frac{x_{n-1}(a)}{x_n(a)}\right) &= ?([0, a, \dots, a]) = \frac{1}{2^{a-1}} - \frac{1}{2^{2a-1}} + - \cdots + \frac{(-1)^{n+1}}{2^{na-1}} \\ &= \frac{1}{2^a + 1} \cdot \frac{2^{na} - (-1)^n}{2^{na-1}} = \frac{2}{2^a + 1} + \frac{(-1)^{n+1}}{(2^a + 1)2^{na-1}} \ . \end{aligned} \tag{36}$$

For example, taking $a = 1$, we have

$$?\left(\frac{F_n}{F_{n+1}}\right) = \frac{2}{3} + \frac{(-1)^{n+1}}{3 \cdot 2^{n-1}}. \tag{37}$$

These will be useful formulas when we look at $?'(x)$. It is also not too difficult to compute explicitly

$$a + \cfrac{1}{b + \cfrac{1}{a + \cfrac{1}{b + \dots}}} = \frac{p_n(a, b, a, b, \dots)}{p_{n-1}(b, a, b, a, \dots)} \ . \tag{38}$$

We consider the numerator first, defining $y_n(a, b)$ by:

$$\begin{gathered} y_0(a, b) = 1, \quad y_1(a, b) = a, \\ y_{2n}(a, b) = by_{2n-1}(a, b) + y_{2n-2}(a, b), \quad y_{2n+1}(a, b) = ay_{2n}(a, b) + y_{2n-1}(a, b). \end{gathered} \tag{39}$$

This doesn't look very promising, but if we suppress the argument for brevity, we can note that

$$\begin{aligned} y_{2n} &= y_{2n}, \\ y_{2n+1} &= ay_{2n} + y_{2n-1}, \\ y_{2n+2} &= by_{2n+1} + y_{2n} = (ab + 1)y_{2n} + by_{2n-1}, \\ y_{2n+3} &= ay_{2n+2} + y_{2n+1} = (a^2b + 2a)y_{2n} + (ab + 1)y_{2n-1}, \\ y_{2n+4} &= by_{2n+3} + y_{2n+2} = (a^2b^2 + 3ab + 1)y_{2n} + (ab^2 + 2b)y_{2n-1}. \end{aligned} \tag{40}$$

This looks even less promising; miraculously though,

$$y_{2n+4} = (ab + 2)y_{2n+2} - y_{2n}, \tag{41}$$

and since this equation is symmetric in $\{a, b\}$, it would also have applied if we had started at y_{2n+1}, with a and b reversed. Therefore, we can say that for $r \geq 4$,

$$y_r = (ab + 2)y_{r-2} - y_{r-4} \tag{42}$$

and in this way obtain explicit formulas. It is not accidental that this might be helpful in some problems on the second homework. Finally, if the limit in (38) is θ, then it is reasonable to expect that

$$\theta = a + \cfrac{1}{b + \frac{1}{\theta}} \implies b\theta^2 - ab\theta - a = 0, \tag{43}$$

and θ would be the positive root of this quadratic. (If $a = b$, this reduces to the previous equation, as it should.)

3. Biting the bullet

We would like to extend the definition of $?(x)$ to the entire interval $[0, 1]$, which means we will have to consider infinite continued fractions. This means that we really should more standard terminology. We shall say that

$$[x_0, x_1, \dots x_n] := x_0 + \cfrac{1}{x_1 + \cfrac{1}{x_2 + \cfrac{1}{\dots + \cfrac{1}{x_n}}}}\,, \tag{44}$$

under the convention that $x_i \in \mathbb{Z}$, with $x_i \geq 1$ for $i \geq 1$. We already know that

$$[x_0, x_1, \dots x_n] = \frac{p_{n+1}(x_0, \dots, x_n)}{p_n(x_1, \dots, x_n)}\,. \tag{45}$$

We should have said explicitly, but haven't yet, that we can also write

$$x_0 + \cfrac{1}{x_1 + \cfrac{1}{x_2 + \cfrac{1}{\dots + \cfrac{1}{x_n}}}} = x_0 + \cfrac{1}{x_1 + \cfrac{1}{\dots + \cfrac{1}{x_k + \cfrac{1}{[x_{k+1}, \dots, x_n]}}}} \tag{46}$$

for any k, $0 < k < n$, with the acknowledgement that the last "denominator" is not an integer.

Lemma 6. *If $x_i \geq 1$, then $p_n(x_1, \dots, x_n) \geq F_{n+1}$.*

Proof. Since p_n is a polynomial with non-negative coefficients, we can only decrease it by replacing each x_i with 1: $p_n(x_1, \dots, x_n) \geq p_n(1^n) = F_{n+1}$. □

Lemma 7. *Let integers x_i be given as above and let $\xi_n = [x_0, x_1, \dots x_n]$. Then (ξ_n) is a convergent sequence.*

Proof. For $n \geq 2$, by the very important Theorem 9 (Notes IV),

$$\begin{aligned}\xi_n - \xi_{n-1} &= \frac{p_{n+1}(x_0, \ldots, x_n)}{p_n(x_1, \ldots, x_n)} - \frac{p_n(x_0, \ldots, x_{n-1})}{p_{n-1}(x_1, \ldots, x_{n-1})} \\ &= \frac{(-1)^{n+1}}{p_{n-1}(x_1, \ldots, x_{n-1}) p_n(x_1, \ldots, x_n)}.\end{aligned} \tag{47}$$

In view of the last lemma, this implies that for $n \geq 2$,

$$|\xi_n - \xi_{n-1}| \leq \frac{1}{F_{n-1} F_n} \leq \frac{1}{2^{n-1}}, \tag{48}$$

the last inequality coming from an easy induction ($\phi^2 = 1 + \phi \approx 2.618 > 2$.) It follows that (ξ_n) is Cauchy, and hence is convergent. □

We have already discussed the continued fraction representations of rational numbers. We now talk about irrationals. Given real $t \notin \mathbb{Q}$, define

$$G(t) := \frac{1}{t - \lfloor t \rfloor} \ . \tag{49}$$

Since $t \notin \mathbb{Q}$, $t > \lfloor t \rfloor$, so that $G(t) \in (1, \infty)$ is well-defined, and since $t = \lfloor t \rfloor + \frac{1}{G(t)}$, we see that $G(t) \notin \mathbb{Q}$ as well. We also see that

$$t = \lfloor t \rfloor + \frac{1}{G(t)} = \lfloor t \rfloor + \cfrac{1}{\lfloor G(t) \rfloor + \cfrac{1}{G(G(t))}}, \qquad \text{etc.} \tag{50}$$

Let $G_n(t)$ denote the n-th iterate of G ($G_1 = G$, $G_n = G \circ G_{n-1}$) and let $x_n = x_n(t) := \lfloor G_n(t) \rfloor$. Then the preceding discussion shows the existence of a family of *exact* formulas:

$$t = x_0(t) + \cfrac{1}{x_1(t) + \cfrac{1}{\cdots + \cfrac{\cdots}{x_{n-1}(t) + \cfrac{1}{G_n(t)}}}}. \tag{51}$$

Again, to be precise, keep in mind that $G_n(t)$ is not an integer, but there should be no confusion. We need some short lemmas.

Lemma 8. *The function $(-1)^n [x_0, x_1, \ldots, x_n]$ is increasing in real x_n.*

Proof. Clearly, x_0 is increasing in x_0 and $x_0 + \frac{1}{x_1}$ is decreasing in x_1. Since

$$[x_0, \ldots, x_n] = x_0 + \frac{1}{[x_1, \ldots, x_n]}, \tag{52}$$

the result is immediate by induction. □

Lemma 9. *Provided $x_j \ge 1$ for all j, we have*

$$[x_0, \dots, x_{2k}] < [x_0, \dots, x_{2k+2}] < [x_0, \dots, x_{2k+3}] < [x_0, \dots, x_{2k+1}]. \tag{53}$$

Proof. We have

$$\begin{aligned} [x_0, \dots, x_{2k}] &< [x_0, \dots, x_{2k} + [x_{2k+1}, \dots, x_{2k+r}]^{-1}] = [x_0, \dots, x_{2k+r}], \\ [x_0, \dots, x_{2k+1}] &> [x_0, \dots, x_{2k+1} + [x_{2k+2}, \dots, x_{2k+1+r}]^{-1}] = [x_0, \dots, x_{2k+1+r}]. \end{aligned} \tag{54}$$

□

We now prove the existence of infinite continued fractions.

Theorem 10. *If $t \notin \mathbb{Q}$, then*

$$t = \lim_{n\to\infty} [x_0(t), \dots, x_n(t)]. \tag{55}$$

Proof. Let $\xi_n = \xi_n(t) := [x_0(t), \dots, x_n(t)]$, then

$$t - \xi_n = x_0 + \cfrac{1}{x_1 + \cfrac{1}{\cdots + \cfrac{\cdots}{x_{n-1} + \cfrac{1}{g_n(t)}}}} \quad - \quad x_0 + \cfrac{1}{x_1 + \cfrac{1}{\cdots + \cfrac{\cdots}{x_{n-1} + \cfrac{1}{x_n}}}}. \tag{56}$$

Since $x_n < g_n(t) < 1 + x_n$, it follows from monotonicity that

$$\begin{aligned} |t - \xi_n| &< \big| [x_0, \dots, x_n] - [x_0, \dots, 1 + x_n] \big| \\ &= \left| \frac{p_{n+1}(x_0, \dots, x_n)}{p_n(x_1, \dots, x_n)} - \frac{p_{n+1}(x_0, \dots, 1 + x_n)}{p_n(x_1, \dots, 1 + x_n)} \right|. \end{aligned} \tag{57}$$

Let us write $a = p_n(x_0, \dots, x_{n-1})$, $b = p_{n-1}(x_0, \dots, x_{n-2})$, $c = p_{n-1}(x_1, \cdots, x_{n-1})$ and $d = p_{n-2}(x_1, \cdots, x_{n-2})$. As previously noted, $|bc - ad| = 1$. so this inequality becomes

$$\begin{aligned} |t - \xi_n| &< \left| \frac{x_n a + b}{x_n c + d} - \frac{(1 + x_n)a + b}{(1 + x_n)c + d} \right| \\ &= \frac{|ad - bc|}{(x_n c + d)((1 + x_n)c + d)} < \frac{1}{F_n^2}. \end{aligned} \tag{58}$$

It follows that $\xi_n \to t$ and, indeed, that

$$\xi_0(t) < \xi_2(t) < \xi_4(t) \cdots < t < \cdots < \xi_3(t) < \xi_1(t). \tag{59}$$

□

It is customary to say that the $\xi_n(t)$'s are the *convergents* of t.

Theorem 11. *The continued fraction representation of an irrational number is unique and infinite; conversely, every infinite continued fraction represents an irrational.*

Proof. Suppose $t = [x_0, x_1, \dots] = [y_0, y_1, \dots]$ with $x_j, y_j \in \mathbb{Z}$, $x_i, y_i \ge 1$ for $i \ge 1$. We prove by induction that $x_j = y_j$. Indeed, since

$$x_0 < [x_0, x_1, \dots] < 1 + x_0, \qquad y_0 < [y_0, y_1, \dots] < 1 + y_0, \tag{60}$$

we have $x_0 = y_0 = m_0$. But now,

$$\frac{1}{t - m_0} = [x_1, x_2, \dots] = [y_1, y_2, \dots], \tag{61}$$

and we may repeat the argument. Suppose on the other hand that $u = [x_0, x_1, \dots] = [y_0, y_1, \dots, y_n]$. Repeating the previous argument n times leads us to $[x_n, x_{n+1}, \dots] = [y_n]$, and $x_n < [x_n, x_{n+1}, \dots] < 1 + x_n$ is not an integer, a contradiction. □

We now return to some topics touched on in the earlier notes. An infinite continued fraction is *purely periodic* of period $d \ge 1$ if all its denominators are repeating blocks of length d:

$$[\overline{x_0, \dots, x_{d-1}}] := [x_0, \dots, x_{d-1}, x_0, \dots, x_{d-1}, \dots]. \tag{62}$$

Let $\theta_k = \theta_k(x_0, \dots, x_{d-1})$ denote the finite continued fraction with kd denominators, comprising k complete cycles of the pattern $(x_0, \dots, x_{d-1})$, or, more formally, $\theta_1 = [x_0, x_1, \dots .x_{d-1}]$ and

$$\theta_k = x_0 + \cfrac{1}{x_1 + \cfrac{1}{\cdots + \cfrac{\cdots}{x_{d-1} + \cfrac{1}{\theta_{k-1}}}}}. \tag{63}$$

If $d = 1$ and $x_0 = a$, we have the familiar

$$\theta_k = a + \frac{1}{\theta_{k-1}} = \frac{x_k(a)}{x_{k-1}(a)}, \tag{64}$$

and if $d \ge 2$,

$$\theta_k = \frac{p_{d+1}(x_0, \dots, x_{d-1}, \theta_{k-1})}{p_d(x_1, \dots, x_{d-1}, \theta_{k-1})} = \frac{\theta_{k-1} p_d(x_0, \dots, x_{d-1}) + p_{d-1}(x_0, \dots, x_{d-2})}{\theta_{k-1} p_{d-1}(x_1, \dots, x_{d-1}) + p_{d-2}(x_1, \dots, x_{d-2})}. \tag{65}$$

If $d = 1$, this is consistent with $p_0 = 1, p_{-1} = 0$, so we won't distinguish that case.

Observe that (θ_k) is the subsequence of the kd-th entries in a convergent sequence and so is also convergent, to, say, $\theta = \theta_{(x_0, \dots, x_{d-1})}$. By continuity of polynomials,

$$\theta = \frac{p_d(x_0, \dots, x_{d-1}) \cdot \theta + p_{d-1}(x_0, \dots, x_{d-2})}{p_{d-1}(x_1, \dots, x_{d-1}) \cdot \theta + p_{d-2}(x_1, \dots, x_{d-2})} := \frac{A_x \theta + B_x}{C_x \theta + D_x}, \tag{66}$$

and so θ is a root of the quadratic:

$$C_x T^2 + (D_x - A_x) T - B_x. \tag{67}$$

If $x_0 \geq 1$, then $B_x > 0$, so that there are two roots, one positive and one negative. Clearly, $\theta > 0$. Thus, every purely periodic continued fraction is a quadratic irrational; $\theta \in \mathbb{Q}(\sqrt{(D-A)^2+4BC})$. The uniqueness of continued fractions rules out the possibility that $\theta \in \mathbb{Q}$.

Example. Suppose $d = 1$, so that

(68) $$\theta_k == \frac{x_k(a)}{x_{k-1}(a)} .$$

Plugging into the previous expression with $x = (a)$, we have $A_x = a, B_x = C_x = 1, D_x = 0$, and the quadratic above is $T^2 - aT - 1$, as we have already seen. If $d = 2$, then $A = ab+1, B = b, C = a, D = 1$, and we get the quadratic seen in (43).

An infinite continued fraction is *periodic* if it is periodic after an initial string. We write

(69) $$[y_0, \dots, y_e, \overline{x_0, \dots, x_{d-1}}] := [y_0, \dots, y_e, x_0, \dots, x_{d-1}, x_0, \dots, x_{d-1}, \dots].$$

If $\theta = [\overline{x_0, \dots, x_{d-1}}]$ as before, and ψ_k denotes the continued fraction with the initial string, followed by k repeated blocks, then as before

(70) $$\psi_k = \frac{p_{e+1}(y_0, \dots, y_{e-1}, \theta_k)}{p_e(y_1, \dots, y_{e-1}, \theta_k)} = \frac{\theta_k p_e(y_0, \dots, y_{d-1}) + p_{e-1}(y_0, \dots, y_{e-2})}{\theta_k p_{e-1}(y_1, \dots, y_{e-1}) + p_{e-2}(e_1, \dots, e_{d-2})}.$$

If $e = 1$, then we simply define

(71) $$[y_0, \overline{x_0, \dots, x_{d-1}}] = y_0 + \frac{1}{\theta}$$

In any event, the convergence of continued fractions implies that (ψ_k) is convergent, to ψ, say, and the convergence of (θ_k) implies

(72) $$\psi = \frac{\theta p_e(y_0, \dots, y_{d-1}) + p_{e-1}(y_0, \dots, y_{e-2})}{\theta p_{e-1}(y_1, \dots, y_{e-1}) + p_{e-2}(e_1, \dots, e_{d-2})}.$$

Since each of the continuants is an integer, we see that so $\psi \in \mathbb{Q}(\sqrt{(D-A)^2+4BC})$ as well. Remarkably enough, the converse is true:

Theorem 12. *If $t = a + b\sqrt{d} \in \mathbb{Q}(\sqrt{d})$ is irrational, then the continued fraction expansion of t is periodic.*

This will be proved in the next section.

4. Lagrange's Theorem

Lagrange's Theorem is the result alluded to at the end of the last section: if u is a quadratic irrational, then the continued fraction expansion of u is periodic. The first thing we need to do is write quadratic irrationals in a peculiar form.

Lemma 13. *Suppose $u \in \mathbb{Q}(\sqrt{n}) \setminus \mathbb{Q}$, where $n \in \mathbb{N}$ is not a square. Then there exist (non-unique) integers m, d, q so that*

$$u = \frac{m + \sqrt{d}}{q}, \tag{73}$$

where $m, d, q \in \mathbb{Z}$, $q \neq 0$, $d > 0$ is not a square and so that $q \mid d - m^2$.

Proof. By hypothesis, we have $u = \alpha + \beta\sqrt{n}$, with $\alpha, \beta \in \mathbb{Q}$ and $\beta \neq 0$. After taking a common denominator, there exist integers c_j, with $c_2, c_3 \neq 0$ so that

$$u = \frac{c_1 + c_2\sqrt{n}}{c_3} = \frac{-c_1 + -c_2\sqrt{n}}{-c_3}\ . \tag{74}$$

We may thus assume without loss of generality that $c_2 > 0$ and so that

$$u = \frac{c_1 + \sqrt{nc_2^2}}{c_3} = \frac{c_1|c_3| + \sqrt{nc_2^2c_3^2}}{c_3|c_3|}\ . \tag{75}$$

If we let $m = c_1|c_3|$, $d = nc_2^2c_3^2$ and $q = c_3|c_3|$, then $m, d, q \in \mathbb{Z}$, $q \neq 0$, $d > 0$ is not a square and

$$d - m^2 = c_3^2(nc_2^2 - c_1)^2 = \pm q(nc_2^2 - c_1)^2. \tag{76}$$

We remark that this representation is not unique: if $r \in \mathbb{N}$, then another valid representation will hold under the substitution $(m, q, d) \to (rm, rq, r^2d)$. □

Theorem 14 (Lagrange's Theorem)**.** *A quadratic irrational u has a periodic continued fraction.*

Proof. By Lemma 1, we may write

$$u = \frac{m + \sqrt{d}}{q} = [a_0, a_1, \dots]. \tag{77}$$

Recall from the last section that

$$\xi_n(u) = \frac{p_{n+1}(a_0, \dots, a_n)}{p_n(a_1, \dots, a_n)} := \frac{p_{n+1}}{q_{n+1}} \to u. \tag{78}$$

(Note that where $p_{n+1}, q_{n+1} \in \mathbb{N}$.) Also, there is an *exact* expression

$$u = [a_0, \dots, a_{n-1}, G_n(u)] = \frac{p_{n+1}(a_0, \dots, a_{n-1}, G_n(u))}{p_n(a_1, \dots, a_{n-1}, G_n(u))} = \frac{G_n(u)p_n + p_{n-1}}{G_n(u)q_n + q_{n-1}}. \tag{79}$$

We claim that we can always write

$$G_n(u) = \frac{m_n + \sqrt{d}}{q_n}, \tag{80}$$

where $m_n, q_n \in \mathbb{Z}$, $q_n \neq 0$ and $q_n \mid d - m_n^2$. This is certainly true for $n = 0$ by hypothesis, taking $m_0 = m$ and $q_0 = q$, as $u = G_n(u)$. Supposing the claim holds for n, let $a_n = x_n(u) = \lfloor G_n(u) \rfloor$. Then we have

$$\begin{aligned} G_{n+1}(u) &= \frac{1}{G_n(u) - a_n} = \frac{q_n}{m_n - a_n q_n + \sqrt{d}} \\ &= \frac{q_n}{m_n - a_n q_n + \sqrt{d}} \cdot \frac{-(m_n - a_n q_n) + \sqrt{d}}{-(m_n - a_n q_n) + \sqrt{d}} \\ &= \frac{-(m_n - a_n q_n) + \sqrt{d}}{(d - (m_n - a_n q_n)^2)/q_n} \ . \end{aligned} \tag{81}$$

It is clear that $m_{n+1} := -(m_n - a_n q_n) \in \mathbb{Z}$. If q_{n+1} is the last denominator above, then

$$q_n q_{n+1} = d - m_{n+1}^2 = d - (m_n - a_n q_n)^2 = (d - m_n^2) + q_n(2a_n m_n - a_n^2 q_n). \tag{82}$$

Since $q_n \mid d - m_n^2$, $q_{n+1} \in \mathbb{Z}$; further, q_{n+1} divides $d - m_{n+1}^2$, as required, completing the proof of the claim.

We can solve for $G_n(u)$ in terms of u, using (7):

$$G_n(u) = -\frac{p_{n-1} - u q_{n-1}}{p_n - u q_n} = -\frac{q_n}{q_{n-1}} \cdot \frac{\xi_{n-2}(u) - u}{\xi_{n-1}(u) - u}. \tag{83}$$

We now invoke some algebraic number theory: conjugate is

$$\bar{u} = \frac{m - \sqrt{d}}{q}. \tag{84}$$

Take the conjugate of both sides in (11), observing that $p_i, q_i \in \mathbb{Z}$ (so $\xi_i(u) \in \mathbb{Q}$), and keeping alert to the fact that, in general $\overline{G_n(u)} \neq G_n(\bar{u})$. Then

$$\overline{G_n(u)} = -\frac{q_n}{q_{n-1}} \cdot \frac{\xi_{n-2}(u) - \bar{u}}{\xi_{n-1}(u) - \bar{u}}. \tag{85}$$

Since $\xi_n(u) \to u$ and since $u \neq \bar{u}$,

$$\frac{\xi_{n-2}(u) - \bar{u}}{\xi_{n-1}(u) - \bar{u}} \to \frac{u - \bar{u}}{u - \bar{u}} = 1. \tag{86}$$

Because $q_n, q_{n-1} > 0$, we may conclude that there exists N so that for $n \geq N$,

$$\overline{G_n(u)} = \frac{m_n - \sqrt{d}}{q_n} < 0. \tag{87}$$

Since $G_n(u) \geq 1$ for $n \geq 1$ by definition, we have

$$G_n(u) - \overline{G_n(u)} = \frac{2\sqrt{d}}{q_n} > 0 \tag{88}$$

for $n \geq N$, hence $q_n > 0$ for $n \geq N$. But now recall from (10) that

$$q_n q_{n+1} + m_{n+1}^2 = d \implies q_n,\ m_{n+1}^2 < d. \tag{89}$$

This implies that there are only finitely many possibilities for (q_n, m_n) for $n \geq N$ and so there exists $M \geq N, r > 0$ so that $(q_M, m_M) = (q_{M+r}, m_{M+r})$. It follows that $G_M(u) = G_{M+r}(u)$, and so, for $n \geq M$, $G_n(u) = G_{n+r}(u)$, implying that $a_n = a_{n+r}$. That is, the continued fraction expression for u is periodic with period r. □

Example. Let

$$u = G_0(u) = \frac{1+\sqrt{2}}{3} = \frac{3+\sqrt{18}}{9} \approx .805. \tag{90}$$

Then $a_0 = 0$ and

$$G_1(u) = \frac{9}{3+\sqrt{18}} = -3+\sqrt{18} \approx 1.242. \tag{91}$$

We see that $a_1 = 1$ and

$$G_2(u) = \frac{1}{-3+\sqrt{18}-1} = \frac{1}{-4+\sqrt{18}} = \frac{4+\sqrt{18}}{2} \approx 4.121. \tag{92}$$

Therefore, $a_2 = 4$ and

$$G_3(u) = \frac{1}{\frac{4+\sqrt{18}}{2}-4} = \frac{2}{-4+\sqrt{18}} = 4+\sqrt{18} \approx 8.243. \tag{93}$$

We're almost done, because $a_3 = 8$ and

$$G_4(u) = \frac{1}{4+\sqrt{18}-8} = \frac{1}{-4+\sqrt{18}} = G_2(u). \tag{94}$$

It follows that $u = [0, 1, \overline{4, 8}]$. We can verify this by our earlier calculation. We have previously shown that, if $\theta = [\overline{4, 8}]$, then

$$8\theta^2 - 32\theta - 4 = 0 \implies \theta = \frac{4+3\sqrt{2}}{2}. \tag{95}$$

It follows that

$$x = 0 + \cfrac{1}{1+\cfrac{1}{4+\cfrac{1}{8+\cfrac{1}{\cdots}}}} = \cfrac{1}{1+\cfrac{1}{\theta}} = \cfrac{1}{1+\cfrac{2}{4+3\sqrt{2}}} = \frac{4+3\sqrt{2}}{6+3\sqrt{2}} = u. \tag{96}$$

Just for completeness, we observe that

$$\frac{1+\sqrt{2}}{3} = 0 + \cfrac{1}{1+\cfrac{1}{4+\cfrac{1}{4+\sqrt{18}}}} \implies \frac{1-\sqrt{2}}{3} = 0 + \cfrac{1}{1+\cfrac{1}{4+\cfrac{1}{4-\sqrt{18}}}}. \tag{97}$$

The actual continued fraction representation for $\bar{u}$ is $[-1,1,6,\overline{4,8}]$, a formula whose resemblance to the earlier one might make want you to explore more deeply in the subject of continued fractions.

