Elliptic curves and related sequences

Christine S. Swart

Royal Holloway and Bedford New College
University of London

Thesis submitted to
The University of London
for the degree of
Doctor of Philosophy
2003.

To my family

Not all those who wander are lost.

Abstract

A *Somos* 4 *sequence* is a sequence (h_n) of rational numbers defined by the quadratic recursion $h_{m+2}\,h_{m-2} = \lambda_1\,h_{m+1}\,h_{m-1} + \lambda_2\,h_m^2$ for all $m \in \mathbb{Z}$ for some rational constants λ_1, λ_2. *Elliptic divisibility sequences* or *EDSs* are an important special case where $\lambda_1 = h_2^2$, $\lambda_2 = -h_1\,h_3$, the h_n are integers and h_n divides h_m whenever n divides m. Somos (4) is the particular Somos 4 sequence whose coefficients λ_i and initial values are all 1. In this thesis we study the properties of EDSs and Somos 4 sequences reduced modulo a prime power p^r.

In chapter 2 we collect some results from number theory, and in chapter 3 we give a brief introduction to elliptic curves.

In chapter 4 we introduce elliptic divisibility sequences, describe their relationship with elliptic curves, and outline what is known about the properties of an EDS modulo a prime power p^r (work by Morgan Ward and Rachel Shipsey).

In chapter 5 we extend the EDS "symmetry formulae" of Ward and Shipsey to higher powers of p. We use this to find the period of $(h_n \bmod p^r)$ in terms of the period of $(h_n \bmod p)$, confirming a conjecture by Shipsey for the $r = 2$ case.

In chapter 6 we give an introduction to Somos 4 sequences and list several conjectures by Raphael Robinson on the modulo p^r periodicity properties of Somos(4). In chapter 7 we describe a recent result by Nelson Stephens relating a Somos 4 sequence (h_n) to the sequence of points $Q + [n]P$ on an elliptic curve. We use this to prove that most of Robinson's conjectures on the pattern of zeroes hold in the sequence $(h_n \bmod p^r)$.

In chapter 8 we consider prime powers p^r dividing some term of a given Somos 4 sequence (h_n), and we find conditions under which Robinson's periodicity conjectures hold for $(h_n \bmod p^r)$ (although we do not always know if Somos(4) satisfies them). We do this by defining an equivalent sequence (ℓ_n), finding an EDS congruent to (ℓ_n) modulo p^r and using our EDS results from chapter 4.

Finally, in chapter 9 we use elliptic curves to prove that a weakened version of Robinson's periodicity conjecture holds for prime powers which are coprime to λ_1 and to every term of a given Somos 4 sequence.

Contents

Chapter 1

Introduction

In this thesis we study elliptic divisibility sequences and Somos 4 sequences, focusing on their properties when reduced modulo a prime power p^r.

1.1 Elliptic divisibility sequences

An *elliptic sequence* is a sequence (h_n) of rational numbers satisfying the quadratic recurrence relation

$$h_{m+n}\,h_{m-n} = h_{m+1}\,h_{m-1}\,h_n^2 - h_{n+1}\,h_{n-1}\,h_m^2 \quad \text{for all } m, n \in \mathbb{Z}. \tag{1.1}$$

An *elliptic divisibility sequence* (or *EDS*) is an integer elliptic sequence with the divisibility property that h_n divides h_m whenever n divides m. It is easy to prove that all elliptic sequences have $h_0 = 0$, $h_1 = -h_{-1} = \pm 1$, and that $h_{-n} = -h_n$ for all $n \in \mathbb{Z}$.

EDSs are a generalisation of a class of integer divisibility sequences called Lucas sequences. These, in common with all divisibility sequences studied previously (like the Fibonacci sequence $F_n = F_{n-1} + F_{n-2}$ and the Mersenne sequence $M_n = 2^n - 1 = 3\,M_{n-1} - 2\,M_{n-2}$), satisfy a linear recurrence relation. EDSs were interesting in being the first non-linear divisibility sequences to be studied.

