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

$$a_{2n} = f(a_n), \quad a_{2n+1} = g(a_n, a_{n+1}), \qquad n \geq 1, \tag{2}$$

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

$$\begin{array}{lllll} a & b & & & \\ a & \beta a + \gamma b & b & & \\ a & \beta a + \gamma(\beta a + \gamma b) & \beta a + \gamma b & \beta(\beta a + \gamma b) + \gamma b & b \\ & \dots & & & \end{array} \tag{12}$$

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

$$\begin{array}{lllll} 1-\gamma & \beta & & & \\ 1-\gamma & \beta & \beta & & \\ 1-\gamma & \beta & \beta & \beta^2+\beta\gamma & \beta \\ \dots & & & & \end{array} \tag{15}$$

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 $(a,b) = (\gamma, 1-\beta)$:

$$\begin{array}{rrrrr} & & & \gamma & 1-\beta \\ & & \gamma & \gamma & 1-\beta \\ \gamma & \beta\gamma+\gamma^2 & \gamma & \gamma & 1-\beta \\ & & & & \dots \end{array} \tag{17}$$

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 $y \le \max\{x, 1-x\}$ 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 $x \le \frac{1}{2}$ and to bet $1-x$ when you have $x \ge \frac{1}{2}$. (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 $0 \le j \le r$, let

$$A_{j,r} = \{2^{i_1} + \cdots + 2^{i_j} : r-1 \ge i_1 > \cdots > i_j \ge 0\} \tag{27}$$

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 $n \sim [a_1,\dots,a_{2v+1}]$ 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 $n = [[a_1,\dots,a_t]]$ is defined recursively by:

$$[[a]] = 2^a - 1; \qquad [[a_1,\dots,a_t]] = 2^{a_1+\cdots+a_t} - [[a_2,\dots,a_t]] \tag{42}$$

If $t = 2v+1$ is odd, then $[[a_1,\dots,a_{2v+1}]] \sim [a_1,\dots,a_{2v+1}]$, but t could be even.

Theorem 1. *Using the preceding notations, if $n = [[a_1,\dots,a_t]]$ and $r = \sum t_i$, then*

$$\frac{S(n;\lambda)}{S(2^r-n;\lambda)} = \frac{S([[a_1,\dots,a_t]];\lambda)}{S([[a_2,\dots,a_t]];\lambda)} = (\lambda)_{a_1} + \cfrac{\lambda^{a_1}}{(\lambda)_{a_2} + \cfrac{\lambda^{a_2}}{\cdots + \cfrac{\lambda^{a_{t-1}}}{(\lambda)_{a_t}}}}. \tag{43}$$

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,

$$\frac{S(2^{a+b+c}-2^{b+c}+2^c-1;\lambda)}{S(2^{a+b+c}-2^{b+c}+2^c;\lambda)} = \frac{(\lambda)_c S(2^{a+b}-2^b+1;\lambda)+\lambda^b(\lambda)_a}{\lambda^c S(2^{a+b}-2^b+1;\lambda)}$$

$$= \lambda^{-c}(\lambda)_c + \cfrac{\lambda^{-c}}{\cfrac{S(2^{a+b}-2^b+1;\lambda)}{\lambda^b(\lambda)_a}} = \tag{55}$$

$$\lambda^{-c}(\lambda)_c + \cfrac{\lambda^{-c}}{\cfrac{(\lambda)_b(\lambda)_a+\lambda^a}{\lambda^b(\lambda)_a}} = \lambda^{-c}(\lambda)_c + \cfrac{\lambda^{-c}}{\lambda^{-b}(\lambda)_b + \cfrac{\lambda^{-b}}{\lambda^{-a}(\lambda)_a}}\,.$$

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

$$\zeta_n = e^{\frac{2\pi i}{n}}, \tag{61}$$

$$\begin{gathered} S(t_n;\lambda) = 0 \iff \left(\frac{1+\sqrt{1+4\lambda}}{2}\right)^n = \left(\frac{1+\sqrt{1-4\lambda}}{2}\right)^n \\ \iff 1 + \sqrt{1+4\lambda} = \zeta_n^k(1 - \sqrt{1+4\lambda}) \iff \sqrt{1+4\lambda} = \frac{\zeta_n^k - 1}{\zeta_n^k + 1}, \end{gathered} \tag{62}$$

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

$$n \sim [1, b_1, 1, \dots, b_t, 1] :=]b_1, \dots, b_t[. \tag{71}$$

Put recursively,

$$]b[= 2^{b+1} + 1; \qquad]b_1, \dots, b_t[= 2^{b_t+1}]b_1, \dots, b_{t-1}[+ 1. \tag{72}$$

Theorem 10. *In the preceding notation, we have*

$$w(]b_1, \dots, b_t[) = p_t(-b_1, \dots, b_t) = (-1)^t p_t(b_1, \dots, b_t). \tag{73}$$

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.