We also state a pretty neat theorem that we will not use directly. A proof can be found in your favorite high-quality number theory textbook:

Theorem 15. *The real quadratic irrational u has a purely periodic continued fraction expansion if and only if $u > 1$ and $-1 < \bar{u} < 0$.*

(Note that for $\theta = [\overline{4,8}]$, $\theta \approx 4.121$ and $\bar{\theta} \approx -.121$.) More generally, for $\theta = [\overline{a,b}]$, θ and $\bar{\theta}$ are the positive and negative roots of the equation

$$p_{a,b}(T) = bT^2 - abT - a = 0, \tag{98}$$

and $p_{a,b}(-1) = b + a(b-1) > 0 > -a = p_{a,b}(0)$, so $\bar{\theta} \in (-1,0)$, and $p_{a,b}(1) = -b(a-1) - a < 0$, so $\theta > 1$. All bets are off if you want to play with periodic continued fractions whose denominators are not in $\mathbb{N}$.

Corollary 16. *If $m \in \mathbb{N}$ is not a perfect square and $a_0 = \lfloor\sqrt{m}\rfloor$, then $\sqrt{m} = [a_0, \overline{a_1,\dots,a_{d-1},2a_0}]$.*

Proof. Observe that $a_0^2 + 1 \le m \le (a_0+1)^2 - 1$. If we let $u = a_0 + \sqrt{m}$, then $u \ge 2a_0 > 1$ and $\bar{u} = a_0 - \sqrt{m}$. Thus

$$0 > a_0 - \sqrt{a_0^2+1} \ge \bar{u} > a_0 - (a_0+1) = -1, \tag{99}$$

verifying the hypotheses of Theorem 3. □

In a few cases, the non-periodic continued fraction representations of non-quadratic irrationals is known. Here is a very brief sampling:

$$\begin{aligned} e &= [2,1,2,1,1,4,1,1,6,1,1,8,1,1,\dots], \\ \frac{e-1}{e+1} &= [0,2,6,10,14,18,22,26,30,34\dots], \\ \tan(1) &= [1,1,1,3,1,5,1,7,1,9,1,11,\dots], \\ \pi &= [3,7,15,1,292,1,1,1,2,1,3,1,14,2,\dots]. \end{aligned} \tag{100}$$

If you can find the pattern for π, wave to me from the stage as you pick up your Fields Medal.

5. A little analysis

We will come across the same situation more than once, so I'd like to formalize it.

Theorem 17. *Suppose X is a dense subset of the real interval $[a,b]$, Y is a dense subset of the real interval $[c,d]$ and f is a strictly increasing bijection of X onto Y. Then f extends to a unique continuous function F from $[a,b]$ to $[c,d]$; this function F is also strictly increasing.*

Proof. We first remark that if $a \in X$, then $f(a) = c$. This is because $f(X) = Y \subseteq [f(a), d]$ and Y is dense in $[c, d]$. A similar argument shows that if $b \in X$, then $f(b) = d$.

We now define $F(x)$ for $u \in [a, b]$. If $a \notin X$, $F(a) = c$ and if $b \notin X$, $F(b) = d$. If $u \in (a, b)$,

$$F_-(u) = \sup\{f(x) : x \in X \cap [a, u]\}; \qquad F_+(u) = \inf\{f(x) : x \in X \cap [u, b]\}. \tag{101}$$

The sup and inf are finite, because $c \le f(a), f(a') \le d$ for $a, a' \in X$. If $x \in X$, then clearly $F_-(x) = F_+(x) = f(x)$, so F extends f. Suppose $u \notin X$. Observe that if $x_0 < u < x_1$ for $x_i \in X$, then $f(x_0) < f(x_1)$, hence $f(x_0) \le F_+(u)$ and $F_-(u) \le f(x_1)$. Taking sups or infs, we see that $F_-(u) \le F_+(u)$. In fact, we show that $F_-(u) = F_+(u)$. Suppose not, then by the denseness of $Y = f(X)$, there exists $z \in X$ so that $f(z)$ is in the open interval $(F_-(u), F_+(u))$. But $f(z) > F_-(x)$ so $x > u$ and $f(z) < F_+(x)$ so $x < u$, a contradiction.

We now define

$$F(u) = F_-(u) = F_+(u). \tag{102}$$

We must prove that F is strictly increasing, continuous and unique with these conditions. First observe that F is non-decreasing, because $a \le u < v \le b$ implies that

$$\{f(x) : x \in X \cap [a, u]\} \subseteq \{f(x) : x \in X \cap [a, v]\}. \tag{103}$$

This set inclusion implies that $F_+(u) \le F_+(v)$. Since X is dense in $[a, b]$, there exist $x_i \in X$ so that $u < x_0 < \frac{u+v}{2} < x_1 < v$. Since f is strictly increasing, it follows that

$$F(u) \le F(x_0) = f(x_0) < f(x_1) = F(x_1) \le F(v), \tag{104}$$

so F is strictly increasing.

The continuity of F on (a, b) can be proved from the definition. Suppose that $F(u) = y$ and let $\epsilon > 0$ be small enough that $(y - \epsilon, y + \epsilon) \subseteq (c, d)$. As before, there exist $x_i \in X$ so that $y - \epsilon < f(x_0) < y < f(x_1) < y + \epsilon$. Since F is strictly increasing, $x_0 < u < x_1$. Let $\delta = \min\{u - x_0, x_1 - u\}$. Then $|u - v| < \delta$ implies that $x_0 < v < x_1$, so $f(x_0) < F(v) < f(x_1)$, and so $|F(u) - F(v)| < \epsilon$.

Finally, suppose G also extends F and is continuous, and suppose $F(u) \ne G(u)$ for some $u \in [a, b]$, say $F(u) > G(u)$. Then by continuity, $F(x) > G(x)$ on some neighborhood of u intersected with $[a, b]$. But this neighborhood will contain $x \in X$, and $F(x) = G(x)$ by hypothesis, a final contradiction. □

Theorem 18. *The Minkowski ?-function extends to a strictly increasing function on* $[0, 1]$ *defined on irrational arguments by*

$$?([0, a_0, a_1, a_2, \dots]) = \sum_{m=1}^{\infty} \frac{(-1)^{m-1}}{2^{a_1 + \cdots + a_m - 1}}. \tag{105}$$

This function maps the rationals in $[0, 1]$ *to the dyadic rationals in* $[0, 1]$ *and the quadratic irrationals in* $[0, 1]$ *to the non-dyadic rationals in* $[0, 1]$.

Proof. First observe that since $a_i > 1$, the series above converges by the ratio test. Let $[a,b] = [c,d] = [0,1]$ in Theorem 5, let $X = \mathbb{Q} \cap [0,1]$ and $Y = \mathbb{Q}_2 \cap [0,1]$, where $\mathbb{Q}_2$ denotes the dyadic rationals. Then X and Y are both dense on $[0,1]$ and ? has already been shown to be strictly increasing on X. If $t = [0, a_0, a_1, \dots]$, then $t = \lim \xi_n(t)$ and $?(\xi_n(t))$ is the n-th partial sum of the series, and since ? extends to a continuous function on $[0,1]$, which we shall also call ?,

$$?(t) = \lim_{n\to\infty} \xi_n(t) = \lim_{n\to\infty} \sum_{m=1}^{n} \frac{(-1)^{m-1}}{2^{a_1+\cdots+a_m-1}} = \sum_{m=1}^{\infty} \frac{(-1)^{m-1}}{2^{a_1+\cdots+a_m-1}}. \tag{106}$$

To show periodicity, we first remark that

$$[0, b_0, \dots, b_{e-1}, \overline{a_0, \dots, a_{d-1}}] = [0, b_0, \dots, b_{e-1}, \overline{a_0, \dots, a_{d-1}, a_0, \dots, a_{d-1}}], \tag{107}$$

so we may assume that any periodic continued fraction has even period. If u is a quadratic irrational and $D = \sum a_j$ is the sum of the denominators in the period and $2s$ is the length of the period, then there exists $N = 2sn_0$ so that for $m \ge N$, if $m = (2s)n + r$, $0 \le r < 2s$, then

$$a_1 + \cdots + a_m - 1 = T_r + (n - n_0)D \tag{108}$$

for some integers T_r, determined by the non-purely periodic part and $a_0, \dots, a_r$. It follows that

$$\sum_{m=1}^{\infty} \frac{(-1)^{m-1}}{2^{a_1+\cdots+a_m-1}} = \sum_{m=1}^{2sn_0-1} \frac{(-1)^{m-1}}{2^{a_1+\cdots+a_m-1}} + \sum_{r=0}^{2s-1} \sum_{k=0}^{\infty} \frac{(-1)^r}{2^{T_r+kD}}. \tag{109}$$

Since

$$\sum_{k=0}^{\infty} \frac{1}{2^{T_r+kD}} = \frac{1}{2^{T_r}} \cdot \frac{2^D}{2^D - 1}, \tag{110}$$

$?(t)$ is a finite sum of dyadic rationals and rationals, and so is rational.

Conversely, suppose $v \in (0,1) \cap \mathbb{Q}$ is not dyadic. We may write

$$v = \frac{p}{q} = \frac{1}{2^n}\frac{p}{q'} = \frac{1}{2^n}\left(c + \frac{p'}{q'}\right), \tag{111}$$

where q' is odd, $0 \le c < 2^n \in \mathbb{N}$ and $0 < p' < q'$. Since q' is odd, there exists r so that $2^r \equiv 1 \pmod{q'}$, and so

$$v = \frac{m}{2^n} + \frac{t}{2^r - 1}. \tag{112}$$

As before, we can write both t and c in the form

$$2^{b_1} - 2^{b_2} + - \cdots + 2^{b_{2k-1}} - 2^{b_{2k}} \tag{113}$$

with $b_1 > b_2 > \dots b_{2k}$, and since

$$\frac{1}{2^r - 1} = \sum_{j=1}^{\infty} \frac{1}{2^{rj}}, \tag{114}$$

we obtain a representation

$$v = \sum_{m=1}^{\infty} \frac{(-1)^{m-1}}{2^{a_1+\cdots+a_m-1}} \tag{115}$$

in which the a_k's are eventually periodic. □

Corollary 19. *If* $u \in [0,1]$, *then*

$$\begin{gathered} ?(u)+?(1-u)=1, \qquad ?\left(\frac{u}{u+1}\right)=\frac{?(u)}{2}, \qquad ?\left(\frac{1}{u+1}\right)=1-\frac{?(u)}{2} \\ ?\left(\frac{v}{1-kv}\right)=2^k\cdot ?(v), \quad 0\le v\le \frac{1}{k+1}. \end{gathered} \tag{116}$$

Proof. The equations on the top row hold when $u = x$ is rational. If $u \notin \mathbb{Q}$, write $u = \lim x_n$ with $x_n \to u$. Since ? is continuous and $\phi_j(t) = 1-t, \frac{t}{t+1}, \frac{1}{t+1}$ are continuous, simply replace u with x_n and take the limit.

For the bottom, first observe that if $0 \le v \le \frac{1}{2}$, then $v = \frac{u}{u+1}$ for some $u \in [0,1]$, and $u = \frac{v}{1-v}$. If $T(z) = \frac{z}{z+1}$, then the k-th iterate is $T_k(z) = \frac{z}{kz+1}$, so that

$$?\left(\frac{u}{ku+1}\right) = \frac{?(u)}{2^k} \implies ?\left(\frac{v}{1-kv}\right) = 2^k\cdot ?(v), \tag{117}$$

provided $\frac{v}{1-kv} \in [0,1]$; that is, $v \in [0, \frac{1}{k+1}]$. □

Example. We have seen above that $\frac{1+\sqrt{2}}{3} = [0,1,\overline{4,8}]$. It follows that

$$\begin{aligned} ?\left(\frac{1+\sqrt{2}}{3}\right) &= \frac{1}{2^{1-1}} - \frac{1}{2^{1+4-1}} + \frac{1}{2^{1+4+8-1}} - \frac{1}{2^{1+4+8+4-1}} + - \dots \\ &= 1 - \frac{1}{2^4} + \frac{1}{2^{12}} - \frac{1}{2^{16}} = \left(1-\frac{1}{2^4}\right)\cdot\frac{2^{12}}{2^{12}-1} = \frac{15\cdot 4096}{16\cdot 4095} = \frac{256}{273}. \end{aligned} \tag{118}$$

Example. We have already seen in Notes, V(37), that

$$?\left(\frac{F_n}{F_{n+1}}\right) = \frac{2}{3} + \frac{(-1)^{n+1}}{3\cdot 2^{n-1}}. \tag{119}$$

By taking the limit, we see that

$$?\left(\frac{\sqrt{5}-1}{2}\right) = \frac{2}{3}. \tag{120}$$

More generally, since

$$\theta_a = \frac{a+\sqrt{a^2+4}}{2} = [\overline{a}], \tag{121}$$

by taking the limit in Notes, V(36) we see that for $a \in \mathbb{N}$,

$$?\left(\frac{\sqrt{4+a^2}-a}{2}\right) = \frac{2}{2^a+1}. \tag{122}$$

It is useful to note a shortcut. define γ_a so that

$$?(\gamma_a) = \frac{1}{2^a+1} \implies ?\left(\frac{\gamma_a}{1-a\gamma_a}\right) = \frac{2^a}{2^a+1} = ?(1-\gamma a). \tag{123}$$

It follows that γ_a is a root of the equation

$$\frac{X}{1-aX} = 1-X \implies aX^2-(a+2)X+1=0 \implies \gamma_a = \frac{a+2-\sqrt{a^2+4}}{2a}. \tag{124}$$

(Since the polynomial has a root between 1 and 2, γ_a must be the smaller root.) Since

$$\frac{\gamma_a}{1-\gamma_a} = \theta_a, \tag{125}$$

this is consistent with the earlier calculation.

Remark. We remark that there is an implicit algorithm at work here. If $\frac{p}{q}$ is a non-dyadic rational, keep doubling until you get past $\frac{1}{2}$. Then take $1-x$ to get below $\frac{1}{2}$, keep doubling again, etc. These fractions all have denominator q and lie in $(0,1)$, hence eventually you get repetition. By solving the resulting equation, you can compute $?^{-1}(\frac{p}{q})$ without having to compute dyadic expansions and continued fractions explicitly. (This information is implicitly contained in the ordering of doubling and folding back.) This ought to be a theorem. Later.

Example. A more complicated object is $\theta_{a,b} = [\overline{a,b}]$. We already know that

$$\theta_{a,b} = \frac{ab+\sqrt{a^2b^2+4ab}}{2b} \implies \theta_{a,b}^{-1} = \frac{-ab+\sqrt{a^2b^2+4ab}}{2a}, \tag{126}$$

The computation of $?(\theta_{a,b})$ is easy because the period has even length. We have

$$?\left(\frac{-ab+\sqrt{a^2b^2+4ab}}{2a}\right) = \sum_{j=0}^{\infty}\left(\frac{1}{2^{(j+1)a+jb-1}} - \frac{1}{2^{(j+1)a+(j+1)b-1}}\right) = \frac{2(2^b-1)}{2^{a+b}-1}. \tag{127}$$

Example. It is a little-appreciated fact that every integer ≤ 10 can be written in the form $2^a(2^b \pm 1)$. It is also little-appreciated that the inverse to the ?-function is also a continuous and strictly-increasing function. Based on the results on the previous pages, we can compute $?^{-1}(\frac{a}{b})$ for every proper fraction in $(0,1)$ with $b \le 10$. For $b = 2,4,8$, see Notes V(13). In view of (33) and (36) above, we find immediately that

$$?\left(\frac{3-\sqrt{5}}{2}\right) = \frac{1}{3}, \qquad ?\left(\frac{5-\sqrt{5}}{10}\right) = \frac{1}{6}, \qquad ?\left(\frac{5+\sqrt{5}}{10}\right) = \frac{5}{6}. \tag{128}$$

Taking $a = 2$ in (38), we get

$$?(\sqrt{2}-1) = \frac{2}{5} \implies ?(2-\sqrt{2}) = \frac{3}{5}, \quad ?\left(\frac{2-\sqrt{2}}{2}\right) = \frac{1}{5}, \quad ?\left(\frac{\sqrt{2}}{2}\right) = \frac{4}{5}. \tag{129}$$

Working backwards from the right-hand-side of (55), with $(a, b) = (2, 1)$ and $(1, 2)$, we see after some simplification that

$$?\left(\frac{-1+\sqrt{3}}{2}\right) = \frac{2}{7}, \qquad ?\left(-1+\sqrt{3}\right) = \frac{6}{7}. \tag{130}$$

More use of Corollary 7 implies that

$$?\left(\frac{3-\sqrt{3}}{2}\right) = \frac{5}{7}, \quad ?\left(2-\sqrt{3}\right) = \frac{1}{7}, \quad ?\left(\frac{\sqrt{3}}{3}\right) = \frac{4}{7}, \quad ?\left(\frac{3-\sqrt{3}}{3}\right) = \frac{3}{7}. \tag{131}$$

Here's a preview from the next homework set: compute $?^{-1}(\frac{a}{b})$ for reduced fractions in $(0, 1)$ with $b = 9$ and $b = 10$.

6. Return to the first section

It turns out that it will be useful to have a proof of Theorem 3, and more.

Proof of Theorem 3. Suppose $u = [\overline{a_0, \dots, a_{n-1}}]$. By repeating the block if necessary, we may assume that $n \ge 3$. Then we've seen in the earlier notes that

$$u = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\dots + \cfrac{1}{a_{n-1} + \cfrac{1}{u}}}} = \frac{p_{n+1}(a_0, \dots, a_{n-1}, u)}{p_n(a_1, \dots, a_{n-1}, u)} \tag{132}$$
$$= \frac{u \cdot p_n(a_0, \dots, a_{n-1}) + p_{n-1}(a_0, \dots, a_{n-2})}{u \cdot p_{n-1}(a_1, \dots, a_{n-1}) + p_{n-2}(a_1, \dots, a_{n-2})} .$$

If, as before, we write this as

$$u = \frac{Au + B}{Cu + D}, \tag{133}$$

then we've seen that u is a (positive) root of the quadratic

$$f(T) = CT^2 + (D - A)T - B \implies u = \frac{A - D + \sqrt{(D-A)^2 + 4BC}}{2C}, \tag{134}$$

and the other root of f must be $\bar{u}$. We have $u > a_0 \ge 1$. Note that $f(0) = -B < 0$ and $f(-1) = (A - B) + (C - D)$, and

$$A - B = p_n(a_0, \dots, a_{n-1}) - p_{n-1}(a_0, \dots, a_{n-2}) \tag{135}$$
$$= (a_{n-1} - 1)p_{n-1}(a_0, \dots, a_{n-2}) + p_{n-2}(a_0, \dots, a_{n-3}) > 0.$$

Similarly, $C - D = (a_{n-1} - 1)p_{n-2}(a_1, \dots, a_{n-2}) + p_{n-3}(a_1, \dots, a_{n-3}) \ge 0$. Thus, $-1 < \bar{u} < 0$. (Alternatively, one can show that $u\bar{u} = \frac{-B}{C} \in (-1, 0)$.)

The converse is somewhat harder. We suppose that $u > 1$ and $\bar{u} \in (-1, 0)$. Since u is a quadratic irrational, its continued fraction expansion is periodic. Thus there exist n_0, r so that $G_{n_0}(u) = G_{n_0+r}(u)$, and so that $G_n(u) = G_{n+r}(u)$ for $n \ge n_0$. Assume without loss of generality that n_0 is minimal with this property. Our goal is to show that $n_0 = 0$.

To do this, we first wish to show that for all $n \ge 0$, we have $G_n(u) > 1$ and $\overline{G_n(u)} \in (-1, 0)$. The first is easy to establish, because, $G_0(u) = u > 1$ by hypothesis, and $G_n(u) > 1$ for $n \ge 1$ by construction. For the second, again, $\overline{G_n(u)} = \bar{u} \in (-1, 0)$ by hypothesis. We have

$$G_n(u) = a_n + \frac{1}{G_{n+1}(u)} \tag{136}$$

and this implies that

$$\overline{G_n(u)} = a_n + \frac{1}{\overline{G_{n+1}(u)}} \implies \overline{G_{n+1}(u)} = \frac{1}{\overline{G_n(u)} - a_n} \ . \tag{137}$$

By hypothesis, $\overline{G_n(u)} \in (-1, 0)$ and $a_n \geq 1$, hence

$$\overline{G_n(u)} - a_n < -1 \implies \overline{G_{n+1}(u)} = \frac{1}{\overline{G_n(u)}} - a_n \in (-1, 0). \tag{138}$$

This completes the induction. Observe that we have now shown that

$$-a_n - \frac{1}{\overline{G_{n+1}(u)}} = -\overline{G_n(u)} \in (0, 1), \tag{139}$$

and since a_n is an integer, this means that

$$a_n = \left\lfloor -\frac{1}{\overline{G_{n+1}(u)}} \right\rfloor. \tag{140}$$

Suppose that $G_{n_0}(u) = G_{n_0+r}(u)$ and $n_0 \geq 1$. It follows from this last equation that $a_{n_0-1} = a_{n_0+r-1}$. But now, applying (64) with $n = n_0 - 1$ and $n = n_0 + r - 1$, we see that $G_{n_0-1}(u) = G_{n_0+r-1}(u)$, contradicting the minimality of n_0. Thus, u is purely periodic. □

The next result is interesting in its own right.

Theorem 20.

$$u = [\overline{a_0, \ldots, a_{n-1}}] \implies -\bar{u}^{-1} = [\overline{a_{n-1}, \ldots, a_0}]. \tag{141}$$

Proof. Keep the notation of the last theorem. Then $u, \bar{u}$ are the roots of $f(T) = 0$. By the same reasoning, if $v = [\overline{a_{n-1}, \ldots, a_0}]$, then

$$v = \frac{v \cdot p_n(a_{n-1}, \ldots, a_0) + p_{n-1}(a_{n-1}, \ldots, a_1)}{v \cdot p_{n-1}(a_{n-2}, \ldots, a_0) + p_{n-2}(a_{n-2}, \ldots, a_1)} \ . \tag{142}$$

But by the reversability of the arguments of continuants, this means that v and $\bar{v}$ are the roots of the equation

$$g(T) = BT^2 + (D - A)T - C = -T^2 f(-T^{-1}). \tag{143}$$

Thus $\{v, \bar{v}\} = \{-u^{-1}, -\bar{u}^{-1}\}$. Since $u, v > 1$, $\bar{u}, \bar{v} \in (-1, 0)$, it follows that $v = -\bar{u}^{-1}$. □

We now improve Corollary 4.