The elliptic sequence equation (1.1) is the same recurrence relation satisfied by the division polynomials of an elliptic curve. Morgan Ward proved in [30] that for every elliptic sequence (h_n) there exists an elliptic curve E and a rational

point $P = (x, y)$ such that

$$h_n = \psi_n(x, y) \quad \text{for all} \quad n \in \mathbb{Z},$$

where ψ_n is the nth division polynomial of E. He then used this relationship to study the properties of an EDS (h_n) reduced modulo a prime p. He first proved that, unless p divides every term h_n with $|n| \geq 3$, the multiples of p are regularly spaced in (h_n), i.e.,

$$h_n \equiv 0 \bmod p \quad \text{if and only if} \quad n \equiv 0 \bmod N$$

for some positive integer N. N is called the *gap* of p, and p is said to be *regular* in (h_n). It can be shown that some multiple of N lies within the "Hasse bound" of $p + 1$.

Ward then looked at the subsequence obtained from $(h_n \bmod p)$ by taking every Nth term from some term h_t, and found that if $N > 3$ then it always has the simple form (his "symmetry formula")

$$h_{t+sN} \equiv \left(c^t\right)^s (-b)^{s^2} h_t \bmod p \quad \text{for all} \quad s \in \mathbb{Z}, \tag{1.2}$$

where b and c are some constants which can be calculated from the sequence modulo p. Finally, he proved that $(h_n \bmod p)$ is periodic with period $N\tau$, where τ is a constant dividing $p - 1$ that can be calculated from the sequence modulo p.

In a later paper [31] Ward turned his attention to higher powers of the prime p. He first considered the sequence obtained from an EDS (h_n) by taking every kth term from h_0 and dividing by h_k, and proved that it is also an EDS. It follows that if $p \nmid \gcd(h_3, h_4)$ then for all $r \in \mathbb{N}$

$$h_n \equiv 0 \bmod p^r \quad \Leftrightarrow \quad n \equiv 0 \bmod N_r,$$

i.e., p^r is regular in (h_n) with gap N_r. In the special case where $k = N_r$, $p > 3$ and $N_1 > 3$ Ward proved that the EDS $\left(\frac{h_{sN_r}}{h_{N_r}}\right)$ always has the simple form

$$\frac{h_{sN_r}}{h_{N_r}} \equiv s\ (-h_{N_r+1}\, h_{N_r-1})^{\frac{1}{2}(s^2-1)} \bmod p^{2r} \quad \text{for all} \quad s \in \mathbb{Z}. \tag{1.3}$$

(Notice that if s is even then this determines h_{sN_r} only up to sign.) Finally, he used (1.3) to prove that if $p > 3$ and p^w is the highest power of p dividing h_{N_1} then

$$N_r = p\,N_{r-1} \quad \text{for all} \quad r > w.$$

Rachel Shipsey continued the study of EDSs in her PhD thesis [23], seeing EDSs as a convenient way to study elliptic curves, with possible application to elliptic curve cryptography. In particular, she used EDS symmetry results in $(h_n \bmod p)$ and $(h_n \bmod p^2)$ to to find elegant alternative attacks on the elliptic curve discrete log problem in the MOV [16] and anomalous curve [26] cases. She proved that if $N_1 > 3$ then

$$h_{s\,N_1} \equiv s(-b)^{s^2-1}\,h_{N_1} \bmod p^2 \quad \text{for all} \quad s \in \mathbb{Z}. \tag{1.4}$$

(Note that this determines the sign of the square root of $-h_{N_r+1}\,h_{N_r-1} \bmod p^{3r}$ in (1.3).) She also conjectured that the period of $(h_n \bmod p^2)$ is $\tau_1\,N_2$; that is, the same as the period of $(h_n \bmod p)$ if $p^2 \mid h_{N_1}$, or p times this period if $p^2 \nmid h_{N_1}$.

Elliptic divisibility sequences have also been studied by Chudnovsky and Chudnovsky in [5] and extensively by Everest et al (see for example [7], [8], [9], [10] and [11]).

In chapter 5 we consider the symmetry and periodicity of EDSs reduced modulo a prime power p^r. We first prove that Ward's symmetry formula (1.2) and Shipsey's formula (1.4) hold modulo p^r if N is replaced by N_r and b and c by new constants b_r and c_r. We extend these results by finding a simple expression for the general term of the subsequence (h_{t+sN_r}) for a fixed term $h_t \not\equiv 0 \bmod p$ that holds modulo p^{3r} instead of just modulo p^r, and we use this to calculate c_r and b_r in terms of c_w and b_w, and hence in terms of the sequence $(h_n \bmod p^w)$.