Corollary 21. *If $m \in \mathbb{N}$ is not a perfect square, and $a_0 = \lfloor \sqrt{m} \rfloor$, then $\sqrt{m} = [a_0, \overline{a_1, \ldots, a_{d-1}, 2a_0}]$, where $a_k = a_{d-k}$ for $1 \leq k \leq d - 1$.*

Proof. We have already proved that $u = a_0 + \sqrt{m}$ is purely periodic, and since $\lfloor u \rfloor = 2a_0$, we have

$$u = [\overline{2a_0, a_1, \dots, a_{d-1}}] \tag{144}$$

for some denominators a_j. We now take away the first denominator and observe that

$$u = 2a_0 + \cfrac{1}{a_1 + \cfrac{1}{\cdots + \cfrac{1}{a_{d-1} + \cfrac{1}{2a_0 + \cfrac{1}{\cdots}}}}} = 2a_0 + \frac{1}{[\overline{a_1, \dots, a_{d-1}, 2a_0}]}. \tag{145}$$

Thus

$$w := [\overline{a_1, \dots, a_{d-1}, 2a_0}] = \frac{1}{u - 2a_0} = \frac{1}{\sqrt{m} - a_0}. \tag{146}$$

It follows from Theorem 8 that

$$\begin{aligned} v := [\overline{2a_0, a_{d-1}, \dots, a_1}] &= -\bar{w}^{-1} = -\cfrac{1}{\cfrac{1}{-\sqrt{m} - a_0}} \\ &= a_0 + \sqrt{m} = [\overline{2a_0, a_1, \dots, a_{d-1}}] = u. \end{aligned} \tag{147}$$

Since continued fractions are unique, we conclude that $a_k = a_{d-k}$ for $1 \le k \le d - 1$. □

7. Bonus leftovers

This accidental section consists of a few items that should have been in section two. The first is a definition of the Minkowski ?-function which is seemingly independent of anything connected with the Stern sequence.

Corollary 22. *The Minkowski ?-function is the unique function F satisfying the following properties:*

(1) *F is continuous on $[0,1]$;*
(2) *$F(0) = 0$, $F(1) = 1$;*
(3) *$F(x) + F(1-x) = 1$ for all $x \in [0,1]$;*
(4) *$F(\frac{x}{x+1}) = \frac{F(x)}{2}$ for all $x \in [0,1]$.*

Proof. Properties (2), (3) and (4) iterate to showing that $F(x) = ?(x)$ for any $x \in \mathbb{Q}$, and so by (1), F must be the unique continuous extension. ◻

Example. The approach in the last remark on p.8 is less useful than perhaps it seems. What it amounts to is a way of quickly determining the repeating portion of the binary representation of $y = \frac{p}{q}$, and then translating that into finding $?^{-1}(y)$. If $q = 2^v m$ for odd m, then the number of steps is v plus the smallest k so that $2^k \equiv \pm 1 \pmod{m}$.

Recall that if $2p < q$, then

$$?\left(\frac{p}{q}\right) = y \iff ?\left(\frac{q-p}{q}\right) = 1-y,\ ?\left(\frac{p}{q-p}\right) = 2y,\ ?\left(\frac{q-2p}{q-p}\right) = 1-2y. \tag{148}$$

We apply this repeatedly to one not-so-easy illustrative example. Suppose

$$?(x) = \frac{13}{44} \iff ?(1-x) = \frac{31}{44}. \tag{149}$$

Since the first of these is less than $\frac{1}{2}$, we "use" it:

$$?\left(\frac{x}{1-x}\right) = \frac{13}{22} \iff ?\left(\frac{1-2x}{1-x}\right) = \frac{9}{22}. \tag{150}$$

Now $\frac{9}{22} < \frac{1}{2}$, so

$$?\left(\frac{1-2x}{x}\right) = \frac{9}{11} \iff ?\left(\frac{3x-1}{x}\right) = \frac{2}{11}. \tag{151}$$

We can already see from this that $x \in (\frac{1}{3}, \frac{1}{2})$. Here, $\frac{2}{11} < \frac{1}{2}$, so

$$?\left(\frac{3x-1}{1-2x}\right) = \frac{4}{11} \iff ?\left(\frac{2-5x}{1-2x}\right) = \frac{7}{11}. \tag{152}$$

Repeating,

$$?\left(\frac{3x-1}{2-5x}\right) = \frac{8}{11} \iff ?\left(\frac{3-8x}{2-5x}\right) = \frac{3}{11}. \tag{153}$$

Don't give up!

$$?\left(\frac{3-8x}{3x-1}\right)=\frac{6}{11}\iff ?\left(\frac{11x-4}{3x-1}\right)=\frac{5}{11}. \tag{154}$$

We now know that $x\in(\frac{4}{11},\frac{3}{8})$. Almost there!

$$?\left(\frac{11x-4}{3-8x}\right)=\frac{10}{11}\iff ?\left(\frac{7-19x}{3-8x}\right)=\frac{1}{11}. \tag{155}$$

(Thus, $\frac{4}{11}=.3636..<x<.3684..=\frac{7}{19}$.) Finally:

$$?\left(\frac{7-19x}{11x-4}\right)=\frac{2}{11}\iff ?\left(\frac{30x-11}{11x-4}\right)=\frac{9}{11}. \tag{156}$$

At last, a match!

$$?\left(\frac{3x-1}{x}\right)=?\left(\frac{7-19x}{11x-4}\right)=\frac{2}{11}\implies \frac{3x-1}{x}=\frac{5-11x}{7x-3}. \tag{157}$$

This gives a quadratic, which is good:

$$\begin{aligned}&(3x-1)(7x-3)=x(5-11x)\implies 52x^2-30x+4=0\\ &\implies x=\frac{15+\sqrt{17}}{52}\approx .36775,\quad x=\frac{15-\sqrt{17}}{52}\approx .20917.\end{aligned} \tag{158}$$

There are two roots, but they imply that

$$\frac{3x-1}{x}=\frac{\pm 17-3}{4}\implies ?\left(\frac{\pm 17-3}{4}\right)=\frac{2}{11}. \tag{159}$$

We must choose the "+" sign; it's also the only one that is in the correct range (otherwise the formulas above are not accurate.) We conclude that

$$?\left(\frac{15+\sqrt{17}}{52}\right)=\frac{13}{44}. \tag{160}$$

Interestingly, it turns out that

$$?\left(\frac{15-\sqrt{17}}{52}\right)=\frac{3}{44}. \tag{161}$$

To confirm these formulas, we make some Mathematica calculations:

$$\frac{15+\sqrt{17}}{52}=[0,2,1,2,\overline{1,1,3}] \tag{162}$$

A calculation similar to that found in the proof of Theorem 6 shows that

$$\begin{aligned}\frac{13}{44}&=\frac{1}{4}\cdot\frac{52}{44}=\frac{1}{4}\left(1+\frac{2}{11}\right)=\frac{1}{4}+\frac{1}{4}\cdot\frac{186}{1023}\\ &=\frac{1}{2^2}+\frac{1}{2^2}\cdot\frac{2^8-2^7+2^6-2^3+2^1}{2^{10}-1}=[.01\overline{0010111010}]_2\ .\end{aligned} \tag{163}$$

These match up. It is worth noting as well that one obtains the following result from a more careful examination of this algorithm. If q is odd and $a \equiv \pm 2^k b \pmod{q}$, then $?^{-1}(\frac{a}{q})$ and $?^{-1}(\frac{b}{q})$ belong to the same quadratic extension of $\mathbb{Q}$. The smallest odd number q for which there exist a, b *not* satisfying this condition is $q = 17$. It's not hard to show that

$$?\left(\frac{3-\sqrt{5}}{4}\right) = \frac{1}{17}, \qquad ?\left(\frac{4-\sqrt{10}}{3}\right) = \frac{3}{17}. \tag{164}$$

Finally, we sketch the proof of a theorem, which I have given a somewhat disparaging name. It is quite striking but is really a trivial extension of what's already known. I don't remember having seen it before but would not be surprised to learn it was 100 years old.

Theorem 23 (Low hanging fruit)**.** *Suppose*

$$?(x) = \frac{2p}{q}. \tag{165}$$

Then p and q are both odd if and only if x is a quadratic irrational and $\bar{x} < -1$.

Proof. Let $u = 1/x \in (1, \infty)$. We have seen that u has a purely periodic continued fraction if and only if $\bar{u} \in (-1, 0)$, and $xu = 1$ implies $\bar{x}\bar{u} = 1$, hence $\bar{x} < -1$. We now want to show that $?(x) = \frac{2p}{q}$ with odd p, q if and only if u is purely periodic.

First suppose $u = [\overline{a_0, \ldots, a_{d-1}}]$. As remarked earlier, by repeating the period, we may assume without loss of generality that d is even. Let

$$T_j = \sum_{i=0}^{j} a_i, \qquad 0 \le j \le d-1. \tag{166}$$

By the periodicity of the a_i's, if $n = dk + j$, $0 \le j \le d-1$, then

$$\sum_{i=0}^{n} a_i = \sum_{i=0}^{dk+j} a_i = kT_{d-1} + T_j, \tag{167}$$

and from (33),

$$\begin{gathered} ?(u) = \sum_{n=0}^{\infty} \frac{(-1)^n}{2^{a_0+\cdots+a_n-1}} = \\ \sum_{j=0}^{d-1} \frac{(-1)^j}{2^{T_j-1}} \sum_{k=0}^{\infty} \frac{1}{2^{kT_{d-1}}} = \left(\sum_{j=0}^{d-1} \frac{(-1)^j}{2^{T_j-1}}\right)\left(\frac{2^{T_{d-1}}}{2^{T_{d-1}}-1}\right). \end{gathered} \tag{168}$$

The proof is complete upon the identification

$$p = \sum_{j=0}^{d-1} (-1)^j 2^{T_{d-1}-T_j}, \qquad q = 2^{T_{d-1}} - 1. \tag{169}$$

(Note that $T_{d-1} > T_j$ for $j < d-1$, hence p is odd.)

Conversely, suppose $?(x) = \frac{2p}{q}$, where p and q are odd. Since $q \mid 2^T - 1$ for some T, we may write

$$\frac{2p}{q} = \frac{2m}{2^T - 1} \tag{170}$$

for some odd m. We can now write

$$m = \sum_{j=0}^{d-1} (-1)^j 2^{T-T_j} \tag{171}$$

where d is even and the T_j's are strictly increasing, with $T_0 \geq 1$. Since m is odd, $T_{d-1} = T$. The argument of the last paragraph reverses, and we find that x is purely periodic. This proves that $\bar{x} < -1$, completing the proof. □

Remark. Checking against the examples on p.9, we see that for 1/3, 2/3, 1/6, 5/6,

$$\frac{-\sqrt{5}-1}{2} < -1, \quad \frac{3+\sqrt{5}}{2} > 1, \quad \frac{5 \pm \sqrt{5}}{10} > 0. \tag{172}$$

In checking $?^{-1}(\frac{a}{b})$ with $b = 5, 7$, we see easily that only $\frac{2}{5}, \frac{2}{7}, \frac{6}{7}$ are the only ones fitting the hypothesis of this theorem, and their images are the only ones whose conjugates are < -1.

Moreover, we've seen that for $\theta_{ab} = [\overline{a, b}]$,

$$?(\theta_{a,b}^{-1}) = \frac{2(2^b - 1)}{2^{a+b} - 1}; \tag{173}$$

fortunately,

$$\overline{\theta_{a,b}^{-1}} = \frac{-ab - \sqrt{a^2b^2 + 4ab}}{2a} < \frac{-2ab}{2a} \leq -1. \tag{174}$$

8. The differentiability of $?(x)$

First, an important correction – the error was noted by Jason Benda after class. I had made a mistake in an earlier draft and not completely replaced it in the final. You should read (85) and (86) as

$$?\left(\frac{3x-1}{x}\right) = ?\left(\frac{7-19x}{11x-4}\right) = \frac{2}{11} \implies \frac{3x-1}{x} = \frac{7-19x}{11x-4}. \tag{175}$$

$$(3x-1)(11x-4) = x(7-19x) \implies 52x^2 - 30x + 4 = 0. \tag{176}$$

Before talking about the differentiability of $?'(x)$, we need a more general consideration. Let $\mathcal{F}$ denote the set of continuous, strictly increasing functions from $[0,1]$ to itself. Observe that $f \in \mathcal{F}$ implies that the inverse $f^{-1} \in \mathcal{F}$ and that both ? and ¿ are in $\mathcal{F}$. (Another nod to Jason for finding the way for me to write the inverse.) As is well-known, if $f \in \mathcal{F}$, then f' exists almost everywhere on $[0,1]$.

For $f \in \mathcal{F}$, We define

$$\begin{aligned} D(f,0) &:= \left\{x_0 : \lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0} = 0\right\}, \\ D(f,\infty) &:= \left\{x_0 : \lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0} = \infty\right\}. \end{aligned} \tag{177}$$

Obviously, $D(f,0)$ is the set of points at which $f' = 0$; but $D(f,\infty)$ is not just the set of points at which f' does not exist. It is possible for the difference quotient to oscillate; we are interested in the set where it goes to ∞. Inasmuch as

$$x \to x_0 \iff f(x) \to f(x_0), \qquad y \to y_0 \iff f^{-1}(y) \to f^{-1}(y_0), \tag{178}$$

we have the immediate lemma.

Lemma 24. *If $f \in \mathcal{F}$, then*

$$\begin{aligned} x_0 \in D(f,0) &\iff f(x_0) \in D(f^{-1},\infty); \\ x_0 \in D(f,\infty) &\iff f(x_0) \in D(f^{-1},0). \end{aligned} \tag{179}$$

Proof. Simply observe that if $y = f(x)$, then

$$\frac{f(x)-f(x_0)}{x-x_0} = \left(\frac{f^{-1}(y)-f^{-1}(y_0)}{y-y_0}\right)^{-1}. \tag{180}$$

□

Because f' exists ae, we must have $\mu(D(f,\infty)) = 0$. However, it does not follow that $f(D(f,\infty)) = D(f^{-1},0)$ also has measure 0. This would imply by symmetry that $\mu(D(f,0)) = 0$ as well. But this is false in general. (For example, the Cantor function maps the Cantor set to $[0,1]$.) In fact, as Salem proves in his paper, $\mu(D(?,0)) = 1$.

Let us say that if $v_m \uparrow x$, $u_m \downarrow x$ and there exists $\lambda > 0$ so that $x - v_{m+1} > \lambda(x - v_m)$ and $u_{m+1} - x > \lambda(u_m - x)$, then (u_m) and (v_m) *slowly approach* x.

Lemma 25. *Suppose* (u_m) *and* (v_m) *slowly approach* x *and suppose* $f \in \mathcal{F}$. *Then*

$$
\begin{aligned}
\lim_{m\to\infty} \frac{f(u_m)-f(x)}{u_m-x} = \lim_{m\to\infty} \frac{f(v_m)-f(x)}{v_m-x} = 0 &\implies x \in D(f,0); \\
\lim_{m\to\infty} \frac{f(u_m)-f(x)}{u_m-x} = \lim_{m\to\infty} \frac{f(v_m)-f(x)}{v_m-x} = \infty &\implies x \in D(f,\infty).
\end{aligned} \tag{181}
$$

Proof. Observe that if $u > x$ is sufficiently close to x, then $u \in (x_{m+1}, x_m)$ for some m and, by the monotonicity of f,

$$
\begin{aligned}
\frac{f(u)-f(x)}{u-x} &< \frac{f(u_m)-f(x)}{u_{m+1}-x} < \frac{1}{\lambda} \cdot \frac{f(u_m)-f(x)}{u_m-x}, \\
\frac{f(u)-f(x)}{u-x} &> \frac{f(u_{m+1})-f(x)}{u_m-x} > \lambda \cdot \frac{f(u_{m+1})-f(x)}{u_{m+1}-x}.
\end{aligned} \tag{182}
$$

Similar inequalities apply if $v < x$. We see that if y is sufficiently close to x, then the "slow" sequential limits to 0 or ∞ imply continuous limits. It is easy to give counterexamples when the difference quotient is finite and non-zero, or when the sequences don't slowly approach x. The reverse implication always holds of course. □

Lemma 26. *(i) If* $x_0 \in D(?,0)$, *then* $1-x_0$, $\frac{x_0}{1+x_0} \in D(?,0)$; *if* $x_0 \in D(?,\infty)$, *then* $1-x_0$, $\frac{x_0}{1+x_0} \in D(?,\infty)$.

(ii) If $x_0 \in D(¿\,,0)$, *then* $1-x_0$, $\frac{1}{2}x_0 \in D(¿\,,0)$; *if* $x_0 \in D(¿\,,\infty)$, *then* $1-x_0$, $\frac{1}{2}x_0 \in D(¿\,,\infty)$.

Proof. Suppose $x \to x_0$. Then $1-x \to 1-x_0$ and

$$
\begin{aligned}
\lim_{x\to x_0} \frac{?(1-x)-?(1-x_0)}{(1-x)-(1-x_0)} &= \lim_{x\to x_0} \frac{(1-?(x))-(1-?(x_0))}{(1-x)-(1-x_0)} \\
&= \lim_{x\to x_0} \frac{?(x_0)-?(x)}{x_0-x} = \lim_{x\to x_0} \frac{?(x)-?(x_0)}{x-x_0}.
\end{aligned} \tag{183}
$$

Similarly, if $x \to x_0$, then $\frac{x}{1+x} \to \frac{x_0}{1+x_0}$ (since $x \in [0,1]$). Observe that

$$
\begin{aligned}
?\left(\frac{x}{1+x}\right) - ?\left(\frac{x_0}{1+x_0}\right) &= \frac{?(x)-?(x_0)}{2}; \\
\frac{x}{1+x} - \frac{x_0}{1+x_0} &= \frac{(x-x_0)}{(1+x)(1+x_0)}.
\end{aligned} \tag{184}
$$

Thus, upon taking the difference quotient, we find that

$$
\begin{aligned}
\lim_{x\to x_0} \frac{?(\frac{x}{1+x})-?(\frac{x_0}{1+x_0})}{\frac{x}{1+x}-\frac{x_0}{1+x_0}} &= \lim_{x\to x_0} \frac{(1+x)(1+x_0)}{2} \cdot \frac{?(x)-?(x_0)}{x-x_0} \\
&= \frac{(1+x_0)^2}{2} \lim_{x\to x_0} \frac{?(x)-?(x_0)}{x-x_0}.
\end{aligned} \tag{185}
$$

The second set of implications follow from Lemma 12. □

Theorem 27. *If $u \in (0,1) \cap \mathbb{Q}$, then $u \in D(?,0)$; that is, $?'(u) = 0$. If $v = \frac{p}{2^n} \in (0,1)$, then $v \in D(¿\,,\infty)$.*

Proof. We first consider $u = \frac{1}{2}$. Let

$$\begin{aligned} v_m &= \frac{1}{[0,2,m]} = \frac{m}{2m+1} \implies ?(v_m) = \frac{1}{2} - \frac{1}{2^{m+1}}, \\ u_m = 1 - v_m &= \frac{1}{[0,1,1,m]} = \frac{m+1}{2m+1} \implies ?(u_m) = \frac{1}{2} + \frac{1}{2^{m+1}}. \end{aligned} \tag{186}$$

Observe that $u_{m+1} - \frac{1}{2} = \frac{2m+1}{2m+3}(u_m - \frac{1}{2})$ and $\frac{1}{2} - v_{m+1} = \frac{2m+1}{2m+3}(\frac{1}{2} - v_m)$, so (u_m) and (v_m) slowly approach $\frac{1}{2}$. Further,

$$\frac{?(u_m) - ?(\frac{1}{2})}{u_m - \frac{1}{2}} = \frac{?(v_m) - ?(\frac{1}{2})}{v_m - \frac{1}{2}} = \frac{4m+2}{2^{m+1}} \to 0. \tag{187}$$

It follows that $\frac{1}{2} \in D(?,0)$.

As we have seen earlier in these notes, every rational number can be constructed from $\frac{1}{2}$ by repeated application of the maps $x \to 1 - x$ and $x \to \frac{x}{1+x}$, and this completes the proof by Lemma 14. Alternatively, we have $\frac{1}{2} \in D(¿\,,\infty)$, and every dyadic rational is derived from $\frac{1}{2}$ by repeated application of $x \to 1-x$ and $x \to \frac{x}{2}$. □

Theorem 28.

$$\frac{1}{\phi} = \frac{\sqrt{5}-1}{2} \in D(?,\infty) \implies \frac{2}{3} \in D(¿\,,0). \tag{188}$$

Proof. We have already shown in (47), (48) that

$$?\left(\frac{F_n}{F_{n+1}}\right) = \frac{2}{3} + \frac{(-1)^{n+1}}{3 \cdot 2^{n-1}}, \qquad ?\left(\frac{1}{\phi}\right) = \frac{2}{3}. \tag{189}$$

We note that, like all finite continued fraction approximations, the sequence F_n/F_{n+1} alternates above and below its limit:

$$\frac{F_0}{F_1} < \frac{F_2}{F_3} < \cdots < \frac{1}{\phi} < \cdots < \frac{F_3}{F_4} < \frac{F_1}{F_2}. \tag{190}$$

Recall the Binet formula for F_n, and the identities $\bar{\phi} = -\phi^{-1}$ and $\phi^2 + 1 = \sqrt{5}\phi$, so

$$\begin{aligned} z_n := \frac{F_n}{F_{n+1}} - \frac{1}{\phi} &= \frac{\frac{1}{\sqrt{5}}(\phi^n + (-1)^{n+1}\phi^{-n})}{\frac{1}{\sqrt{5}}(\phi^{n+1} + (-1)^{n+2}\phi^{-(n+1)})} - \frac{1}{\phi} \\ &= \frac{\phi^{2n+1} + (-1)^{n+1}\phi}{\phi^{2n+2} + (-1)^{n+2}} - \frac{1}{\phi} = \frac{(-1)^{n+1}\sqrt{5}}{\phi^{2n+2} + (-1)^{n+2}}. \end{aligned} \tag{191}$$

It follows that

$$\left|\frac{z_{n+1}}{z_n}\right| \to \frac{1}{\phi^2} > 0. \tag{192}$$

and so if

$$v_m = \frac{F_{2m}}{F_{2m+1}} \uparrow \frac{1}{\phi}, \qquad u_m = \frac{F_{2m+1}}{F_{2m+2}} \downarrow \frac{1}{\phi}, \tag{193}$$

then (u_m) and (v_m) slowly approach $\frac{1}{\phi}$. We also have

$$\frac{?\left(\frac{F_n}{F_{n+1}}\right) - ?\left(\frac{1}{\phi}\right)}{\frac{F_n}{F_{n+1}} - \frac{1}{\phi}} = \frac{\phi^{2n+2} + (-1)^{n+2}}{3\sqrt{5}\cdot 2^{n-1}} \to \infty, \tag{194}$$

since $\phi^2 = \frac{3+\sqrt{5}}{2} > 2$. The result now follows by Lemma 13. The divergence isn't particularly rapid. With $n = 10$, we have

$$\frac{?(\frac{55}{89}) - ?(\frac{1}{\phi})}{\frac{55}{89} - \frac{1}{\phi}} = \frac{\frac{341}{512} - \frac{2}{3}}{\frac{55}{89} - \frac{1}{\phi}} \approx \frac{-.000651}{-.000056} \approx 11.53. \tag{195}$$

□

Corollary 29. *If* $x = [0, a_0, \dots, a_n, \overline{1}]$ *for any* $a_j \in \mathbb{N}$*, then* $x \in D(?, \infty)$*. If* $u = \frac{p}{3\cdot 2^r} \in (0,1)$ *with* $\gcd(p,3) = 1$*, then* $u \in D($¿ $, 0)$*.*

Proof. This is a direct consequence of Lemma 14. The first part may be clearer than the second. If $u < \frac{1}{2}$, then $u = \frac{v}{2}$, where $v = \frac{p}{3\cdot 2^{r-1}}$. If $u > \frac{1}{2}$, then $u = 1 - \frac{v}{2}$, where $v = \frac{3\cdot 2^r - p}{3\cdot 2^{r-1}}$. □

Corollary 30. *The Minkowski* $?$*-function is nowhere continuously differentiable.*

Proof. The rationals of the form $\frac{p}{3\cdot 2^r}$ are dense in $[0,1]$, hence so are the quadratic irrationals of the form $[0, a_0, \dots, a_n, \overline{1}]$. It follows that every open interval of $(0,1)$ contains points where $?' = 0$ and points where $?'$ is not differentiable. □

We sketch a stab at generalizing Theorem 16. Fix $a, b \in \mathbb{N}$ and let

$$\theta_{a,b} = \frac{1}{[\overline{a,b}]} = \frac{-ab + \sqrt{a^2b^2 + 4ab}}{2a}. \tag{196}$$

(This was discussed on p.9 of this supplement, see (54) and (55).) More specifically, define sequences $(u_m), (v_m)$ as above, with

$$\begin{aligned} u_0 = 1, \quad u_m &= \cfrac{1}{a + \cfrac{1}{b + u_{m-1}}} := \frac{p_{2m}}{q_{2m}}; \\ v_0 = \frac{1}{a}, \quad v_m &= \cfrac{1}{a + \cfrac{1}{b + v_{m-1}}} := \frac{p_{2m+1}}{q_{2m+1}}. \end{aligned} \tag{197}$$

Then, as before $v_m \uparrow \theta_{a,b}$ and $u_m \downarrow \theta_{a,b}$. Further, as these are consecutive approximants, we have

$$\left|\frac{p_n}{q_n} - \frac{p_{n+1}}{q_{n+1}}\right| = \frac{1}{q_n q_{n+1}} \ . \tag{198}$$

As consecutive approximants are on alternate sides of their limit, we have

$$\left|\frac{p_n}{q_n} - \frac{p_{n-2}}{q_{n-2}}\right| < \left|\frac{p_n}{q_n} - \theta_{a,b}\right| < \left|\frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}}\right| \ . \tag{199}$$

It is (and shall be!) an exercise to show that there exists a constant $c_{a,b} > 0$ so that

$$q_n q_{n+1} = c_{a,b} \left(\frac{ab + 2 + \sqrt{a^2b^2 + 4ab}}{2}\right)^n (1 + o(1)). \tag{200}$$

(Hints: see equation (41) of these notes, V. You need to look at $q_{2k}q_{2k+1}$ and $q_{2k+1}q_{2k+2}$ separately. This formula is valid for the product; q_{2k} and q_{2k+1} have a different constant in the asymptotics. Note also that if $a = b$, then

$$\frac{ab + 2 + \sqrt{a^2b^2 + 4ab}}{2} = \frac{a^2 + 2 + a\sqrt{a^2 + 4}}{2} = \left(\frac{a + \sqrt{a^2 + 4}}{2}\right)^2, \tag{201}$$

so this result is consistent with Theorem 16. In any case, this is a fairly routine computation. It follows that (u_m) and (v_m) slowly approach $\theta_{a,b}$, and after a bit more work, that

$$\left|\frac{p_n}{q_n} - \theta_{a,b}\right| \sim c' \left(\frac{ab + 2 + \sqrt{a^2b^2 + 4ab}}{2}\right)^{-n}. \tag{202}$$

It is also easy to see that

$$\left|?\left(\frac{p_n}{q_n}\right) - ?(\theta_{a,b})\right| \sim 2^{-n(a+b)/2}. \tag{203}$$

Thus, $\theta_{a,b} \in D(?, \infty)$ if and only if

$$\frac{ab + 2 + \sqrt{a^2b^2 + 4ab}}{2} > 2^{(a+b)/2}. \tag{204}$$

In our earlier case with $a = b = 1$, this condition reverts to $\frac{3+\sqrt{5}}{2} > 2$. Because of the exponential growth on the right-hand side in (a, b), it is easy to believe that this inequality is satisfied for only finitely many pairs (a, b). Noting the symmetry in (a, b), a Mathematica check shows that the inequality holds precisely when $1 \le a \le b \le 4$, together with $(2, 5)$ and $(3, 5)$. This gives several more countably infinite families of elements in $D(?, \infty)$ as well as irrational elements in $D(?, 0)$.

There is an extensive literature on the differentiability of $?(x)$, and it is not our intention to cover it all. Salem, in his paper, first cites the result from metric continued fraction theory that, for $x \in (0,1)$,

$$x = [a_0(x), a_1(x), \dots], \tag{205}$$

if we let

$$A = \{x : \limsup_{n\to\infty} a_n(x) = \infty\}, \tag{206}$$

then $\mu(A) = 1$; alternatively, the set of x for which $a_n(x)$ is bounded has measure zero. He then proves that if $x \in A$ and if $?$ is differentiable at x, then $?'(x) = 0$. Since $?'$ exists a.e., it follows that $\mu(D(?,0)) = 1$.

There are two fairly recent papers by Paradis, Viader and Biblioni, preprints of which will be distributed. the main result of interest is a generalization of the foregoing. Let

$$T(x) := \limsup_{n\to\infty} \frac{a_0(x) + \cdots + a_n(x)}{n+1}. \tag{207}$$

If $?'(x)$ exists and $T(x) > k_0 \approx 5.31972...$, then $x \in D(?,0)$; if $?'(x)$ exists and $T(x) < k_1 \approx 1.38848...$, then $x \in D(?,\infty)$. Here,

$$2\log_2(1+k_0) = k_0, \qquad k_1 = 2\log_2\left(\frac{1+\sqrt{5}}{2}\right). \tag{208}$$

These conditions would resolve the behavior of $?'(\theta_{a,b})$ when $a+b \geq 11$ and $a+b \leq 2$, under the condition that we knew that $?'$ existed.

I believe, but have not been able to find yet in the literature, or prove, that there exist points x with the property that the difference quotients oscillate arbitrarily close to zero and arbitrarily large. To make such a point, we consider first consider a sequence of positive integers n_k which grows very rapidly. (I don't know yet how rapidly.) We then consider

$$u_{2k} = [0, 1^{n_1}, n_2, \dots, n_{2k}]; u_{2k+1} = [0, 1^{n_1}, n_2, \dots, n_{2k+1}]; \tag{209}$$

The intuition is that in u_{2k}, the behavior is dominated by the large final denominator, which makes the difference quotient very small, but in u_{2k+1}, the behavior is dominated by the large number of 1's in the denominator, which makes the difference quotient very large.

STERN NOTES, CHAPTER 6 (FIRST DRAFT)

BRUCE REZNICK, UIUC, MATH 595

1. The summatory function

The fact that there is such a simple formula as

$$\sum_{k=2^r}^{2^{r+1}}{}^{*}\, s(n) = 3^r \tag{1}$$

suggests that it might be interesting to see if there is an underlying measure on $[0,1]$ induced by the Stern sequence. To this end, we define the function f on $[0,1]$ by

$$f(\lambda) := \lim_{r\to\infty} \frac{1}{3^r} \cdot \sum_{k=2^r}^{\lfloor (1+\lambda)2^r \rfloor}{}^{*}\, s(n). \tag{2}$$

Except for the trivial values $f(0) = 0$, $f(\frac{1}{2}) = \frac{1}{2}$ (by symmetry) and $f(1) = 1$, there is no *a priori* reason that this limit should exist, although the pictures of the Stern sequence distributed earlier suggested a regular pattern.

We first consider dyadic rationals. Suppose $\lambda = \frac{k}{2^v}$, so that the upper limit in (2) has a simple expression for $r \geq v$:

$$f\left(\frac{k}{2^v}\right) = \lim_{r\to\infty} \frac{1}{3^r} \cdot \sum_{k=2^r}^{2^r+2^{r-v}k}{}^{*}\, s(n) = \lim_{r\to\infty} \frac{1}{3^r} \cdot \sum_{j=0}^{k-1} \left(\sum_{k=2^r+2^{r-v}j}^{2^r+2^{r-v}(j+1))}{}^{*}\, s(n) \right). \tag{3}$$

Recalling our earlier notation

$$\Sigma(f; m, r) = \sum_{n=2^r m}^{2^r(m+1)}{}^{*}\, f(n), \tag{4}$$

this becomes

$$f\left(\frac{k}{2^v}\right) = \lim_{r\to\infty} \frac{1}{3^r} \cdot \sum_{j=0}^{k-1} \Sigma(s; 2^v + j, r - v). \tag{5}$$

Further, recall that we observed that

$$
\begin{aligned}
\Sigma(s;m,1) &= \frac{s(2m)}{2} + s(2m+1) + \frac{s(2m+2)}{2} \\
&= \frac{s(m)}{2} + s(m) + s(m+1) + \frac{s(m+1)}{2} \\
&= \frac{3}{2}\left(s(m)+s(m+1)\right) = 3\Sigma(s,m,0),
\end{aligned} \tag{6}
$$

hence $\Sigma(s;m,t) = 3^t\Sigma(s;m,0)$, and the limit in (5) is that of a *constant* sequence (at least for r sufficiently large). We find that for $r \ge v$,

$$
\begin{aligned}
f\left(\frac{k}{2^v}\right) &= \frac{3^{r-v}}{3^r}\cdot\sum_{j=0}^{k-1}\Sigma(s;2^v+j,0) = \frac{1}{3^v}\cdot{\sum_{n=2^v}^{2^v+k}}^{*} s(n) \\
&= \frac{1}{3^v}\cdot\sum_{n=2^v}^{2^v+k-1}\frac{s(n)+s(n+1)}{2} = \frac{1}{2\cdot 3^v}\cdot\sum_{n=2^v}^{2^v+k-1}s(2n+1).
\end{aligned} \tag{7}
$$

It's also useful to observe that for $k \ge 1$,

$$
f\left(\frac{k}{2^v}\right) - f\left(\frac{k-1}{2^v}\right) = \frac{s(2^{v+1}+2k-1)}{2\cdot 3^v}, \tag{8}
$$

and since $f(0) = 0$, this can serve as an alternate definition of f.

To compute this function, we need to look at the odd elements of the Stern sequence. In the table below, the first row is $s(3)$, the second is $s(5), s(7)$, etc.

$$
\begin{array}{c}
2 \\
3 \quad 3 \\
4 \quad 5 \quad 5 \quad 4 \\
5 \quad 7 \quad 8 \quad 7 \quad 7 \quad 8 \quad 7 \quad 5 \\
6 \quad 9 \quad 11 \quad 10 \quad 11 \quad 13 \quad 12 \quad 9 \quad 9 \quad 12 \quad 13 \quad 11 \quad 10 \quad 11 \quad 9 \quad 6
\end{array} \tag{9}
$$

Reading from the bottom row, $f(\frac{1}{16}) = \frac{6}{162}$, $f(\frac{2}{16}) = \frac{6+9}{162} = \frac{5}{54} = f(\frac{1}{8})$, $f(\frac{3}{16}) = \frac{6+9+11}{162}$ and so on. We also can easily see that

$$
f\left(\frac{1}{2^v}\right) = \frac{v+2}{2\cdot 3^v}. \tag{10}
$$

It is worth noting here that this table of the odd Stern sequence can be generated autonomously. We have $s(2^r+1) = s(2^{r+1}-1) = r+1$,

$$
\begin{aligned}
s(4n+1)+s(4n+3) &= (2s(n)+s(n+1)) + (s(n)+2s(n+1)) \\
&= 3\left(s(n)+s(n+1)\right) = 3s(2n+1)
\end{aligned} \tag{11}
$$

and

$$
\begin{aligned}
s(4n+5)-s(4n+3) &= (s(2n+2)+s(2n+3)) - (s(2n+1)+s(2n+2)) \\
&= s(2n+3)-s(2n+1).
\end{aligned} \tag{12}
$$

Thus, if you have a row of odd terms, then you can make the next row by starting appropriately and then following these rules. To be specific,

$$(13)\qquad \begin{aligned} &s(2n+1) = a, \quad s(2n+3) = b, \quad s(4n+1) = c \\ \implies &s(4n+3) = 3a - c, \quad s(4n+5) = s(4n+3) + (b-a). \end{aligned}$$

Theorem 1. *The function f defined above extends to a continuous, strictly increasing function F from $[0,1]$ to itself, satisfying $F(1-x) = 1 - F(x)$.*

Proof. We apply Theorem 5, from the Notes V, supplement. It is certainly the case that f is defined on a dense subset X of $[0,1]$, namely the dyadic rationals, and that f is strictly increasing on its image. We need only show that the image, Y, is dense in $[0,1]$. But it follows from (8) that

$$(14)\qquad f\left(\frac{k}{2^v}\right) - f\left(\frac{k-1}{2^v}\right) \le \frac{F_{v+3}}{2\cdot 3^v} \approx \frac{\phi^3}{2\sqrt{5}}\left(\frac{\phi}{3}\right)^v,$$

which goes to 0 as $v \to \infty$. Thus, Y is indeed dense in $[0,1]$.

The mirror symmetry of the r-th row of the diatomic array implies that $f(1-x) = 1 - f(x)$ for any dyadic rational x, and this is inherited by F by continuity. □

In the notation of the last notes, $F \in \mathcal{F}$. There are now two natural questions. What is the formula for $F(\lambda)$ when λ is *not* a dyadic rational? What can be said about the differentiability of F? We can answer the first question more easily than the second.

We begin by rephrasing the definition of f. Let (r_j) be a strictly increasing sequence of positive integers and let $\lambda_0 = 0$ and

$$(15)\qquad \lambda_m = \frac{1}{2^{r_1}} + \cdots + \frac{1}{2^{r_m}}, \quad \text{for } m \ge 1,$$

so that $\lambda_m = \lambda_{m-1} + \frac{1}{2^{r_m}}$. We see from (8) that

$$(16)\qquad f(\lambda_m) - f(\lambda_{m-1}) = \frac{s(2^{r_m+1}(1+\lambda_m) - 1)}{2\cdot 3^{r_m}},$$

and since $f(0) = 0$, it follows that

$$(17)\qquad f(\lambda_m) = \sum_{j=1}^{m} \frac{s(2^{r_j+1}(1+\lambda_j) - 1)}{2\cdot 3^{r_j}}\ .$$

Since every $x \in [0,1]$ is a limit of dyadic rationals x_m given as above, we now have our definition of F:

$$(18)\qquad x_m = \sum_{j=1}^{m} \frac{1}{2^{r_j}}, \quad x = \lim_{m\to\infty} x_m \implies F(x) = \sum_{j=1}^{\infty} \frac{s(2^{r_j+1}(1+\lambda_j) - 1)}{2\cdot 3^{r_j}}\ .$$

We note that this gives a peculiar recurrence satisfied by F.

Theorem 2. *We have $F(x) - 6F(\frac{x}{2}) + 9F(\frac{x}{4}) = 0$ for $x \in [0,1]$; that is, $3^n F(\frac{x}{2^n})$ is linear in n.*

Proof. Let x be given, with r_j and λ_j as above, then for $t = 1, 2$

$$\frac{x}{2^t} = \sum_{j=1}^{\infty} \frac{1}{2^{r_j+t}}, \quad \lambda_{m,t} = \frac{\lambda_m}{2^t} \implies F(x) - 6F\left(\frac{x}{2}\right) + 9F\left(\frac{x}{4}\right) = \sum_{j=1}^{\infty} \frac{W_j}{3^{r_j}}, \tag{19}$$

where

$$\begin{aligned} W_j &= s(2^{r_j+1}(1+\lambda_j) - 1) - 2s(2^{r_j+2}(1+\lambda_j/2) - 1) + s(2^{r_j+3}(1+\lambda_j/4) - 1) \\ &= s(2^{r_j+1} + 2^{r_j+1}\lambda_j - 1) - 2s(2^{r_j+2} + 2^{r_j+1}\lambda_j - 1) + s(2^{r_j+3} + 2^{r_j+1}\lambda_j - 1). \end{aligned} \tag{20}$$

But we have already seen that $s(2^n + k)$ is linear in n, when $2^n > k$ (as is the case here), and if g is any linear function, then $g(r_j + 1) - 2g(r_j + 2) + g(r_j + 2) = 0$. Thus, $W_j = 0$ for all j, completing the proof of the identity. If $a_n = 3^n F(\frac{x}{2^n})$, then $a_{n+2} - 2a_{n+1} + a_n = 0$, so a_n is linear in n. □

There is already enough information to prove the following theorem.

Theorem 3. *If x is a dyadic rational, then $F'(x) = 0$.*

Proof. We will use the "slowly approaching" lemmas of Notes, V, supplement, pp. 17–18. Let $x = \frac{k}{2^r}$, and for $m > r$, define

$$v_m = \frac{k}{2^r} - \frac{1}{2^m}, \qquad u_m = \frac{k}{2^r} + \frac{1}{2^m}. \tag{21}$$

Then $x - v_{m+1} = \frac{1}{2}(x - v_m)$ and $u_{m+1} - x = \frac{1}{2}(u_m - x)$, so Lemma 13 applies. Now observe that

$$\begin{aligned} F(x) - F(v_m) &= F\left(\frac{k}{2^r}\right) - F\left(\frac{k}{2^r} - \frac{1}{2^m}\right) = \frac{s(2^{m+1} + k2^{m-r+1} - 1)}{2 \cdot 3^m}, \\ F(u_m) - F(x) &= F\left(\frac{k}{2^r} + \frac{1}{2^m}\right) - F\left(\frac{k}{2^r}\right) = \frac{s(2^{m+1} + k2^{m-r+1} + 1)}{2 \cdot 3^m}. \end{aligned} \tag{22}$$

But

$$\begin{aligned} s(2^{m+1} + k2^{m-r+1} \pm 1) &= s(2^{m-r+1}(2^r + k) \pm 1) = \\ & s(2^{m-r+1} - 1)s(2^r + k) + s(1)s(2^r + k \pm 1) \\ &= (m - r + 1)s(2^r + k) + s(2^r + k \pm 1), \end{aligned} \tag{23}$$

so, for appropriate constants c_j we have

$$\frac{F(x) - F(v_m)}{x - v_m} = \frac{2^m(c_1 m + c_2)}{3^m} \to 0, \quad \frac{F(u_m) - F(x)}{u_m - x} = \frac{2^m(c_1 m + c_3)}{3^m} \to 0, \tag{24}$$

completing the proof. □

We shall show later that, if $gcd(k, 3) = 1$, then $\frac{k}{3 \cdot 2^r} \in F(f, \infty)$. But in order to do this, we need a way to calculate F at non-dyadic rationals.

2. F AT NON-DYADIC RATIONALS

(Please disregard pp. 5 → 7 from the last handout; lots of typos!)

It follows from the geometric series that

$$\frac{1}{3} = \frac{1}{2^2} + \frac{1}{2^4} + \frac{1}{2^6} + \cdots , \tag{25}$$

and so from the discussion above, we have

$$F\left(\frac{1}{3}\right) = \frac{s(2^3+1)}{2\cdot 3^2} + \frac{s(2^5+2^3+1)}{2\cdot 3^4} + \frac{s(2^7+2^5+2^3+1)}{2\cdot 3^6} + \cdots . \tag{26}$$

The numerator of the r-th term is $s(w_r)$, where

$$\begin{aligned} w_r = 2^{2r+1} + 2(2^{2r-2} + \cdots + 1) - 1 &= 2^{2r+1} + 2^{2r-1} + \cdots + 2^3 + 2^1 - 1 \\ &= 2^{2r+1} + \cdots + 2^3 + 1. \end{aligned} \tag{27}$$

By a stroke of luck (I can't plan things this well!), we have already calculated this expression: $w_r = n_{2r+1} - 2$ in the notation of equations (71)–(73) on p.11 of the Notes, IV (keeping the typos in mind), and

$$s(w_r) = 3F_{2r} + F_{2r-1}. \tag{28}$$

(We have another computation of this later in the section.) It follows that

$$F\left(\frac{1}{3}\right) = \sum_{r=1}^{\infty} \frac{3F_{2r} + F_{2r-1}}{2\cdot 3^{2r}} . \tag{29}$$

In one sense, this is a routine computation, if one breaks down the Fibonacci numbers via the Binet formula. It's also interesting to see how this can be done without using "$\sqrt{5}$" at all, but entirely via the even and odd parts of generating functions:

$$\sum_{n=0}^{\infty} F_n x^n = \frac{x}{1-x-x^2} \implies \sum_{n=0}^{\infty} (-1)^n F_n x^n = \frac{-x}{1+x-x^2} . \tag{30}$$

By adding and subtracting these equations, we find that

$$\begin{aligned} \sum_{n=0}^{\infty} F_{2n} x^{2n} &= \frac{1}{2}\left(\frac{x}{1-x-x^2} - \frac{x}{1+x-x^2}\right) = \frac{x^2}{1-3x^2+x^4}; \\ \sum_{n=0}^{\infty} F_{2n+1} x^{2n+1} &= \frac{1}{2}\left(\frac{x}{1-x-x^2} + \frac{x}{1+x-x^2}\right) = \frac{x-x^3}{1-3x^2+x^4}. \end{aligned} \tag{31}$$

Since $F_0 = 0$, a mild amount of reindexing gives

$$\begin{aligned} \sum_{n=0}^{\infty} \frac{F_{2n}}{3^{2n}} &= \frac{1/9}{1-1/3+1/81} = \frac{9}{55} \implies \sum_{r=1}^{\infty} \frac{3F_{2r}}{2\cdot 3^{2r}} = \frac{27}{110}; \\ \sum_{n=0}^{\infty} \frac{F_{2n+1}}{3^{2n+1}} &= \frac{1/3-1/27}{1-1/3+1/81} = \frac{24}{55} \implies \sum_{r=1}^{\infty} \frac{F_{2r-1}}{2\cdot 3^{2r}} = \sum_{r=0}^{\infty} \frac{F_{2r+1}}{2\cdot 3^{2r+2}} = \frac{8}{110}. \end{aligned} \tag{32}$$

We conclude that

$$F\left(\frac{1}{3}\right) = \frac{27}{110} + \frac{8}{110} = \frac{35}{110} = \frac{7}{22}. \tag{33}$$

You may be forgiven for not quite believing this. As a nod towards the skeptical, $2^{16} = 65536$, $\frac{4}{3} \cdot 2^{16} = 87381\frac{1}{3}$, and a Mathematica calculation shows that

$$\frac{1}{3^{16}} \cdot \sum_{n=65536}^{87381}{}^{*} s(n) - \frac{7}{22} = \frac{27391903}{86093442} - \frac{7}{22} = -\frac{8057}{473513931}, \tag{34}$$

if I've copied right. The difference is in the fifth decimal place, which is pretty close.

What does this mean? Since $F(\frac{1}{3}) = \frac{7}{22}$, we must have $F(\frac{2}{3}) = \frac{15}{22}$, so there is slightly more Stern "stuff" in the middle than at the ends. This is also shown by $F(\frac{1}{4}) = \frac{2}{9}$ and $F(\frac{3}{4}) = \frac{7}{9}$.