We then prove that for $r \in \mathbb{N}$ the sequence $(h_n \bmod p^r)$ is periodic with period $\pi_r = \tau_r N_r$ for some constant τ_r that can be calculated from the sequence. Finally, we prove that $\pi_r = \pi_1$ for $r = 1, 2, \ldots, u$ for some $u \geq 1$, and $\pi_r = p\,\pi_{r-1}$ for $r > u$, confirming Shipsey's conjecture about π_2.

Our proofs of the above results are elementary; that is, they do not rely on the relationship between EDSs and elliptic curves but simply use the elliptic

sequence formula (1.1) directly.

1.2 Somos 4 sequences

A *Somos* 4 *sequence* is a sequence of rational numbers defined by the recursion

$$h_{m+2}\, h_{m-2} = \lambda_1\, h_{m+1}\, h_{m-1} + \lambda_2\, h_m^2 \quad \text{for all} \quad m \in \mathbb{Z}, \tag{1.5}$$

where λ_1 and λ_2 are given rational numbers (and the sequence terminates if we get a zero term). Elliptic sequences are an important special case with $\lambda_1 = h_2^2$, $\lambda_2 = -h_1\, h_3$, $h_0 = 0$ and $h_1 = \pm 1$. The other important special case which has been studied is the particular Somos 4 sequence called Somos (4) whose coefficients λ_1, λ_2 and four initial values are all 1.

Somos 4 sequences form part of a larger group of Somos k sequences introduced by Michael Somos in [27]. Initial interest in Somos k sequences was in the surprising fact that, if the initial values are all 1 and the coefficients integers, then for $4 \leq k \leq 7$ the Somos k recursion produces an integer sequence, despite the fact that computing each term involves dividing by another term. Some history of this problem is given by David Gale in [13] and [12].

It was proved independently by Nelson Stephens and Noam Elkies that Somos 4 sequences are related to the weighted Z- coordinates of the sequence of points $Q + [n]P$, $n \in \mathbb{Z}$, for two rational points Q and P on an elliptic curve. (This work is as yet unpublished, but see the mailing list on bilinear sequences maintained by Jim Propp [19].) A different approach to the same problem is taken by Andy Hone in [14]. Most recently, Somos sequences have been studied in connection with the number of matchings in planar graphs [3].

We are interested in the properties of Somos 4 sequences reduced modulo a prime power p^r, and particularly in how these change as r increases. This problem was posed by Raphael Robinson [21], who ran computer experiments on Somos(4) and made the following conjectures about its modulo p^r properties:

Robinson's conjectures:

1. If p is an odd prime dividing some term of Somos(4), then for all $r \in \mathbb{N}$, p^r also divides some term.

2. Every prime power dividing some term of Somos(4) is regular in Somos(4).

3. If p is a regular prime with gap N_1 in Somos(4) then some multiple of N_1 is close to p.

4. Let p be an odd prime with gap N_1 in Somos(4), and let p^w be the highest power of p such that all multiples of p in Somos(4) are divisible by p^w. Then for $r > w$, the gap of p^r in Somos(4) is $N_r = p\,N_{r-1}$.

5. Let (h_n) be Somos(4). Then for all odd primes p and $r \in \mathbb{N}$, the period of $(h_n \bmod p^r)$ is equal to p^{r-1} times the period of $(h_n \bmod p)$.

6. If (h_n) is Somos(4) and p is a regular prime with gap N_1 in (h_n), then the period π_1 of $(h_n \bmod p)$ is a multiple of N_1, and a divisor of $(p-1)\,N_1$.