We generalize, and reaffirm these numbers. Suppose $k \ge 2$. Then

$$\frac{1}{2^k - 1} = \frac{1}{2^k} + \frac{1}{2^{2k}} + \frac{1}{2^{3k}} + \cdots . \tag{35}$$

It follows that

$$F\left(\frac{1}{2^k - 1}\right) = \frac{s(2^{k+1} + 1)}{2 \cdot 3^k} + \frac{s(2^{2k+1} + 2^{k+1} + 1)}{2 \cdot 3^{2k1}} + \cdots . \tag{36}$$

We have made a very similar computation recently. If we take $m = 1$ in the "Bonus round" Theorem of HW2 solutions, and define $a_r = \frac{2^{kr}-1}{2^k-1} = 2^{k(r-1)} + \cdots + 2^k + 1$, then since $s(m+1) + s(2^k - m) = s(2) + s(2^k - 1) = k + 1$, we have that

$$s(a_0) = 0, \quad s(a_1) = 1, \qquad s(a_{r+2}) - (k+1)s(a_{r+1}) + s(a_r) = 0. \tag{37}$$

In this notation, if we let $c_j = 2^{jk+1} + \cdots + 2^{k+1} + 1 = 2a_{j+1} - 1$, then

$$s(c_j) = s(a_{j+1}) + s(a_{j+1} - 1) = s(a_{j+1}) + s(2^k a_j) = s(a_{j+1}) + s(a_j), \tag{38}$$

and so $s(c_{j+2}) - (k+1)s(c_{j+1}) + s(c_j) = 0$ as well. That is,

$$F\left(\frac{1}{2^k - 1}\right) = \sum_{j=1}^{\infty} \frac{s(c_j)}{2 \cdot 3^{jk}}, \tag{39}$$

and can now use the method of generating functions to evaluate the sum:

$$\begin{aligned} &\Phi(X) = \sum_{j=1}^{\infty} s(c_j) X^j \\ &\implies (1 - (k+1)X + X^2)\Phi(X) = s(c_1)X + (s(c_2) - (k+1)s(c_1))X^2. \end{aligned} \tag{40}$$

Now $s(c_1) = s(a_2) + s(a_1) = (k+1) + 1 = k+2$ and $s(c_2) = s(a_3) + s(a_2) = (k^2 + 2k) + (k+1) = k^2 + 3k + 1$, so

$$\Phi(X) = \frac{(k+2)X - X^2}{1 - (k+1)X + X^2} \tag{41}$$
$$\implies F\left(\frac{1}{2^k - 1}\right) = \frac{1}{2} \cdot \Phi\left(\frac{1}{3^k}\right) = \frac{(k+2)3^k - 1}{2(3^{2k} - (k+1)3^k + 1)}.$$

As a check, for $k = 2$, this becomes

$$F\left(\frac{1}{3}\right) = \frac{36-1}{2(81-27+1)} = \frac{35}{110} = \frac{7}{22}. \tag{42}$$

The next step would be to calculate $F(\frac{2}{2^k-1})$, because then the recurrence in Theorem 3 will give us $F(\frac{2^j}{2^k-1})$ for all integral $j < k$, including negative j. We don't need this for $k = 2$, because $\frac{2}{3} = 1 - \frac{1}{3}$. Let

$$a_n = 3^n F\left(\frac{1}{2^n} \cdot \frac{2}{3}\right). \tag{43}$$

Theorem 3 shows that a_n is linear in n, and since $a_0 = F(\frac{2}{3}) = \frac{15}{22}$ and $a_1 = 3F(\frac{1}{3}) = \frac{21}{22}$, we have $a_n = \frac{15+6n}{22}$. Thus,

$$F\left(\frac{1}{6}\right) = \frac{1}{9} \cdot \frac{27}{22} = \frac{3}{22}, \quad F\left(\frac{1}{12}\right) = \frac{1}{27} \cdot \frac{33}{22} = \frac{1}{18}, \quad F\left(\frac{1}{3 \cdot 2^r}\right) = \frac{7+2r}{22 \cdot 3^r}. \tag{44}$$

We now wish to show that F is not differentiable at $\lambda = \frac{m}{3 \cdot 2^v}$, provided $gcd(3, m) = 1$. In the interest of time, and interest, we shall only write out the details in one direction. Recall our notations from p.3 of these notes.

Lemma 4. *If*

$$\lim_{n\to\infty} \frac{2^{r_n} \cdot s(2^{r_n+1}(1+\lambda_n) - 1)}{3^{r_n}} = \infty, \tag{45}$$

then F is not differentiable at x.

Proof. We have the following inequalities:

$$F(\lambda) - F(\lambda_{n-1}) \geq \frac{s(2^{r_n+1}(1+\lambda_n) - 1)}{2 \cdot 3^{r_n}} \tag{46}$$
$$x - \lambda_{n-1} < \frac{1}{2^{r_n}} + \frac{1}{2^{r_n+1}} + \frac{1}{2^{r_n+2}} \cdots = \frac{2}{2^{r_n}}.$$

If (45) holds, then it follows that

$$\frac{F(\lambda) - F(\lambda_{n-1})}{\lambda - \lambda_{n-1}} > \frac{2^{r_n} \cdot s(2^{r_n+1}(1+\lambda_n) - 1)}{4 \cdot 3^{r_n}} \to \infty \tag{47}$$

as $n \to \infty$, and hence F is not differentiable at λ. □

A more careful argument can show that the difference quotient goes to ∞ as λ is approached from the left. (Note that there is no guarantee here that λ_m converges slowly to λ, because the difference between the r_m's might be unbounded if λ is irrational.)

Theorem 5. *If $\lambda = \frac{m}{3\cdot 2^v}$ and $gcd(3,m) = 1$, then F is not differentiable at λ.*

Proof. First suppose $\lambda = \frac{k}{2^v} + \frac{1}{3\cdot 2^v}$. We have

$$\lambda = \frac{k}{2^v} + \frac{1}{2^{v+2}} + \frac{1}{2^{v+4}} + \cdots, \tag{48}$$

so

$$\lambda_j = \frac{2^{2j}k + 2^{2j-2} + \cdots + 1}{2^{v+2j}}. \tag{49}$$

If

$$z_j = 2^{v+2j+1} + 2^{2j+1}k + 2^{2j-1} + \cdots + 2^3 + 1, \tag{50}$$

then

$$F(\lambda) = F\left(\frac{k}{2^v}\right) + \sum_{j=1}^{\infty} \frac{s(z_j)}{2\cdot 3^{v+2j}}. \tag{51}$$

However, $z_j = [\cdots(10)^{j-2}1001]_2$, so $z_j \sim [\cdots 1^{2j-3}21]$, and so $s(z_j) \ge F_{2j}$ – a very crude bound! Applying the lemma, we have

$$\frac{2^{r_m} s(2^{r_m+1}(1+\lambda_m)-1)}{3^{r_m}} \ge \frac{2^{v+2m}F_{2m}}{3^{v+2m}} = \left(\frac{2}{3}\right)^v \cdot \frac{2^{2m}F_{2m}}{3^{2m}}. \tag{52}$$

The numerator grows like $(1+\sqrt{5})^{2m}$, and since $1+\sqrt{5} > 3$, it follows that the quotient goes to ∞ and F is not differentiable at λ.

If $\lambda = \frac{k}{2^v} + \frac{2}{3\cdot 2^v}$, then $\lambda = 1-\lambda'$, where $\lambda' = \frac{2^v-k-1}{2^v} + \frac{1}{3\cdot 2^v}$, and F is not differentiable at λ'. But since $F(1-x) = 1-F(x)$, it follows that F is not differentiable at λ either. $\square$

We conclude this section with a more careful computation of F. Let $s(2^v+k) = a$ and $s(2^v+k+1) = b$ and recall the definition of w_r from equation (27) of these notes. We have

$$\begin{aligned} z_1 = 2^3(2^v+k)+1 &\implies s(z_1) = s(7)a + s(1)b = 3a+b, \\ z_2 = 2^5(2^v+k)+2^3+1 &\implies s(z_2) = s(23)a + s(9)b = 7a+4b, \\ z_3 = 2^7(2^v+k)+2^5+2^3+1 &\implies s(z_3) = s(87)a + s(41)b = 18a+11b. \end{aligned} \tag{53}$$

Much of this should look familiar: these are elements of the Lucas sequence (recall: $L_0 = 2, L_1 = 1$, and $L_n = L_{n-1} + L_{n-2}$ for $n \ge 2$, or $L_n = F_{n+1} + F_{n-1}$.)

There are many ways to prove this (I'm sure it follows from earlier notes), but here's a different way, using the basic recurrence. We see directly that $z_{n+1} = 4z_n + 5$, so $z_{n+2} = 16z_n + 21$ and $z_{n+3} = 64z_n + 85$. Thus,

$$\begin{aligned} s(z_{n+1}) &= 2s(z_n+1) + s(z_n+2), \\ s(z_{n+2}) &= 5s(z_n+1) + 3s(z_n+2), \\ s(z_{n+3}) &= 13s(z_n+1) + 8s(z_n+2) \\ \implies s(z_{n+3}) &- 3s(z_{n+2}) + s(z_{n+1}) = 0 \end{aligned} \tag{54}$$

Any sequence (x_n) which satisfies $x_{n+2} = x_{n+1} + x_n$ will also satisfy $x_{n+4} = 3x_{n+2} - x_n$, and so it follows from the last two equations that

$$s(z_j) = L_{2j}a + L_{2j-1}b. \tag{55}$$

The next step is to form the generating function

$$\begin{aligned} \Psi(X) &= \sum_{j=1}^{\infty} s(z_j)X^j \\ \implies (1 - 3X + X^2)\Psi(x) &= s(z_1)X + (s(z_2) - 3s(z_1))X^2 \\ \implies \Psi(X) &= \frac{(3a+b)X + (-2a+b)X^2}{1 - 3X + X^2} \\ \implies \sum_{j=1}^{\infty} \frac{s(z_j)}{3^{2j}} &= \frac{9(3a+b) + (-2a+b)}{81 - 27 + 1} = \frac{5a+2b}{11}. \end{aligned} \tag{56}$$

Theorem 6. *Suppose $a = s(2^v + k)$ and $b = s(2^v + k + 1)$. Then*

$$\begin{aligned} F\left(\frac{k}{2^v} + \frac{1}{3\cdot 2^v}\right) - F\left(\frac{k}{2^v}\right) &= \frac{5a+2b}{22\cdot 3^v}; \\ F\left(\frac{k}{2^v} + \frac{2}{3\cdot 2^v}\right) - F\left(\frac{k}{2^v} + \frac{1}{3\cdot 2^v}\right) &= \frac{4a+4b}{22\cdot 3^v}; \\ F\left(\frac{k+1}{2^v}\right) - F\left(\frac{k}{2^v} + \frac{2}{3\cdot 2^v}\right) &= \frac{2a+5b}{22\cdot 3^v}. \end{aligned} \tag{57}$$

Proof. The first equation follows from (56), applied to (50). The third follows from symmetry by considering $F(1-x) = 1 - F(x)$. Finally, the second follows from subtracting the sum of the first and the last from (8), when rewritten in the way it will appear in the second draft:

$$F\left(\frac{k+1}{2^v}\right) - F\left(\frac{k}{2^v}\right) = \frac{a+b}{2\cdot 3^v}. \tag{58}$$

□

It seems surely possible to continue this work for any odd division of a dyadic interval, but we'll stop here.

We close this section with some numbers. We have already seen that when $[0,1]$ is broken up into $\frac{1}{8}$-ths, the fraction of the Stern mass is divided in ratio $5:7:8:7:7:8:7:5$. Using Theorem 6, we can compute the mass when $[0,1]$ is broken up into $\frac{1}{24}$-ths. By symmetry, we only pressent the first 12: is divided in ratio

$$13:20:22:26:28:23:25:32:31:29:28:20. \tag{59}$$

The denominator here is $11 \cdot 54 = 594$. Thus, $F(\frac{1}{24}) = \frac{13}{594}$, $F(\frac{2}{24}) = F(\frac{1}{12}) = \frac{33}{594} = \frac{1}{18}$, as we have seen, etc. And in the largest six 24-ths of the interval – $x \in [\frac{7}{24}, \frac{10}{24}] \cup [\frac{14}{24}, \frac{17}{24}]$, we find $\frac{2(32+31+29)}{594} \approx 31\%$ of the total mass. In other words, it's pretty well distributed, at least on this level of granularity.

3. More to do

We bid a farewell to F with these notes, but there is more work to do.

• The closed form for F is unsatisfactory in some ways, and it would be good to find a version in which the infinite sum did not involve unevaluated elements of the Stern sequence.

• It is almost certainly true that if $x \in \mathbb{Q}$, then $F(x) \in \mathbb{Q}$. This would follow from the eventual periodicity of the binary expansion of x, when combined with the (unproved) lemma that if (y_n) is a sequence given by the recurrence $y_{n+1} = 2^r y_n + a$, with $0 \le a < 2^r$, then $s(y_n)$ satsifies a second order recurrence. This can be proved using a version of the bonus round theorem, which will undoubtedly have a more dignified name in the second draft.

• I'll put the computation of $F(\frac{1}{5})$ on the homework. Once this is done, it is also possible, if tedious, to write a version of Theorem 6 for 5-ths.

• It is totally natural to define the measure μ_F on $[0,1]$ by "differentiating' F:

$$\mu([0,x]) := F(x) = \int_0^x d\mu_F \tag{60}$$

It is also totally natural to make a pun out of a cliche and call this the Stern measure. I have not found (yet) anything really interesting which will let me say that the "Stern measure must be applied." The most natural questions about this measure would be to find its moments; that is

$$\int_0^1 t^k \; d\mu_F. \tag{61}$$

• A picture of F was given in the first day's handout. Some questions which come to mind there are these: what is the maximum of $f(x) - x$? where does it occur? what can be said about the places where $f(x) = x$, besides $x = 0, \frac{1}{2}, 1$?

4. $F(\frac{1}{5})$ and more

We really ought to work out $F(\frac{1}{5})$. First observe that

$$\frac{1}{5} = \frac{3}{15} = \frac{2^1 + 2^0}{2^4 - 1} = \frac{1}{2^3} + \frac{1}{2^4} + \frac{1}{2^7} + \frac{1}{2^8} + \frac{1}{2^{11}} + \frac{1}{2^{12}} + \cdots . \tag{62}$$

We define the sequence of consecutive dyadic approximations to $\frac{1}{5}$:

$$\lambda_0 = 0, \quad \lambda_1 = \frac{1}{2^3}, \quad \lambda_2 = \lambda_1 + \frac{1}{2^4}, \quad \lambda_3 = \lambda_2 + \frac{1}{2^7}, \quad \lambda_4 = \lambda_3 + \frac{1}{2^8} \dots; \tag{63}$$

that is,

$$\lambda_{2m} = \lambda_{2m-1} + \frac{1}{2^{4m}}, \qquad \lambda_{2m+1} = \lambda_{2m} + \frac{1}{2^{4m+3}}. \tag{64}$$

As we've seen before

$$\begin{aligned} F\left(\frac{1}{5}\right) &= \sum_{k=1}^{\infty} F(\lambda_k) - F(\lambda_{k-1}) \\ &= \sum_{m=0}^{\infty} \frac{s(2^{4m+4}(1+\lambda_{2m+1}) - 1)}{2 \cdot 3^{4m+3}} + \sum_{m=1}^{\infty} \frac{s(2^{4m+1}(1+\lambda_{2m}) - 1)}{2 \cdot 3^{4m}}. \end{aligned} \tag{65}$$

Let

$$\alpha_m = 2^{4m+1}(1+\lambda_{2m}) - 1 = 2y_m - 1, \qquad \beta_m = 2^{4m+4}(1+\lambda_{2m+1}) - 1 = 2z_m - 1. \tag{66}$$

Then

$$\begin{aligned} y_m &= 2^{4m}(1+\lambda_{2m}) = 2^{4m}\left(1 + \lambda_{2m-1} + \frac{1}{2^{4m}}\right) = 2z_{m-1} + 1 \\ z_m &= 2^{4m+3}(1+\lambda_{2m+1}) = 2^{4m+3}\left(1 + \lambda_{2m} + \frac{1}{2^{4m+3}}\right) = 8y_m + 1. \end{aligned} \tag{67}$$

Thus, $y_{m+1} = 2z_m + 1 = 16y_m + 3$ (so $y_{m+2} = 256y_m + 51$), and $z_{m+1} = 8y_{m+1} + 1 = 16z_m + 9$ (so $z_{m+2} = 256z_m + 153$). We have, in a disturbingly familiar calculation:

$$\begin{aligned} s(y_{m+1}) &= 5s(y_m) + 2s(y_m+1), \qquad & s(y_{m+2}) &= 29s(y_m) + 12s(y_m+1), \\ s(z_{m+1}) &= 3s(z_m) + 4s(z_m+1), \qquad & s(z_{m+2}) &= 17s(z_m) + 24s(z_m+1). \end{aligned} \tag{68}$$

It follows that the sequences $\{s(y_m)\}$ and $\{s(z_m)\}$ each satisfy the recurrence

$$x_{m+2} - 6x_{m+1} + x_m = 0. \tag{69}$$

Observe as well that

$$\begin{aligned} s(\alpha_m) &= s(2y_m - 1) = s(y_m) + s(y_m - 1) = s(y_m) + s(2z_{m-1}), \\ s(\beta_m) &= s(2z_m - 1) = s(z_m) + s(z_m - 1) = s(z_m) + s(8y_m). \end{aligned} \tag{70}$$

Returning to the task at hand, we have

$$F\left(\frac{1}{5}\right) = \sum_{m=0}^{\infty} \frac{s(y_m)+s(z_m)}{2\cdot 3^{4m+3}} + \sum_{m=1}^{\infty} \frac{s(y_m)+s(z_{m-1})}{2\cdot 3^{4m}}. \tag{71}$$

Finally, let

$$\begin{aligned}\Phi(X) &= \sum_{m=0}^{\infty}(s(y_m)+s(z_m))X^m = 5+31X+181X^2+\dots\\ \Psi(X) &= \sum_{m=1}^{\infty}(s(y_m)+s(z_{m-1})X^m = 11X+65X^2+379X^3+\dots.\end{aligned} \tag{72}$$

The coefficients of Φ and Ψ each satisfy the recurrence (69), and so when they are multiplied by $1-6X+X^2$ yield a polynomial:

$$\Phi(X) = \frac{5-X}{1-6X+X^2}, \qquad \Psi(X) = \frac{11X-X^2}{1-6X+X^2}. \tag{73}$$

A final computation gives

$$F\left(\frac{1}{5}\right) = \frac{\Phi(3^{-4})}{54} + \frac{\Psi(3^{-4})}{2} = \frac{87}{868} + \frac{445}{6076} = \frac{17}{98}. \tag{74}$$

I can assert with some confidence that someone some day will find a more direct way of computing $F(\frac{1}{5})$ without needing denominators as large as 6076. Perhaps this would follow from

$$\frac{1}{5} = \frac{1}{2^2} - \frac{1}{2^4} + \frac{1}{2^6} - + \dots. \tag{75}$$

In any event, some skepticism is appropriate, and we again check numerically: $\frac{6}{5}\cdot 2^{16} = 78643.2$, and another Mathematica calculation shows that

$$\frac{1}{3^{16}} \cdot \sum_{n=65536}^{78643}{}^{*} s(n) - \frac{17}{98} = \frac{7467080}{43046721} - \frac{17}{98} = -\frac{20417}{4218578658}, \tag{76}$$

if I've copied right. The difference again is in the fifth decimal place, which is pretty close, so this is probably right.

We have now by symmetry that $F(\frac{4}{5}) = 1 - F(\frac{1}{5}) = \frac{81}{98}$. Theorem 2 says that $F(\frac{4}{5}), 3F(\frac{2}{5})$ and $9F(\frac{1}{5})$ are in arithmetic progression, hence $3F(\frac{2}{5}) = \frac{117}{98}$ and so $F(\frac{2}{5}) = \frac{39}{98}$ and $F(\frac{3}{5}) = \frac{59}{98}$. Thus the Stern division by fifths is

$$17 : 22 : 20 : 22 : 17, \tag{77}$$

which is both regular and irregular enough to hint at some deeper structures.

More generally, and in a nod towards the second draft of these notes, let

$$x = \sum_{k=1}^{\infty} \frac{1}{2^{r_k}}, \quad \lambda_n = \sum_{k=1}^{n} \frac{1}{2^{r_k}}, \qquad \delta_1 = r_1, \quad \delta_k = r_k - r_{k-1}\ (k \ge 2). \tag{78}$$

Then as usual,

$$F(x)=\sum_{n=1}^{\infty}\frac{s(2^{r_n+1}(1+\lambda_n)-1)}{2\cdot 3^{r_n}}. \tag{79}$$

Following the pattern for $\frac{1}{5}$, for $n \geq 1$, let

$$2y_n - 1 = 2^{r_n+1}(1+\lambda_n)-1. \tag{80}$$

It then follows that for $n \geq 2$,

$$y_n = 2^{r_n}(1+\lambda_n) = 2^{\delta_n}2^{r_{n-1}}(1+\lambda_{n-1}+2^{-r_n}) = 2^{\delta_n}y_{n-1}+1. \tag{81}$$

Since $y_1 = 2^{r_1}+1$, this recurrence is also valid if we set $y_0 = 1$. Now, we have

$$s(2y_n-1) = s(y_n)+s(y_n-1) = s(y_n)+s(2^{\delta_n}y_{n-1}) = s(y_n)+s(y_{n-1}), \tag{82}$$

and the expression for F simplifies:

$$F(x)=\sum_{n=1}^{\infty}\frac{s(y_{n-1})+s(y_n))}{2\cdot 3^{r_n}}. \tag{83}$$

We have $s(y_0) = s(1) = 1$ and $s(y_1) = s(2^{r_1}+1) = \delta_1 + 1$. In general, if $x_1 = 2^a x_0 + 1$ and $x_2 = 2^b x_1 + 1 = 2^{a+b}x_0 + 2^b + 1$, then since

$$s(2^{a+b}-2^b-1) = s(2^b(2^a-1)-1) = s(2^b-1)s(2^a-1)+s(1)s(2^a-2) = ab+a-1, \tag{84}$$

we have

$$\begin{aligned} s(x_1) = as(x_0)+s(x_0+1), \quad s(x_2) &= (ab+a-1)s(x_0)+(b+1)s(x_0+1) \\ \implies s(x_2) &= (b+1)s(x_1)-s(x_0). \end{aligned} \tag{85}$$

It follows that the sequence $s(y_m)$ satisfies the recurrence

$$s(y_m) = \delta_m s(y_{m-1}) - s(y_{m-2}), \qquad \text{for } m \geq 2. \tag{86}$$

If x is irrational, then the dyadic expression for x is not repeating, but if $x \in \mathbb{Q}$, then eventually the δ_m's are periodic, with period p say. It follows that each sequence $s(y_{pt+j})$ will satisfy a second-order recurrence of some kind and the method shown above will allow a computation of $F(x)$ via a number of generating functions. Also,

$$\frac{s(y_m)}{s(y_{m-1})} = \delta_m - \frac{1}{\frac{s(y_{m-1})}{s(y_{m-2})}} = \dots \tag{87}$$

Every positive integer can be written in this way:

$$n = 2^{r_m}+\dots+2^{r_1}+1 = 2^{r_1}(2^{r_m-r_1}+\dots+2^{r_2-r_1})+1, \dots \tag{88}$$

and so this gives a less canonical form for $s(n)$ as hte numerator of a (non-simple) continued fraction. If n has this form, then

$$n \sim [1, r_m - r_{m-1}-1, 1, \dots, r_2 - r_1 - 1, 1, r_1 - 1, 1]_2 \tag{89}$$

If $r_k = r_{k-1}+1$, then the 1's are consecutive. However, as we have observed, the continuant is smart enough to notice: see Notes, IV, Theorem 6.

Congruence properties of binary partition functions

Katherine Anders, Melissa Dennison, Jennifer Weber Lansing and Bruce Reznick

Abstract. Let $\mathcal{A}$ be a finite subset of $\mathbb{N}$ containing 0, and let $f(n)$ denote the number of ways to write n in the form $\sum \epsilon_j 2^j$, where $\epsilon_j \in \mathcal{A}$. We show that there exists a computable $T = T(\mathcal{A})$ so that the sequence $(f(n) \bmod 2)$ is periodic with period T. Variations and generalizations of this problem are also discussed.

Mathematics Subject Classification (2010). Primary:11A63,11P81, Secondary: 11B34, 11B50.

Keywords. partitions, digital representations, Stern sequence.

1. Introduction

Let $\mathcal{A} = \{0 = a_0 < a_1 < \cdots\}$ denote a finite or infinite subset of $\mathbb{N}$ containing 0, and fix an integer $b \geq 2$. Let $f_{\mathcal{A},b}(n)$ denote the number of ways to write n in the form

$$n = \sum_{k=0}^{\infty} \epsilon_k b^k, \qquad \epsilon_k \in \mathcal{A}. \tag{1.1}$$

The uniqueness of the standard base-b representation of $n \geq 0$ reflects the fact that $f_{\mathcal{A},b}(n) = 1$ for $\mathcal{A} = \{0, \dots, b-1\}$. For non-standard bases, the behavior of $f_{\mathcal{A},b}(n)$ has been studied primarily when $\mathcal{A} = \mathbb{N}$ or $b = 2$, in terms of congruences at special values, and also asymptotically. In this paper, we are concerned with the behavior of $f_{\mathcal{A},b}(n) \pmod{d}$, especially when $b = d = 2$, and when $\mathcal{A}$ is finite.

We associate to $\mathcal{A}$ its characteristic function $\chi_{\mathcal{A}}(n)$, and the generating function

$$\phi_{\mathcal{A}}(x) := \sum_{n=0}^{\infty} \chi_{\mathcal{A}}(n) x^n = \sum_{a \in \mathcal{A}} x^a = 1 + x^{a_1} + \cdots . \tag{1.2}$$

Anders and Weber Lansing received support from National Science Foundation grant DMS 0838434 EMSW21MCTP: Research Experience for Graduate Students.

Let

$$F_{\mathcal{A},b}(x) := \sum_{n=0}^{\infty} f_{\mathcal{A},b}(n)x^n \tag{1.3}$$

denote the generating function of $f_{\mathcal{A},b}(n)$. Viewing (1.1) as a partition problem, we find an immediate infinite product representation for $F_{\mathcal{A},b}(x)$:

$$F_{\mathcal{A},b}(x) = \prod_{k=0}^{\infty}\left(1 + x^{a_1 b^k} + \cdots\right) = \prod_{k=0}^{\infty} \phi_{\mathcal{A}}(x^{b^k}). \tag{1.4}$$

Observe that (1.1) implies that $n \equiv \epsilon_0 \pmod b$. Thus, every such representation may be rewritten as

$$n = \sum_{j=0}^{\infty} \epsilon_j b^j = \epsilon_0 + b\left(\sum_{j=0}^{\infty} \epsilon_{j+1} b^j\right). \tag{1.5}$$

Since $f_{\mathcal{A},b}(n) = 0$ for $n < 0$, we see that (1.5) gives the recurrence

$$f_{\mathcal{A},b}(n) = \sum_{\substack{a\in\mathcal{A},\\ n\equiv a \bmod b}} f_{\mathcal{A},b}\left(\tfrac{n-a}{b}\right), \quad \text{for} \quad n \ge 1. \tag{1.6}$$

Alternatively, decompose $\mathcal{A}$ into residue classes mod b and write

$$\mathcal{A} = \bigcup_{i=0}^{b-1} \mathcal{A}_i, \quad \text{where} \quad \mathcal{A}_i := \mathcal{A} \cap (b\mathbb{Z} + i). \tag{1.7}$$

If we write $\mathcal{A}_i = \{bv_{k,i} + i\}$, then for $m \ge 0$ and $0 \le i \le b-1$:

$$f_{\mathcal{A},b}(bm + i) = \sum_k f_{\mathcal{A},b}(m - v_{k,i}). \tag{1.8}$$

The initial condition $f_{\mathcal{A},b}(0) = 1$, combined with (1.6) or (1.8), is sufficient to determine $f_{\mathcal{A},b}(n)$ for all $n > 0$.

We say that a sequence (u_n) is *ultimately periodic* if there exist integers $N \ge 0$, $T \ge 1$ so that, for $n \ge N$, $u_{n+T} = u_n$. The *period* of an ultimately periodic sequence is the smallest such T. By extension, we say that the set $\mathcal{A}$ is *ultimately periodic* if the sequence of its characteristic function, $(\chi_{\mathcal{A}}(n))$, is ultimately periodic. Equivalently, $\mathcal{A}$ is ultimately periodic if there exists T, and $k \ge 1$ integers $r_1, \ldots, r_k$, $0 \le r_i \le T-1$, so that the symmetric set difference of $\mathcal{A}$ and $\cup(T\mathbb{N} + r_i)$ is finite. In particular, if $\mathcal{A}$ is finite or the complement of a finite set, then $\mathcal{A}$ is ultimately periodic.

Theorem 1.1. *As elements of* $\mathbb{F}_2[[x]]$,

$$F_{\mathcal{A},2}(x)\phi_{\mathcal{A}}(x) = 1. \tag{1.9}$$

This theorem also appears as [5, Lemma 2.2(ii)], although the implications we discuss here for digital representations are not pursued there in detail. Theorem 1.1 has an immediate corollary.

Corollary 1.2.

1. *If* $\mathcal{A}$ *is finite, then there is a computable integer* $T = T(\mathcal{A}) > 0$ *so that for all* $n \geq 0$*,* $f_{\mathcal{A},2}(n) \equiv f_{\mathcal{A},2}(n+T) \pmod 2$*.*
2. *If* $\mathcal{A}$ *is infinite, then the following are equivalent:*
   (i) *The sequence* $(f_{\mathcal{A},2}(n) \pmod 2)$ *is ultimately periodic.*
   (ii) $\phi_{\mathcal{A}}(x)$ *is the power series of a rational function in* $\mathbb{F}_2(x)$*.*
   (iii) *The set* $\mathcal{A}$ *is ultimately periodic.*

It will follow from Corollary 1.2(1) that if $\mathcal{A}$ is a finite set, and $T = T(\mathcal{A})$, then there is a *complementary* finite set $\mathcal{A}' = \{0 = b_0 < b_1 < \cdots\}$ so that

$$\begin{aligned} f_{\mathcal{A},2}(n) \quad &\text{is odd} \iff n \equiv b_k \pmod T \quad \text{for some } b_k; \\ f_{\mathcal{A}',2}(n) \quad &\text{is odd} \iff n \equiv a_k \pmod T \quad \text{for some } a_k. \end{aligned} \tag{1.10}$$

Complementary sets needn't look very much alike. If $\mathcal{A} = \{0, 1, 4, 9\}$, then $T = 84$ and $|\mathcal{A}'| = 41$, with elements ranging from 0 to 75 (see Example 4.3).

One instance of Theorem 1.1 in the literature comes from the *Stern sequence* $(s(n))$ (see [13, 8, 11]), which is defined by

$$\begin{gathered} s(0) = 0, \; s(1) = 1; \\ s(2n) = s(n), \; s(2n+1) = s(n) + s(n+1) \quad \text{for } n \geq 1. \end{gathered} \tag{1.11}$$

It was proved in [10] that $s(n) = f_{\{0,1,2\},2}(n-1)$, under which the recurrence (1.11) is a translation of (1.8). It is easy to prove, and has basically been known since [13, p.197], that $s(n)$ is even if and only if n is a multiple of three. A simple application of Theorem 1.1 shows that in $\mathbb{F}_2(x)$,

$$F_{\{0,1,2\},2}(x) = \frac{1}{1+x+x^2} = \frac{1+x}{1+x^3} = 1+x+x^3+x^4+x^6+x^7+\cdots. \tag{1.12}$$

This result was generalized in [10, Th.2.14], using the infinite product (1.4). Here, let $\mathcal{A}_d = \{0, \ldots, d-1\}$. Then $\phi_{\mathcal{A}_d}(x) = \frac{1-x^d}{1-x}$, so in $\mathbb{F}_2(x)$,

$$F_{\mathcal{A}_d,2}(x) = \frac{1+x}{1+x^d} = 1 + x + x^d + x^{d+1} + x^{2d} + x^{2d+1} + \cdots. \tag{1.13}$$

Thus, $f_{\mathcal{A}_d,2}(n)$ is odd if and only if $n \equiv 0, 1 \pmod d$.

We also show that there is no obvious "universal" generalization of Theorem 1.1 to $f_{\mathcal{A},b}(n) \pmod d$, except for the case $b = d = 2$.

Theorem 1.3.

1. *If* $(f_{\{0,1,2\},2}(n) \pmod d)$ *is ultimately periodic with period* T*, then* $d = 2$ *and* $T = 3$*.*
2. *If* $d \geq 2$ *and* $b \geq 3$*, then* $(f_{\{0,1\},b}(n) \pmod d)$ *is never ultimately periodic.*

Thus, the Stern sequence has no periodicities mod $d \geq 3$ and, there exists a set $\mathcal{A}$ with the property that the number of its representations in any base $b \geq 3$ is never ultimately periodic modulo any $d \geq 2$.

Let $\nu_2(m)$ denote the largest power of 2 dividing m. In 1969, Churchhouse [4] conjectured, based on numerical evidence, that $f_{\mathbb{N},2}(n)$ is even for $n \geq 2$, that $4 \mid f_{\mathbb{N},2}(n)$ if and only if either $\nu_2(n-1)$ or $\nu_2(n)$ is a positive even integer, and that 8 never divides $f_{\mathbb{N},2}(n)$. He also conjectured that, for all even m,

$$\nu_2(f_{\mathbb{N},2}(4m)) - \nu_2(f_{\mathbb{N},2}(m)) = \lfloor \tfrac{3}{2}(3\nu_2(m)+4) \rfloor. \tag{1.14}$$

This conjecture was proved in the next few years by Rødseth, and by Gupta and generalized by Hirschhorn and Loxton, Rødseth, Gupta, Andrews, Gupta and Pleasants, and most recently by Rødseth and Sellers [12]. We refer the reader to [10, 12] for detailed references. The statements in Theorem 1.3 about the non-existence of recurrences do not apply to formulas such as (1.14). On the other hand, $\phi_{\mathbb{N}}(x) = (1+x)^{-1}$, so Theorem 1.1 implies that $f_{\mathbb{N},2}(n)$ is even for $n \ge 2$.

The paper is organized as follows. In section two, we review some familiar facts about polynomials and rational functions over $\mathbb{F}_2$. Most of this material can be found in [9], and is included here for the sake of completeness. In section three, we give two proofs of Theorem 1.1 and then prove Corollary 1.2 and Theorem 1.3. In section four, we present several examples and applications of Theorem 1.1.

Portions of the research in this paper were contained in Dennison's UIUC Ph.D. dissertation [6] and in the UIUC Summer 2010 Research Experiences for Graduate Students (REGS) project [1] of Anders and Weber Lansing. These projects were written under Reznick's supervision.

The authors thank Bob McEliece and Kevin O'Bryant for helpful correspondence, and the referee for an insightful report.

2. Background

There is an important relationship between rational functions in $\mathbb{F}_2[[x]]$ and ultimately periodic sequences. (For additional information about most of the material in this section, see [9], especially Chapter 8, "Linear Recurring Sequences".) We first recall some familiar facts about finite fields, identifying $\mathbb{Z}/p\mathbb{Z}$ with $\mathbb{F}_p$ for prime p. The binomial theorem implies that for $a, b \in \mathbb{F}_p$, $(a+b)^p = a^p + b^p$, hence $(\sum a_i)^p = \sum a_i^p$. It follows from this fact and Fermat's Little Theorem that for any polynomial $f \in \mathbb{F}_p[x]$,

$$f(x)^p = f(x^p). \tag{2.1}$$

If $f \in \mathbb{F}_2[x]$ is an irreducible polynomial of degree $d \ge 2$ (so $f(0) \ne 0$), then it is well-known that $f(x) \mid 1 + x^{2^d-1}$. Repeated application of (2.1) for $p = 2$ shows that $(1+x^M)^{2^k} = 1 + x^{2^k \cdot M}$, hence if $f \in \mathbb{F}_2[x]$ is irreducible and $j \le 2^k$, then $f(x)^j \mid 1 + x^{2^k \cdot (2^d-1)}$. This leads immediately to the following lemma (see [2, Thm.6.21]):

Lemma 2.1. *Suppose $h \in \mathbb{F}_2[x]$, $h(0) \ne 0$ and h can be factored over $\mathbb{F}_2[x]$ as*

$$h = \prod_{i=1}^{s} f_i^{e_i}, \tag{2.2}$$

where the f_i are distinct irreducible polynomials with $\deg(f_i) = d_i$*, and suppose* $2^k \geq e_i$ *for all* i *and some* $k \in \mathbb{N}$*. Then*

$$h(x) \mid 1 + x^M, \text{ where } \quad M := M(h) = 2^k \cdot lcm(2^{d_1} - 1, \dots, 2^{d_s} - 1). \quad (2.3)$$

Suppose $h \in \mathbb{F}_2[x]$ and $h(0) = 1$. The *period* of h is the smallest $T \geq 1$ so that $h(x) \mid 1 + x^T$; this definition does not assume that h is irreducible. The period of h can be much smaller than $M(h)$, however it is always a divisor of $M(h)$.

Lemma 2.2. *If* h *has period* T*, then* $h(x) \mid 1 + x^V$ *in* $\mathbb{F}_2[x]$ *if and only if* $T \mid V$*.*

Proof. We first note that $(1 + x^T) \mid (1 + x^{kT})$, proving one direction. For the other, suppose $h(x) \mid 1 + x^V$; then $V \geq T$. Write $V = kT + r$, where $0 \leq r \leq T - 1$. Then $h(x)$ also divides

$$x^r(1 + x^{kT}) + 1 + x^V = 1 + x^r, \quad (2.4)$$

which violates the minimality of T unless $r = 0$. □

If $h \in \mathbb{F}_2[x]$ is irreducible, $\deg h = r$ and the period of h is $2^r - 1$, then h is called *primitive*; see e.g.[9, §3.15]. Primitive trinomials have attracted much recent interest, especially when $2^r - 1$ is a Mersenne prime (see [3]); Lemma 2.1 implies that all such irreducible h are primitive. In coding theory, h is called the *generator* polynomial and

$$q(x) = \frac{1 + x^T}{h(x)} \quad (2.5)$$

is called the *parity-check* polynomial.

Consider a rational function in $\mathbb{F}_2(x)$:

$$\frac{g(x)}{h(x)} = a(x) + \frac{r(x)}{h(x)}, \quad (2.6)$$

where g, h, a, r are polynomials, and $\deg r < \deg h$. We make the additional assumption that $h(0) \neq 0$. Lemma 2.1 leads to an important relationship between rational functions and ultimately periodicity.

Lemma 2.3. *Suppose* $b(x) = \sum b_n x^n \in \mathbb{F}_2[[x]]$ *with* $b_0 = 1$*. Then* $b(x)$ *is a rational function if and only if* $\{n : b_n = 1\}$ *is ultimately periodic.*

Proof. First suppose there exists T, N so that $b_n = b_{n+T}$ for $n \geq N$. Then the coefficient of x^{n+T} in

$$(1 + x^T)\left(\sum_{n=0}^{\infty} b_n x^n\right) \quad (2.7)$$

is $b_{n+T} + b_n = 0$ for $n \geq N$. Hence, $b(x)$ is the quotient of a polynomial of degree less than N and $1 + x^T$, and is thereby a rational function. Conversely, suppose $b = g/h$ is rational and is given by (2.6) with $h(0) = 1$. Then by

Lemma 2.1 and the division algorithm, there exists $q(x) \in \mathbb{F}_2[x]$ and T so that

$$b(x) = a(x) + \frac{r(x)}{h(x)} = a(x) + \frac{r(x)q(x)}{1 + x^T}, \tag{2.8}$$

hence $(1+x^T)b(x)$ is a polynomial of degree less than M (say), so $b_n = b_{n+T}$ for $n \geq M$. □

3. Proofs

We start this section with two proofs of Theorem 1.1. The first one is somewhat longer, but yields a recurrence of independent interest.

As in (1.7), write

$$\begin{aligned} \mathcal{A} = \{0 = a_0 < a_1 < \cdots\} = \mathcal{A}_0 \cup \mathcal{A}_1; \\ \mathcal{A}_0 = \{0 = 2b_0 < 2b_1 < \cdots\}, \qquad \mathcal{A}_1 = \{2c_1 + 1 < \cdots\}. \end{aligned} \tag{3.1}$$

We will write $f_{\mathcal{A},2}(n)$ as $f(n)$ when there is no ambiguity. By (1.8), we have:

$$f(2n) = \sum_i f(n - b_i), \qquad f(2n+1) = \sum_j f(n - c_j). \tag{3.2}$$

Theorem 3.1. *For all $n \in \mathbb{Z}$, $n \neq 0$,*

$$\Theta(n) := \sum_k f(n - a_k) \equiv 0 \pmod 2. \tag{3.3}$$

Proof. If $n < 0$, then $f(n) = 0$, so this is immediate; also $\Theta(0) = f(0) = 1$. Suppose $n > 0$. We distinguish two cases: $n = 2m$ and $n = 2m + 1$, and put (3.2) back into itself. We then diagonalize the double sums below; for each fixed m, these sums are finite:

$$\begin{aligned} \Theta(2m) &= \sum_k f(2m - a_k) = \sum_i f(2m - 2b_i) + \sum_j f(2m - 2c_j - 1) \\ &= \sum_i \sum_u f(m - b_i - b_u) + \sum_j \sum_v f(m - c_j - 1 - c_v) \\ &= \sum_i f(m - 2b_i) + 2 \sum_{i<u} f(m - b_i - b_u) \\ &+ \sum_j f(m - 2c_j - 1) + 2 \sum_{j<v} f(m - c_j - c_v - 1) \\ &\equiv \Theta(m) \pmod 2. \end{aligned} \tag{3.4}$$

Similarly,

$$\begin{aligned}\Theta(2m+1) &= \sum_k f(2m+1-a_k) \\ &= \sum_i f(2m+1-2b_i) + \sum_j f(2m-2c_j) \\ &= \sum_i \sum_j f(m-b_i-c_j) + \sum_j \sum_i f(m-c_j-b_i) \\ &= 2\sum_{i,j} f(m-b_i-c_j) \equiv 0 \pmod 2.\end{aligned} \tag{3.5}$$

Since $\Theta(2m) \equiv \Theta(m)$ and $\Theta(2m+1) \equiv 0$, it follows by induction that $\Theta(m) \equiv 0$ for $m \geq 1$. □

We give two proofs of Theorem 1.1. The first uses Theorem 3.1; the second uses the generating function (1.3) and is also [5, Lemma 2.1].

First proof of Theorem 1.1. Write out the product in (1.9) and use Theorem 3.1.

$$F_{\mathcal{A},2}(x)\phi_{\mathcal{A}}(x) = \left(\sum_{n=0}^{\infty} f(n)x^n\right)\left(1+\sum_{i\geq 1} x^{a_i}\right) = \sum_{n=0}^{\infty} \Theta(n)x^n \equiv 1. \tag{3.6}$$

□

Second proof of Theorem 1.1. By repeated use of (1.4) and (2.1),

$$\phi_{\mathcal{A}}(x)F_{\mathcal{A},2}^2(x) \equiv \phi_{\mathcal{A}}(x)F_{\mathcal{A},2}(x^2) = \phi_{\mathcal{A}}(x)\prod_{k=0}^{\infty}\phi_{\mathcal{A}}\left(x^{2^{k+1}}\right) = F_{\mathcal{A},2}(x). \tag{3.7}$$

□

The second proof generalizes to primes $p > 2$ via (2.1).

Theorem 3.2. *If $b = p$ is prime, then $F_{\mathcal{A},p}^{p-1}(x)\phi_{\mathcal{A}}(x) = 1 \in \mathbb{F}_p[x]$.*

Proof. As before, we have

$$\phi_{\mathcal{A}}(x)F_{\mathcal{A},p}^p(x) = \phi_{\mathcal{A}}(x)F_{\mathcal{A},p}(x^p) = \phi_{\mathcal{A}}(x)\prod_{k=0}^{\infty}\phi_{\mathcal{A}}(x^{p^{k+1}}) = F_{\mathcal{A},p}(x). \tag{3.8}$$

□

This result may fail if b is not prime. For example, if $\mathcal{A} = \{0,1\}$ and $b = 4$, then $\phi_A(x) = 1+x$ and the coefficient of x^2 in $F_{\mathcal{A},4}^3(x)\phi_{\mathcal{A}}(x)$ is $6 \not\equiv 0$ mod 4. Theorem 3.2 implies that $F_{\mathcal{A},p}(x) = \phi_{\mathcal{A}}^{-1/(p-1)}(x)$ as an element of $\mathbb{F}_p[[x]]$.

Proof of Corollary 1.2(1). Suppose $\mathcal{A}$ is finite and T is the period of $\phi_{\mathcal{A}}(x)$. Then by Theorem 1.1, we have in $\mathbb{F}_2[x]$

$$F_{\mathcal{A},2}(x) = \frac{1}{\phi_{\mathcal{A}}(x)} = \frac{q(x)}{1+x^T}, \tag{3.9}$$

where $(1+x^T)F_{\mathcal{A},2}(x) = q(x) = 1+\sum x^{b_k}$ and $\deg q < T$. Since the coefficient of x^{n+T} in q is $f(n+T) - f(n) = 0$, $(f(n) \pmod 2)$ is periodic with period T. □

Let $\mathcal{A}' = \{0 = b_0 < b_1 < \cdots\}$ denote the (finite) set of exponents which occur in q in (3.9); $q(x) = \phi_{\mathcal{A}'}(x)$. It follows from Theorem 1.1 that

$$F_{\mathcal{A}',2}(x) = \frac{1}{\phi_{\mathcal{A}'}(x)} = \frac{1}{q(x)} = \frac{\phi_{\mathcal{A}}(x)}{1+x^T}. \tag{3.10}$$

Equation (1.10) now follows from (3.9) and (3.10). One might hope that $(\mathcal{A}')' = \mathcal{A}$, but that will not be the case if $\mathcal{A}'$ has a smaller period than $\mathcal{A}$. For example, if $\mathcal{A}_d = \{0, \dots, d-1\}$, then $\phi_{\mathcal{A}_d}(x)(1+x) = 1 + x^d$, so, regardless of d, $\mathcal{A}'_d = \{0, 1\}$. In terms of (1.10), $f_{\mathcal{A}_d,2}(n)$ is odd if and only if $n \equiv 0, 1 \pmod d$ (as proved in [10]) and $f_{\mathcal{A}'_d,2}(n)$ is odd if and only if $n \equiv 0, 1, \dots, d-1 \pmod d$. That is, $f_{\mathcal{A}'_d,2}(n)$ is odd for all $n \ge 0$, which is true, because it always equals 1.

Since $(f_{\mathcal{A},2}(n) \pmod 2)$ is periodic, it is natural to ask for the proportion of even and odd values. It follows immediately from (1.10) that the density of n for which $f_{\mathcal{A}}(n)$ is odd is equal to $|\mathcal{A}'|/T$. Computations with small examples lead to the conjecture that $|\mathcal{A}'| \le \frac{T+1}{2}$. This conjecture is false. The smallest such example we have found is $\mathcal{A}_0 = \{0, 1, 5, 9, 10\}$. It turns out that the period of $\mathcal{A}_0$ is 33 and $|\mathcal{A}'_0| = 18 > \frac{33+1}{2}$. On the other hand, it is well-known that if $\phi_{\mathcal{A}}$ is primitive, then $|\mathcal{A}'| = \frac{T+1}{2}$; see [5, §4] and [9, p.449].

Proof of Corollary 1.2(2). By Lemma 2.3, if $\mathcal{A}$ is infinite, then the sequence $(f_{\mathcal{A},2}(n) \pmod 2)$ is ultimately periodic if and only if $F_{\mathcal{A},2}(x)$ is a rational function, and by Theorem 1.1, this is so if and only if $\phi_{\mathcal{A}}(x)$ is a rational function. Suppose

$$\phi_{\mathcal{A}}(x) = a(x) + \frac{q(x)}{1+x^T} \in \mathbb{F}_2(x), \tag{3.11}$$

where $a, q \in \mathbb{F}_2[x]$, $\deg a < N$ and $\deg q < T$ and $q(x) = 1 + \sum_i x^{b_i}$. Recall that $m \in \mathcal{A}$ if and only if x^m appears in $\phi_{\mathcal{A}}(x)$. By (3.11), this holds for $m > N$ if and only if there exists $b_i \in \mathcal{A}'$ so that $N \equiv b_i \pmod T$. □

We conclude this section with proofs of Theorem 1.3(1) and (2).

Proof of Theorem 1.3(1). Let $f(n) := f_{\{0,1,2\},2}(n)$ and suppose $f(n+T) \equiv f(n) \pmod d$ for all sufficiently large n, where T is minimal. By (1.8),

$$f(2m) = f(m) + f(m-1) \quad \text{and} \quad f(2m+1) = f(m) \tag{3.12}$$

for all m. If $T = 2k$ is even, then for all sufficiently large m,

$$\begin{aligned} f(2m+2k+1) \equiv f(2m+1) \pmod{d} \implies \\ f(m+k) \equiv f(m) \pmod{d}, \end{aligned} \tag{3.13}$$

violating the minimality of T, since $k = T/2$.

If $T = 2k+1$ is odd, then for all sufficiently large m,

$$\begin{aligned} f(2m+2k+2) \equiv f(2m+1) \pmod{d} \implies \\ f(m+k+1) + f(m+k) \equiv f(m) \pmod{d}, \end{aligned} \tag{3.14}$$

and

$$\begin{aligned} f(2m+2k+3) \equiv f(2m+2) \pmod{d} \implies \\ f(m+k+1) \equiv f(m) + f(m+1) \pmod{d}. \end{aligned} \tag{3.15}$$

Together, these imply that for all sufficiently large m,

$$\begin{aligned} f(m+k) \equiv -f(m+1) \pmod{d} \implies \\ f(m+1) \equiv f(m+1+(2k-1)) \pmod{d}, \end{aligned} \tag{3.16}$$

which implies that f has a period of $2k-2$. If $k > 1$, then $0 < 2k-2 < 2k+1$ gives a contradiction. If $k = 1$, then $T = 3$.

We now show that $d = 2$. First, $f(2^r - 1) = f(2^{r-1} - 1)$ and so by induction, $f(2^r - 1) = f(1) = 1$. Thus, $f(2^r) = f(2^{r-1}) + f(2^{r-1} - 1) = f(2^{r-1}) + 1$ and so by induction, $f(2^r) = r + 1$, implying that $f(2^r + 1) = f(2^{r-1}) = r$ and $f(2^r+2) = f(2^{r-1}) + f(2^{r-1}+1) = r + r - 1$. Thus, d divides each $f(2^r + 2) - f(2^r - 1) = 2r - 1 - 1$ for sufficiently large r. Therefore, $d = 2$. □

Proof of Theorem 1.3(2). Suppose $\mathcal{A} = \{0, 1\}$ and $b \geq 3$. Then $f(n) := f_{\mathcal{A},b}(n) = 1$ if n is a sum of distinct powers of b, and 0 otherwise. Suppose that for $n > U$,

$$f(n+T) \equiv f(n) \pmod{d} \tag{3.17}$$

and $d \geq 2$. Then, $f(m) \in \{0, 1\}$ implies that $f(n+T) = f(n)$. Choose j so large that $b^j > T, U$ and suppose that f satisfies (3.17). Then $f(b^j) = 1$, hence $f(b^j + T) = 1$, and so $T = \sum_k b^{r_k}$ with distinct $r_k < j$. But then $f(b^j + 2T) = 1$ by periodicity, and so $b^j + 2\sum_k b^{r_k}$ must be also a sum of distinct powers of b, violating the uniqueness of the (standard) base-b representation. □

4. Examples

Example 4.1. The periodicity of $f_{\mathcal{A}_d,2}(n)$ was established in [10], motivated by the interpretation of the Stern sequence. In her dissertation, Dennison [6] studied a variation on the Stern sequence defined by flipping the recurrence (1.11) to a two-parameter family of sequences. The periodicities discovered in [6] for $\mathcal{A} = \{0, 1, 3\}$ and $\mathcal{A} = \{0, 2, 3\}$ led Reznick to suggest that Anders and Weber Lansing look at generalizations as the topic for their 2010 summer research project [1].

For $\alpha, \beta \in \mathbb{C}$, define $b_{\alpha,\beta}(n)$ by

$$\begin{aligned} b_{\alpha,\beta}(1) = \alpha, \qquad b_{\alpha,\beta}(2) = \beta, \\ b_{\alpha,\beta}(2n) = b_{\alpha,\beta}(n) + b_{\alpha,\beta}(n+1) \text{ for } n \geq 2, \\ b_{\alpha,\beta}(2n+1) = b_{\alpha,\beta}(n) \text{ for } n \geq 1. \end{aligned} \tag{4.1}$$

(In order for the recurrence to be unambiguous, it cannot be applied to $b_{\alpha,\beta}(2)$; the value of $b_{\alpha,\beta}(0)$ plays no further role.) It is proved in [6] that $b_{0,1}(n+2) = f_{\{0,2,3\},2}(n)$ for $n \geq 0$. It was also proved there by an argument similar to the proof of Theorem 3.1 that $b_{0,1}(n) \equiv b_{0,1}(n+7) \mod 2$, and is odd when $n \equiv 0, 2, 3, 4 \pmod 7$. This suggested looking at $f_{\{0,1,3\},2}(n)$, which is also periodic with period 7, and is odd when $n \equiv 0, 1, 2, 4 \pmod 7$.

The proofs of these facts are now straightforward in view of Theorem 1.1; we have in $\mathbb{F}_2(x)$:

$$F_{\{0,2,3\}}(x) = \frac{1}{1+x^2+x^3} = \frac{(1+x+x^3)(1+x)}{1+x^7} = \frac{1+x^2+x^3+x^4}{1+x^7};$$

$$F_{\{0,1,3\}}(x) = \frac{1}{1+x+x^3} = \frac{(1+x^2+x^3)(1+x)}{1+x^7} = \frac{1+x+x^2+x^4}{1+x^7}.$$

Thus, $\{0,2,3\}' = \{0,2,3,4\}$ and $\{0,1,3\}' = \{0,1,2,4\}$.

Example 4.2. For $r \geq 2$, define the sets $\mathcal{A}_r = \{0,1,2,\ldots,2^r\}$ and $\mathcal{B}_r = \{0,1,3,\ldots,2^r-1\}$, and let $g_r = \phi_{\mathcal{A}_r}$ and $h_r = \phi_{\mathcal{B}_r}$ for short. Then $g_r(x) = 1 + xh_r(x)$, so in $\mathbb{F}_2[x]$,

$$\begin{aligned} g_r(x)h_r(x) = h_r(x) + xh_r^2(x) = h_r(x) + xh_r(x^2) = \\ 1 + \sum_{\ell=1}^{r} x^{2^\ell - 1} + x + \sum_{\ell=1}^{r} x^{2^{\ell+1}-2+1} = 1 + x^{2^{r+1}-1}. \end{aligned} \tag{4.2}$$

This in itself does not establish that $\mathcal{A}_r, \mathcal{B}_r$ are complementary, or that they both have period $2^{r+1}-1$. If either period T were a proper factor of $2^{r+1}-1$, then since T is odd, $T \leq \frac{1}{3}(2^{r+1}-1) < 2^r - 1 < 2^r$, a contradiction. Thus g_r and h_r each have period $2^{r+1}-1$.

We may interpret this result combinatorially: $f_{\mathcal{A}_r,2}(n)$ is the number of ways to write

$$n = \sum_{i=0}^{\infty} \epsilon_i 2^{i+k_i}, \tag{4.3}$$

where $\epsilon_i \in \{0,1\}$ and $0 \leq k_i \leq r$, and $f_{\mathcal{A}_r,2}(n)$ is even, except when there exists $\ell < r$ so that $n \equiv 2^\ell - 1 \pmod{2^{r+1}-1}$. The infinite version of this example can be found in [5, §5].

Example 4.3. We return to $\mathcal{A} = \{0,1,4,9\}$; in $\mathbb{F}_2[x]$,

$$\phi_{\mathcal{A}}(x) = 1 + x + x^4 + x^9 = (1+x)^4(1+x+x^2)(1+x^2+x^3). \tag{4.4}$$

Note that $1+x$ has period 1, $1+x+x^2$ has period 3, and we have already seen that $1+x+x^3$ has period 7. Since the maximum exponent in (4.4) is $\leq 2^2$, Lemma 2.1 implies that the period of $\mathcal{A}$ divides $4 \cdot \mathrm{lcm}(1,3,7) = 84$. Another calculation shows that $\phi_{\mathcal{A}}(x)$ does not divide $1+x^{\frac{84}{p}}$ for $p = 2, 3, 7$, and so 84 is actually the period. A computation shows that $\mathcal{A}' = \{0, 1, 2, 3, \dots, 70, 75\}$ has 41 terms, as noted earlier. Thus $f_{\mathcal{A}}(n)$ is odd $\frac{41}{84}$ of the time and even $\frac{43}{84}$ of the time. The infinite version of this example can be found in [5, §6.1].

Example 4.4. Although Theorem 1.1 does not generalize to all $\mathcal{A}$ if $(b, d) \neq (2, 2)$, there are a few exceptional cases. Problem B2 on the 1983 Putnam [7] in effect asked for a proof that for $\mathcal{A} = \{0, 1, 2, 3\}$,

$$f_{\mathcal{A},2}(n) = \left\lfloor \frac{n}{2} \right\rfloor + 1. \tag{4.5}$$

This can be seen directly from (1.4), since $\phi_{\mathcal{A},2}(x) = (1 + x)(1 + x^2) = \frac{1-x^4}{1-x}$, hence $F_{\mathcal{A},2}(x)$ telescopes to $\frac{1}{(1-x)(1-x^2)}$. It follows immediately that $f_{\mathcal{A},2}(n+2d) = f_{\mathcal{A},2}(n)+d$, and hence $f_{\mathcal{A},2}$ is periodic mod d with period $2d$, for each $d \geq 2$. A similar phenomenon occurs for $\mathcal{A}_b = \{0, 1, \dots, b^2 - 1\}$, so that $\phi_{\mathcal{A}_b,b}(x) = \frac{1-x^{b^2}}{1-x}$ and $F_{\mathcal{A}_b,b}(x) = \frac{1}{(1-x)(1-x^b)}$, implying that $f_{\mathcal{A}_b,b}(n) = \lfloor \frac{n}{b} \rfloor + 1$ and $f_{\mathcal{A}_b,b}(n + bd) = f_{\mathcal{A}_b,b}(n) + d$.

Example 4.5. Let $\mathcal{A} = \{0\} \cup (2\mathbb{N} + 1)$ (all non-zero digits in (1.1) are odd). Then

$$\phi_{\mathcal{A}}(x) = 1 + \sum_{i=0}^{\infty} x^{2i+1} = 1 + \frac{x}{1 - x^2} = \frac{1 + x - x^2}{1 - x^2}. \tag{4.6}$$

Working in $\mathbb{F}_2(x)$, we have

$$F_{\mathcal{A},2}(x) = \frac{1 - x^2}{1 + x - x^2} = \frac{(1 + x)^2}{1 + x + x^2} = 1 + \frac{x}{1 + x + x^2} = 1 + \frac{x + x^2}{1 + x^3}. \tag{4.7}$$

Thus, $f_{\mathcal{A},2}(n)$ is odd if and only if $n = 0$ or n is not a multiple of 3.

Example 4.6. Let $\mathcal{A}^{\{k\}} := \mathbb{N} \setminus \{k\}$. By Theorem 1.1,

$$\begin{aligned} \phi_{\mathcal{A}^{\{k\}}} &= \frac{1}{1+x} - x^k = \frac{1 - x^k - x^{k+1}}{1+x} \\ \implies F_{\mathcal{A}^{\{k\}}}(x) &= \frac{1+x}{1 + x^k + x^{k+1}} \\ \implies F_{\mathcal{A}^{\{1\}}}(x) &= \frac{1+x}{1+x+x^2} = \frac{(1+x)^2}{1+x^3} = \frac{1+x^2}{1+x^3}. \end{aligned} \tag{4.8}$$

Thus $f_{\mathbb{N}\setminus\{1\},2}(n)$ is odd precisely when $n \equiv 0, 2 \pmod 3$. This may be contrasted with $f_{\{0,1,2\},2}(n)$, which is odd precisely when $n \equiv 0, 1 \pmod 3$.

References

[1] K. Anders and J. Weber, *The parity of generalized binary representations*, REGS report, Aug. 30, 2010.

[2] E. Berlekamp, *Algebraic coding theory, Revised 1984 edition* Aegean Park Press, Laguna Hills, CA, 1984. MR0238597 (38 #6873) (review of first edition).

[3] R. P. Brent and P. Zimmermann, *The great trinomial hunt*, Trans. Amer. Math. Soc., **58** (2011), 233-239, MR2768116.

[4] R. F. Churchhouse, *Congruence properties of the binary partition function*, Proc. Cambridge Philos. Soc., **66** (1969), 371376, MR0248102 (40 #1356).

[5] J. N. Cooper, D. Eichhorn and K. O'Bryant, *Reciprocals of binary series*, Int. J. Number Theory **2**, (2006), no. 4, 499522, MR2281861 (2007h:11015).

[6] M. Dennison, *A sequence related to the Stern sequence*, Ph.D. dissertation, University of Illinois at Urbana-Champaign, 2010.

[7] L. F. Klosinski, G. L. Alexanderson and A. P. Hillman, *The William Lowell Putnam Mathematical Competition*, Amer. Math. Monthly, **91** (1984), 487495, MR1540495.

[8] D. H. Lehmer, *On Stern's diatomic series*, Amer. Math. Monthly, **36** (1929), 59–67, MR1521653.

[9] R. Lidl and H. Niederreiter, *Finite Fields, 2nd ed.* Encyclopedia of Mathematics and its Applications, **20**. Cambridge University Press, Cambridge, 1997, MR1429394 (97i:11115).

[10] B. Reznick, *Some binary partition functions*, Analytic number theory (Allerton Park, IL, 1989), 451477, Progr. Math., 85, Birkhäuser Boston, Boston, MA, 1990, MR1084197 (91k:11092).

[11] B. Reznick, *Regularity properties of the Stern enumeration of the rationals*, J. Integer Seq. **11** (2008), no. 4, Article 08.4.1, MR2447843 (2009g:11016).

[12] Ø. J. Rødseth and J. A. Sellers, *On m-ary partition function congruences: a fresh look at a past problem*, J. Number Theory, **87** (2001), 270281, MR1824148 (2001m:11177).

[13] M. A. Stern, *Ueber eine zahlentheoretische Funktion*, J. Reine Angew. Math., **55** (1858) 193–220.

Katherine Anders
Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, IL 61801
e-mail: kaanders@illinois.edu

Melissa Dennison
Department of Math and Computer Science, Baldwin-Wallace College, Berea, OH 44017
e-mail: mdenniso@bw.edu

Jennifer Weber Lansing
Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, IL 61801
e-mail: jlweber@illinois.edu

Bruce Reznick
Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, IL 61801
e-mail: reznick@illinois.edu

STERN NOTES, CHAPTER 8 (FIRST DRAFT)

BRUCE REZNICK, UIUC

1. Generalizations

The generalizations of any mathematical object are limited only by semantics and the imagination. We are interested here in presenting some situations which can be specialized to the Stern sequence and which preserve some of its properties.

Here is one which keeps the binary nature. Define two functions

$$f : \mathbb{C} \to \mathbb{C}, \qquad g : \mathbb{C}^2 \to \mathbb{C}, \tag{1}$$

and define the sequence (a_n) by

$$a_{2n} = f(a_n), \quad a_{2n+1} = g(a_n, a_{n+1}), \qquad n \ge 1, \tag{2}$$

with (a_0, a_1) to be determined as initial conditions. If we now define $L, R : \mathbb{C}^2 \to \mathbb{C}^2$ by

$$L(u, v) := (f(u), g(u, v)), \qquad R(u, v) := (g(u, v), f(v)), \tag{3}$$

then the fundamental dyadic nature of the Stern sequence is preserved, inasmuch as

$$(a_{2n}, a_{2n+1}) = L(a_n, a_{n+1}), \quad (a_{2n+1}, a_{2n+2}) = R(a_n, a_{n+1}) \tag{4}$$

The value of a_n is determined by encoding the binary representation of n into a word of operators taken from the alphabet $\{L, R\}$, as applied to the initial conditions. This is nice as far as it goes, but probably too general to be very interesting.

If we assume that f and g are linear; specifically,

$$a_{2n} = \alpha a_n, \qquad a_{2n+1} = \beta a_n + \gamma a_{n+1}, \tag{5}$$

then so are L and R, and we can copy a picture from the early notes: the mappings of the consecutive pairs has the repeated pattern

$$\begin{array}{ccc} & \begin{bmatrix} x \\ y \end{bmatrix} & \\ \swarrow & & \searrow \\ L\begin{bmatrix} x \\ y \end{bmatrix} & & R\begin{bmatrix} x \\ y \end{bmatrix} \end{array} \tag{6}$$

where

$$L = \begin{bmatrix} \alpha & 0 \\ \beta & \gamma \end{bmatrix}, \qquad R = \begin{bmatrix} \beta & \gamma \\ 0 & \alpha \end{bmatrix}. \tag{7}$$

The analysis of this situation depends on the semigroup of 2×2 matrices generated by L and R. The most interesting cases would seem to occur when all entries are roots of unity. Even this appears to have too many cases to support a unifying analysis.

We shall discuss here two somewhat more restrictive generalizations of the Stern sequence. In the first, we define the sequence by its power series and work backwards to find the recurrence. For $\alpha, \beta \in \mathbb{C}$, let

$$\Phi_{\alpha,\beta}(X) := \sum_{n=0}^{\infty} a_{\alpha,\beta}(n) X^n = X \prod_{j=0}^{\infty} (1 + \alpha X^{2^j} + \beta X^{2^{j+1}}). \tag{8}$$

By the same reasoning as before, $a_{\alpha,\beta}(0) = 0$ and, for $n \ge 1$,

$$a_{\alpha,\beta}(n) = \sum_{r,s\ge 0} c_{r,s}(n) \alpha^r \beta^s, \tag{9}$$

where $c_{r,s}(n)$ is the number of ways to write

$$n - 1 = \sum_{j=0}^{\infty} \epsilon_j 2^j, \qquad \epsilon_j \in \{0,1,2\}, \tag{10}$$

using exactly r 1's and s 2's.

To find the recurrence, observe that

$$\begin{aligned} &(1 + \alpha X + \beta X^2)\Phi_{\alpha,\beta}(X^2) = X\Phi_{\alpha,\beta}(X) \\ \implies\ &(1 + \alpha X + \beta X^2) \sum_{n=0}^{\infty} a_{\alpha,\beta}(n) X^{2n} = X \sum_{n=0}^{\infty} a_{\alpha,\beta}(n) X^n \\ \implies\ &a_{\alpha,\beta}(2n) = \alpha a_{\alpha,\beta}(n), \qquad a_{\alpha,\beta}(2n+1) = \beta a_{\alpha,\beta}(n) + 1 \cdot a_{\alpha,\beta}(n+1). \end{aligned} \tag{11}$$

In other words, this is (5) in the case that $\gamma = 1$. (Of course, if we set $\alpha = \beta = 1$, we recover the Stern sequence.) If $\gamma \ne 1$, we have neither a reasonable power series factorization, nor a reasonable combinatorial interpretation.

The second general class in which we can say something is when we can retain part of the diatomic array structure by taking $\alpha = 1$. Consider

$$\begin{array}{lllll} a & b & & & \\ a & \beta a + \gamma b & b & & \\ a & \beta a + \gamma(\beta a + \gamma b) & \beta a + \gamma b & \beta(\beta a + \gamma b) + \gamma b & b \\ & \cdots & & & \end{array} \tag{12}$$

More formally, let $Z(r,k) = Z(r,k;a,b;\beta,\gamma)$ for $r \ge 0$ and $0 \le k \le 2^r$ be defined by:

$$\begin{aligned} Z(0,0) = a, &\quad Z(0,1) = b; \\ Z(r,2k) &= Z(r-1,k), \ \text{for } r \ge 1; \\ Z(r,2k+1) &= \beta Z(r-1,k) + \gamma Z(r-1,k+1), \ \text{for } r \ge 1. \end{aligned} \tag{13}$$

As with the simpler diatomic array, the entries are still linear in the initial conditions and have the same self-similarity, but the mirror condition flips the new parameters as well:

$$Z(r, k; a, b; \beta, \gamma) = Z(r, 2^r - k; b, a; \gamma, \beta). \tag{14}$$

This is important as we look for the "eigenarrays" that reduce the analysis of that array to a sequence.

First observe that if $\beta a + \gamma b = b$, then $Z(0, 1) = Z(1, 1)$, and as before, this means that the rows of the array will nest, with the first half of each reproducing the previous row. We can realize this by taking $(a, b) = (1 - \gamma, \beta)$, so that the first few rows of the array are

$$\begin{array}{lllll} 1-\gamma & \beta & & & \\ 1-\gamma & \beta & \beta & & \\ 1-\gamma & \beta & \beta & \beta^2+\beta\gamma & \beta \\ \cdots & & & & \end{array} \tag{15}$$

If $\beta = \gamma = 1$, this recovers the standard Stern array. In any case, the underlying sequence is given by:

$$\begin{gathered} t_{\beta,\gamma}(0) = 1 - \gamma, \quad t_{\beta,\gamma}(1) = \beta; \\ t_{\beta,\gamma}(2n) = t_{\beta,\gamma}(n), \qquad t_{\beta,\gamma}(2n+1) = \beta t_{\beta,\gamma}(n) + \gamma t_{\beta,\gamma}(n+1). \end{gathered} \tag{16}$$

The eigenarray is trivially zero only if $\beta = 0$ and $\gamma = 1$, but the analysis of $Z(r, k; a, b; 0, 1)$ is also trivial: $Z(r, 0) = a$ and $Z(r, k) = b$ for $k \geq 1$. This gives what might be called the *forward eigenarray.* Eventually, we'll give a closed formula, but not in these notes. (It probably makes sense to divide by β.)

The *backward eigenarray* comes from solving $\beta a + \gamma b = a$; so that we might take $(a, b) = (\gamma, 1 - \beta)$:

$$\begin{array}{rrrrr} & & & \gamma & 1-\beta \\ & & \gamma & \gamma & 1-\beta \\ \gamma & \beta\gamma+\gamma^2 & \gamma & \gamma & 1-\beta \\ & & & & \cdots \end{array} \tag{17}$$

This nests from the right, not the left, and again, with $\beta = \gamma = 1$, is familiar in the Stern situation as the reversal of the basic array.

The mirror symmetry connection of these two eigenarrays is clear and will not be elaborated on. We remark that $(1 - \gamma, \beta)$ and $(\gamma, 1 - \beta)$ are linearly independent unless $\beta + \gamma = 1$, in which case an alternate approach is useful.

In the rest of this section, we shall discuss three specific cases:

(1) Hellinger's Function: $Z(r, k; a, b, 1 - p, p)$;
(2) Stern Polynomials: $S(0; \lambda) = 0, S(1; \lambda) = 1$ and $S(2n; \lambda) = \lambda S(n; \lambda)$ and $S(2n + 1; \lambda) = S(n; \lambda) + S(n + 1; \lambda)$;

(3) The sullen cousin (name subject to change) : $w(0) = 0, w(1) = 1$ and $w(2n) = w(n)$ and $w(2n+1) = w(n+1) - w(n)$.

The labeling of these sequences is inconsistent and ought to be corrected in the "second edition". I will also refer to class handouts in some cases without directly putting them in.

2. Hellinger's function

In this section we discuss the arrays $Z(r, k; a, b, \beta, \gamma)$ with $\beta + \gamma = 1$. For reasons that will become clear, it is sensible to write $\beta = 1 - p$ and $\gamma = p$. It follows from linearity that

$$Z(r, k; a, b; 1-p, p) = aZ(r, k; 1, 1; 1-p, p) + (b-a)Z(r, k; 0, 1; 1-p, p), \tag{18}$$

and it is trivial to verify that for all (r, k),

$$Z(r, k; 1, 1; 1-p, p) = 1. \tag{19}$$

This leaves $Z(r, k; 0, 1; 1-p, p)$, for which we give the first few rows:

$$\begin{array}{lllllllll} 0 & 1 & & & & & & & \\ 0 & p & 1 & & & & & & \\ 0 & p^2 & p & 2p-p^2 & 1 & & & & \\ 0 & p^3 & p^2 & 2p^2-p^3 & p & p+p^2-p^3 & 2p-p^2 & 3p-3p^2+p^3 & 1 \\ & \dots & & & & & & & \end{array} \tag{20}$$

At this point it makes sense to exploit the recurrence and explicitly define a family of functions $f_p : [0, 1] \to \mathbb{C}$ by

$$f_p\left(\frac{k}{2^r}\right) := Z(r, k; 0, 1; 1-p, p). \tag{21}$$

This is well-defined, as we have seen in other cases, because $\frac{k}{2^r} = \frac{2k}{2^{r+1}}$, and the recurrence gives

$$f_p\left(\frac{2k+1}{2^{r+1}}\right) = (1-p)f_p\left(\frac{k}{2^r}\right) + pf_p\left(\frac{k+1}{2^r}\right). \tag{22}$$

Furthermore, the self-similarity gives for $0 \le k \le 2^r$:

$$\begin{gathered} Z(r+1, k; 0, 1; 1-p, p) = Z(r, k; 0, p; 1-p, p) \\ \implies f_p\left(\frac{k}{2^{r+1}}\right) = p \cdot f_p\left(\frac{k}{2^r}\right); \\ Z(r+1, 2^r+k; 0, 1; 1-p, p) = Z(r, k; p, 1; 1-p, p) \\ = Z(r, k; p, p; 1-p, p) + Z(r, k; 0, 1-p; 1-p, p) \implies \\ f_p\left(\frac{2^r+k}{2^{r+1}}\right) = p + (1-p) \cdot f_p\left(\frac{k}{2^r}\right). \end{gathered} \tag{23}$$

In other words, for *dyadic* $x \in [0,1]$, we have

$$f_p\left(\frac{x}{2}\right) = pf_p(x); \qquad f_p\left(\frac{1+x}{2}\right) = p + (1-p)f_p(x). \tag{24}$$

It is a routine exercise to prove by induction that

$$f_p\left(\frac{k+1}{2^r}\right) - f_p\left(\frac{k}{2^r}\right) = p^{r-a}(1-p)^a, \tag{25}$$

where a is the number of 1's in the binary expansion of k. Thus, for example, in the third row of the array, the differences are, in order

$$p^3,\ p^2(1-p),\ p^2(1-p),\ p(1-p)^2,\ p^2(1-p),\ p(1-p)^2,\ p(1-p)^2,\ (1-p)^3. \tag{26}$$

In the special case that $p \in (0,1)$, the preceding argument is enough to show that f_p extends to a continuous strictly increasing function from $[0,1]$ to itself, which is singular for $p \neq \frac{1}{2}$. In this case, there is also a probabilistic interpretation. Consider a game in which you start with $x \in [0,1]$ units and are allowed to bet $y \leq \max\{x, 1-x\}$ units, with the goal of reaching "1", and loss if you hit "0". With probability p you win, and have $x+y$ units, and with probability $1-p$ you lose, and have $x-y$ units. The "bold" strategy (a technical term) is to bet x when you have $x \leq \frac{1}{2}$ and to bet $1-x$ when you have $x \geq \frac{1}{2}$. (This is the optimal strategy for reaching "1", if $p < \frac{1}{2}$.) In this case, reference to equation (24) shows that the probability of victory starting with x units is exactly $f_p(x)$.

For $0 \leq j \leq r$, let

$$A_{j,r} = \{2^{i_1} + \cdots + 2^{i_j} : r-1 \geq i_1 > \cdots > i_j \geq 0\} \tag{27}$$

be the integers in $[0, 2^r - 1]$ with j 1's in their binary expansions. We then have

$$\Delta_{j,r} := \sum_{k \in A_{j,r}} \left(f_p\left(\frac{k+1}{2^r}\right) - f_p\left(\frac{k}{2^r}\right)\right) = \binom{r}{j} p^{r-j}(1-p)^j. \tag{28}$$

It follows from the Law of Large Numbers that this increase is concentrated on $A_{j,r}$ where $j \approx r(1-p)$. In fact, it can be proved that the measure df_p determined by f_p is supported on those $x \in [0,1]$ for which the density of 1's in its dyadic expansion is $1-p$. If $p \neq \frac{1}{2}$, this set has measure zero and if $p_1 \neq p_2$, the corresponding sets of support are disjoint. There is a fair bit of literature on this topic, and references will show up in the later versions. In particular, the first person to have studied it appears to have been Ernst Hellinger.

3. Stern polynomials

We define the Stern polynomials by

$$\begin{gathered} S(0;\lambda) = 0, \qquad S(1;\lambda) = 1, \\ S(2n;\lambda) = \lambda S(n;\lambda), \qquad S(2n+1;\lambda) = S(n;\lambda) + S(n+1;\lambda). \end{gathered} \tag{29}$$

This is (5) with $\alpha = \lambda$, $\beta = \gamma = 1$, and so has the generating function

$$S(X;\lambda) := \sum_{n=0}^{\infty} S(n;\lambda)X^n = X\prod_{j=0}^{\infty}(1 + \lambda X^{2^j} + X^{2^{j+1}}). \tag{30}$$

These polynomials can be explicitly evaluated for several values of λ. Of course, $S(n;1) = s(n)$. We have already seen that

$$S(3n;-1) = 0, \qquad S(3n+1;-1) = 1, \qquad S(3n+2;-1) = -1, \tag{31}$$

and two easy inductions (or an appeal to the generating function) show that

$$S(2n;0) = 0, \qquad S(2n+1;0) = 1 \tag{32}$$

and

$$S(n;2) = n. \tag{33}$$

This last identity implies that the $S(n,\lambda)$'s are distinct for distinct n.

The Stern polynomials satisfy many identities which might be considered the "explanation" for identities satisfied by the Stern sequence. For example, it is easy to show by induction that for $0 \le k \le 2^r$,

$$S(2^r n \pm k;\lambda) = S(2^r - k;\lambda)S(n;\lambda) + S(k;\lambda)S(n \pm 1;\lambda). \tag{34}$$

It may be more helpful in understanding this to note the analogy to the Stern sequence: the Stern polynomials can be construed as coming from a modified diatomic array in which consecutive terms are added, but the previous row is multiplied by λ before coming down.

$$\begin{array}{lllll} a & b & & & \\ \lambda a & a+b & \lambda b & & \\ \lambda^2 a & (1+\lambda)a+b & \lambda(a+b) & a+(1+\lambda)b & \lambda^2 b \\ & \dots & & & \end{array} \tag{35}$$

As before, if $(a,b) = (S(n;\lambda), S(n+1;\lambda))$, then the r-th row above lists $S(2^r n + k;\lambda)$ for $0 \le k \le 2^r$.

It is also easy to prove by induction that

$$S(2^r - 1;\lambda) = 1 + \lambda + \cdots + \lambda^{r-1} := (\lambda)_r. \tag{36}$$

(We view $(\lambda)_r \in \mathbb{Z}[\lambda]$; if it is to be evaluated at $\lambda = 1$, we simply replace it by r.) Thus, we have the following specializations of (33):

$$S(2^r n \pm 1;\lambda) = (\lambda)_r \cdot S(n;\lambda) + S(n \pm 1;\lambda). \tag{37}$$

Another family of identities generalizes. Let $t_n = \frac{2^n - (-1)^n}{3}$; then as we have previously seen, $t_n = 2t_{n-1} - (-1)^n = t_{n-1} + 2t_{n-2}$. This implies that

$$S(t_n;\lambda) = S(t_{n-1};\lambda) + S(2t_{n-2};\lambda) = S(t_{n-1};\lambda) + \lambda S(t_{n-2};\lambda). \tag{38}$$

For fixed λ, this is a linear recurrence with characteristic equation $X^2 - X - \lambda$, and since $S(t_0;\lambda) = 0$ and $S(t_1;\lambda) = 1$, we obtain a closed form. If $\lambda \neq -\frac{1}{4}$, then

$$S(t_n;\lambda) = \frac{1}{\sqrt{1+4\lambda}}\left(\left(\frac{1+\sqrt{1+4\lambda}}{2}\right)^n - \left(\frac{1+\sqrt{1-4\lambda}}{2}\right)^n\right), \tag{39}$$

and $S(t_n;-\frac{1}{4}) = \frac{n}{2^{n-1}}$. (In this case the characteristic equation has a double root.)

One of my favorite Stern identities generalizes:

$$\begin{aligned} S((2^r-1)^2;\lambda) &= S(2^{r+1}(2^{r-1}-1)+1;\lambda) = (\lambda)_{r+1}(\lambda)_{r-1} + 1\cdot\lambda^{r-1} \\ &= \frac{(1-\lambda^{r+1})(1-\lambda^{r-1}) + \lambda^{r-1}(1-\lambda)^2}{(1-\lambda)^2} = \frac{(1-\lambda^r)^2}{(1-\lambda)^2} = S(2^r-1;\lambda)^2. \end{aligned} \tag{40}$$

Since

$$S(2^r n \pm n;\lambda) = S(n;\lambda)(S(2^r-n;\lambda) + S(n\pm 1;\lambda)), \tag{41}$$

it follows that each $S(n;\lambda)$ is a factor of infinitely many other ones. (We won't prove it here, but this is also true for $S(n;\lambda)^k$, where k is any positive integer.)

There are two closed forms for $S(n;\lambda)$. Both are based on assuming that n is odd. We have previously written $n \sim [a_1,\ldots,a_{2v+1}]$ to indicate that the binary representation of n consists of a_1 1's, a_2 0's, a_3 1's, etc. It is convenient for the first case to say that $n = [[a_1,\ldots,a_t]]$ is defined recursively by:

$$[[a]] = 2^a - 1; \qquad [[a_1,\ldots,a_t]] = 2^{a_1+\cdots+a_t} - [[a_2,\ldots,a_t]] \tag{42}$$

If $t = 2v+1$ is odd, then $[[a_1,\ldots,a_{2v+1}]] \sim [a_1,\ldots,a_{2v+1}]$, but t could be even.

Theorem 1. *Using the preceding notations, if $n = [[a_1,\ldots,a_t]]$ and $r = \sum t_i$, then*

$$\frac{S(n;\lambda)}{S(2^r-n;\lambda)} = \frac{S([[a_1,\ldots,a_t]];\lambda)}{S([[a_2,\ldots,a_t]];\lambda)} = (\lambda)_{a_1} + \cfrac{\lambda^{a_1}}{(\lambda)_{a_2} + \cfrac{\lambda^{a_2}}{\cdots + \cfrac{\lambda^{a_{t-1}}}{(\lambda)_{a_t}}}}. \tag{43}$$

Proof. The first inductive step is for $t = 1$. If $n = [[a]] = 2^a - 1$, then $2^r - n = 1$, and, indeed,

$$\frac{S(n;\lambda)}{S(2^r-n;\lambda)} = \frac{S(2^a-1;\lambda)}{S(1;\lambda)} = (\lambda)_a. \tag{44}$$

For $t = 2$, we have $2^{a+b} - 2^b + 1 = 2^b(2^a-1)+1$, so that

$$\begin{aligned} \frac{S([[a,b]];\lambda)}{S([[b]];\lambda)} &= \frac{S(2^{a+b}-2^b+1;\lambda)}{S(2^b-1;\lambda)} \\ &= \frac{S(2^b-1;\lambda)S(2^a-1;\lambda) + S(1;\lambda)S(2^a;\lambda)}{S(2^b-1;\lambda)} = (\lambda)_a + \frac{\lambda^a}{(\lambda)_b}, \end{aligned} \tag{45}$$

as desired.

Now assume the inductive hypothesis, let $n^* = 2^r - n$ and $n^{**} = 2^{r-a_1} - n^*$. Then

$$n = 2^r - 2^{r-a_1} + n^{**} = 2^{r-a_1}(2^{a_1} - 1) + n^{**}. \tag{46}$$

It follows that

$$\begin{aligned} S(n;\lambda) &= S(2^{r-a_1} - n^{**};\lambda)S(2^{a_1} - 1;\lambda) + S(n^{**};\lambda)S(2^{a_1};\lambda) \\ &= (\lambda)_{a_1} S(n^*;\lambda) + \lambda^{a_1} S(n^{**};\lambda), \end{aligned} \tag{47}$$

and so,

$$\frac{S(n;\lambda)}{S(n^*;\lambda)} = (\lambda)_{a_1} + \frac{\lambda^{a_1}}{\frac{S(n^*;\lambda)}{S(n^{**};\lambda)}}, \tag{48}$$

as is needed to complete the induction. □

Of course, if $\lambda = 1$, then $(\lambda)_a = a$, and this reduces to one of the formulas earlier in the notes, from the Brocot array.

The second formula is both more familiar and considerably messier, and we omit the proof. Let

$$[\lambda]_a := \lambda^{-a}(\lambda)_a; \tag{49}$$

once again, for $\lambda = 1$, this reduces to a, and we have to omit $\lambda = 0$. We again write $n \sim [a_1, \dots, a_{2v+1}]$:

$$\frac{S(n;\lambda)}{S(n+1;\lambda)} = [\lambda]_{a_{2t+1}} + \cfrac{\lambda^{-a_{2t+1}}}{[\lambda]_{a_{2t}} + \cfrac{\lambda^{-a_{2t}}}{\cdots + \cfrac{\lambda^{-a_2}}{[\lambda]_{a_1}}}}. \tag{50}$$

This gives $S(n;\lambda)$ and $S(n+1;\lambda)$, once the negative powers of λ have been cancelled out. We remark in support of this formula that

$$\frac{S(2^a - 1;\lambda)}{S(2^a;\lambda)} = \frac{(\lambda)_a}{\lambda^a}. \tag{51}$$

and, as in (45) above,

$$S(2^{a+b} - 2^b + 1;\lambda) = S(2^b(2^a - 1) + 1;\lambda) = (\lambda)_b(\lambda)_a + \lambda^a \tag{52}$$

Further, $2^{a+b+c} - 2^{b+c} + 2^c - 1 \sim [a, b, c]$. We note that

$$S(2^{a+b+c} - 2^{b+c} + 2^c;\lambda) = \lambda^c S(2^b(2^a - 1) + 1;\lambda) \tag{53}$$

and

$$\begin{aligned} S(2^{a+b+c} - 2^{b+c} + 2^c - 1;\lambda) &= S(2^c(2^{a+b} - 2^b + 1) - 1;\lambda) \\ &= (\lambda)_c S(2^{a+b} - 2^b + 1;\lambda) + S(2^{a+b} - 2^b;\lambda) \\ &= (\lambda)_c S(2^{a+b} - 2^b + 1;\lambda) + \lambda^b(\lambda)_a. \end{aligned} \tag{54}$$

Thus,

$$\frac{S(2^{a+b+c}-2^{b+c}+2^c-1;\lambda)}{S(2^{a+b+c}-2^{b+c}+2^c;\lambda)} = \frac{(\lambda)_c S(2^{a+b}-2^b+1;\lambda)+\lambda^b(\lambda)_a}{\lambda^c S(2^{a+b}-2^b+1;\lambda)}$$

$$= \lambda^{-c}(\lambda)_c + \frac{\lambda^{-c}}{\frac{S(2^{a+b}-2^b+1;\lambda)}{\lambda^b(\lambda)_a}} = \tag{55}$$

$$\lambda^{-c}(\lambda)_c + \frac{\lambda^{-c}}{\frac{(\lambda)_b(\lambda)_a+\lambda^a}{\lambda^b(\lambda)_a}} = \lambda^{-c}(\lambda)_c + \frac{\lambda^{-c}}{\lambda^{-b}(\lambda)_b + \frac{\lambda^{-b}}{\lambda^{-a}(\lambda)_a}}.$$

The general proof runs much the same way.

It is easy to show that $\gcd(S(n;\lambda), S(n+1;\lambda)) = 1$, but this is less impressive for polynomials than it is for integers.

For $n \geq 1$, let

$$d(n) := deg(S(n;\lambda)). \tag{56}$$

Since all coefficients of $S(n;\lambda)$ are non-negative integers, there can be no cancellation of the leading terms when two Stern polynomials are added. Thus it follows immediately from the recurrence that

$$d(2n) = d(n)+1, \qquad d(2n\pm 1) = \max\{d(n), d(n\pm 1)\} \tag{57}$$

Lemma 2. *For $n \geq 1$, we have $d(n+1)-d(n) \in \{-1,0,1\}$.*

Proof. As the distributed tables show, this is true for small n. Suppose it is true inductively, and $d(n) = d$, say. Then $d(n\pm 1) = d-1, d$ or $d+1$ by the inductive hypothesis, so that $d(2n) = d+1$ and $d(2n\pm 1) = d, d$ or $d+1$, respectively, and the proof is complete. □

Lemma 3. *For all $n \geq 1$, $d(4n\pm 1) = d(n)+1$.*

Proof. By the recurrence, we have

$$S(4n\pm 1;\lambda) = (1+\lambda)S(n;\lambda) + S(n+1;\lambda), \tag{58}$$

and since $deg((1+\lambda)S(n;\lambda)) = 1 + d(n) \geq d(n+1)$, we are done. □

Now let $\mathcal{D}(d) := \{n : d(n) = d\}$ be the set of Stern polynomials of degree d.

Theorem 4.

$$|\mathcal{D}(d)| = 3^d, \qquad \sum_{n\in\mathcal{D}(d)} n = 10^d. \tag{59}$$

Proof. We first observe that $\mathcal{D}(0) = \{1\}$, and then note that by the last lemma, each $n \in \mathcal{D}(d)$ induces $2n, 4n-1, 4n+1 \in \mathcal{D}(d+1)$. Since every integer m can be expressed as exactly one of $\{2n, 4n-1, 4n+1\}$, all cases are accounted for. □

The most strikingly interesting fact about the Stern polynomials is the fractal-like nature of its zero set, as shown in a handout. Let

$$\mathcal{Z} := \{\lambda \in \mathbb{C} : S(n;\lambda) = 0 \text{ for some } n\}. \tag{60}$$

Theorem 5. *The set* $\overline{\mathcal{Z}}$ *contains the unit circle and the interval* $(-\infty, -\frac{1}{4}]$.

Proof. The first assertion follows from the fact that every root of unity $\zeta = e^{2\pi i k/r} \neq 1$ is a root of $S(2^r - 1, \lambda)$. (It is a bit counterintuitive that $1 \in \overline{\mathcal{Z}}$, to be sure.)

For the second, we use the formula for $S(t_n;\lambda)$. We have seen that $S(t_n; -\frac{1}{4}) > 0$ for $n \geq 1$ and that for $\lambda \neq -\frac{1}{4}$, with

$$\zeta_n = e^{\frac{2\pi i}{n}}, \tag{61}$$

$$\begin{aligned} S(t_n;\lambda) = 0 \iff \left(\frac{1+\sqrt{1+4\lambda}}{2}\right)^n = \left(\frac{1+\sqrt{1-4\lambda}}{2}\right)^n \\ \iff 1 + \sqrt{1+4\lambda} = \zeta_n^k(1 - \sqrt{1+4\lambda}) \iff \sqrt{1+4\lambda} = \frac{\zeta_n^k - 1}{\zeta_n^k + 1}, \end{aligned} \tag{62}$$

with the understanding that $\zeta_n^k \neq 1$ (because $\lambda \neq -\frac{1}{4}$) and $\zeta_n^k \neq -1$ (because $\lambda \in \mathbb{C}$). A bit of algebra shows that

$$\sqrt{1+4\lambda} = \frac{\xi - 1}{\xi + 1} \iff 1 + 4\lambda = \frac{1 - 2\xi + \xi^2}{1 + 2\xi + \xi^2} \iff \lambda = \frac{-1}{\xi + 2 + \xi^{-1}} \tag{63}$$

In particular, with $\xi = \zeta_n^k$, we have

$$\lambda = \frac{-1}{2 + 2\cos(\frac{2k\pi}{n})} = \frac{-1}{4\cos^2(\frac{k\pi}{n})}. \tag{64}$$

The restriction $\xi \neq -1, 1$ implies that $\cos^2(\frac{k\pi}{n}) \neq 0, 1$. For $n \in \mathbb{N}$, the union of the points $\{\zeta_n^k\}$ is dense in the unit circle, and so the roots of $S(t_n;\lambda)$ are dense in the real interval $[-\infty, -\frac{1}{4}]$. □

4. Stern's sullen cousin

Finally, we collect some information about the sequence (w_n), defined by

$$w(0) = 0, \quad w(1) = 1, \qquad w(2n) = w(n), \quad w(2n+1) = w(n+1) - w(n). \tag{65}$$

I call this the sullen cousin, because none of the proofs are exciting (so far).

The previous techniques combine to show that the generating function is

$$W(X) := \sum_{n=0}^{\infty} w(n) X^n = X \prod_{j=0}^{\infty} (1 + X^{2^j} - X^{2^{j+1}}), \tag{66}$$

so that $w(n)$ is the number of representations (10) with an even number of 2's, minus the number with an odd number of 2's. It follows immediatedly that

$$|w(n)| \leq s(n), \tag{67}$$

and this can be seen on the class handout. The associated diatomic array shows some interesting patterns:

$$\begin{array}{ccccc} a & b & & & \\ a & b-a & b & & \\ a & b-2a & b-a & a & b \\ & & \cdots & & \end{array} \tag{68}$$

As before, if the first row is $(w(m), w(m+1))$, then the r-th row will be

$$w(2^r m), ..., w(2^r m + 2^r). \tag{69}$$

Since $w(2) = 1$ and $w(3) = 0$, this array can be used to prove a peculiar addition formula for $0 \le k \le 2^r$. We turn the proofs of the remaining properties into homework problems!

Lemma 6. *If $0 \le k \le 2^r$, then*

$$w(2^r m + k) = w(2 \cdot 2^r + k)w(m) + w(k)w(m+1). \tag{70}$$

An immediate consequence is the following:

Lemma 7. *If $w(m) = w(m')$ and $w(m+1) = w(m'+1)$, then $w(2^r m + k) = w(2^r m' + k)$.*

Observe that $w(4n+3) = w(2n+2) - w(2n+1) = w(n+1) - (w(n+1) - w(n)) = w(n)$ and $w(4n+4) = w(n+1)$. Thus, for each n, $n' = 4n+3$ satisfies the hypotheses of the lemma. An almost immediate consequence of this is

Theorem 8. *If n' is derived from n by the deletion (or addition) of two consecutive "1"'s in its binary expansion, then $w(n) = w(n')$.*

Corollary 9. *If, in the binary expansion of n, "1"'s occur only in blocks of even length, then $w(n) = 0$. (In fact, this is an "if and only if" result.)*

In view of the preceding results, it suffices to find a closed formula for $w(n)$ when the binary expansion of n does not have consecutive 1's. For this reason, we consider

$$n \sim [1, b_1, 1, \ldots, b_t, 1] :=]b_1, \ldots, b_t[. \tag{71}$$

Put recursively,

$$]b[= 2^{b+1} + 1; \qquad]b_1, \ldots, b_t[= 2^{b_t+1}]b_1, \ldots, b_{t-1}[+ 1. \tag{72}$$

Theorem 10. *In the preceding notation, we have*

$$w(]b_1, \ldots, b_t[) = p_t(-b_1, \ldots, b_t) = (-1)^t p_t(b_1, \ldots, b_t). \tag{73}$$

Here, we return to the continuant notation from earlier in the semester. This formula has some interesting consequences. Let $d_2(n)$ be the sum of the binary digits of n. Then

Corollary 11. *If $d_2(n)$ is even, then $w(n) \geq 1$; if $d_2(n)$ is odd, then $w(n) \leq 0$. In particular, exactly one of $\{w(2n), w(2n+1)\}$ is ≥ 1 and the other is ≤ 0.*

It is certainly the case that $\gcd(w(n), w(n+1)) = 1$ for all n, but it's clear from this that not all pairs of relatively prime integers occur. A relevant fact in this direction is:

Theorem 12. *For all n, if $w(n) \neq 0$, then*

$$\frac{w(n+1)}{w(n)} \leq 1. \tag{74}$$

One can also parallel the summation formulas for the Stern sequence and show that

Lemma 13. *For $r \geq 1$,*

$$\sum_{n=2^r}^{2^{r+1}-1} w(n) = 1,$$

$$\sum_{n=2^r}^{2^{r+1}-1} |w(n)| = \frac{(1+\sqrt{2})^n + (1-\sqrt{2})^n}{2}, \tag{75}$$

$$\sum_{n=2^r}^{2^{r+1}-1} w(n)^2 = 3^{r-1}.$$

Further, if

$$\sum_{n=2^r}^{2^{r+1}-1} w(n)^3 = c_r, \tag{76}$$

then $c_r = 3c_{r-1} - 4c_{r-2} - 4c_{r-3}$. (The exact formula for c_r involves roots of an irreducible cubic.)

We also showed in class using elementary estimates that

Theorem 14. *As a complex function, $W(z)$ is bounded as $z = x \to 1-$.*

Finally, we remark that my 1985 paper *Some extremal problems for continued fractions* contains a theorem equivalent to the assertion that, if $2^r \leq n \leq 2^{r+1}$, then

$$|w(r)| \leq \left(\frac{3+\sqrt{13}}{2}\right)^{r/4} \approx 1.348^r, \tag{77}$$

and this is best possible.

Much of the description of the Stern sequence can be carried over to its sullen cousin. Usually, it isn't quite as interesting, but I hope to have more to say in the next iteration of these notes.