These conjectured properties of Somos(4) are so similar to the properties of an EDS which we proved in chapter 5 that they prompt us to look for a generalisation to all Somos 4 sequences. We cannot generalise our EDS proofs directly, because Somos 4 sequences only satisfy the local recursion (1.5), while our proofs often use the elliptic sequence formula (1.1) with values of n other than 2. Instead, if p^r divides some term h_k of a Somos 4 sequence (h_n) and $N_1 > 4$, then in chapter 8 we find a mapping from $(h_n \bmod p^r)$ to $(Z_n \bmod p^r)$, where (Z_n) is an EDS. This allows us to use our EDS symmetry and periodicity results from chapter 5 to prove similar results for $(h_n \bmod p^r)$. In particular, we prove that Conjecture 5 holds for (h_n) in the case where some term is divisible by p but not by p^2, and a slightly more complicated result holds otherwise. We also prove that π_1 is a multiple of N_1 and a divisor of $2\,(p-1)\,N_1$, and find the conditions under which π_1 divides $(p-1)\,N_1$. It remains open whether Somos(4) satisfies these conditions and has $w = 1$ for every prime dividing some term, but we give examples of Somos 4 sequences which do not.

In chapter 7 we use Stephens' result on the relationship between a Somos 4 sequence and the sequence of points $Q + [n]P$ on an elliptic curve to prove Robinson's Conjectures 2–4 on the pattern of zeroes in Somos(4) and extend them to a general Somos 4 sequence S. We also prove that Conjecture 1 holds for a general Somos 4 sequence S unless all multiples of p in S are divisible by exactly the same power of p. We do not know if this happens for any prime p other than 2 in Somos(4), but we give an example of a Somos 4 sequence S and an odd prime p for which it does.

The only one of Robinson's conjectures that applies to primes not dividing any term of Somos(4) is Conjecture 5. In chapter 9 we again use elliptic curves to prove that if p is such a prime then $\pi_r = \pi_w$ for $r \leq w$, and $\pi_{r+1} = p$ for all $r > w$, although we have not yet proved that $\pi_{w+1} = p\,\pi_w$. This result holds for any prime p coprime to λ_1 and all terms of a general Somos 4 sequence S.

1.3 Thesis outline

In chapter 2 we collect the results about $\mathbb{Z}_{p^r}^*$, the multiplicative group of integers modulo a prime power p^r, that we will need in the rest of the thesis.

In chapter 3 we give a brief introduction to elliptic curves, and to the division polynomials in particular.

In chapter 4 we introduce elliptic divisibility sequences, give some basic properties, and describe their relationship with elliptic curves. We outline what is known about the properties of an EDS (h_n) reduced modulo a prime power p^r (work by Morgan Ward and Rachel Shipsey).

In chapter 5 we extend the EDS "symmetry formulae" of Ward and Shipsey to higher powers of p. We use this to find the period of $(h_n \bmod p^r)$ in terms of the period of $(h_n \bmod p)$, confirming a conjecture by Shipsey for $r = 2$.

In chapter 6 we give an introduction to Somos 4 sequences. We list Robinson's conjectures on the properties of Somos(4) modulo p^r, and outline the results we have obtained in trying to prove them for all Somos 4 sequences.

In chapter 7 we describe Stephens' result relating a Somos 4 sequence S to the sequence of points $Q + [n]P$ on an elliptic curve. We use this to prove that Robinson's conjectures on the pattern of zeroes modulo p^r hold for S.

In chapter 8 we consider prime powers p^r which divide some term of a given Somos 4 sequence (h_n) (where p has gap $N_1 > 4$) and we find simple conditions under which Robinson's periodicity conjectures hold for (h_n) (although we do not yet know if Somos(4) satisfies these conditions). We do this by defining an equivalent sequence (ℓ_n), finding an EDS congruent to (ℓ_n) modulo p^r and using our EDS results from chapter 4.

In chapter 9 we use the relationship between Somos 4 sequences and elliptic curves to prove that a weakened version of Robinson's periodicity conjecture 5 holds for prime powers which do not divide λ_1 or any term of a given Somos 4 sequence.

Chapter 2

Some number theory

In this chapter we establish some conventions and collect the results from number theory that we will need in the rest of the thesis.

2.1 Definitions and notation

We denote the field of rational numbers by $\mathbb{Q}$, the ring of integers by $\mathbb{Z}$, and the ring of integers modulo an integer m by $\mathbb{Z}_m$. If (h_n) is a sequence of rational numbers then we denote by $(h_n \bmod m)$ the sequence of elements of $\mathbb{Z}_m$ whose nth term is $h_n \bmod m$.

Let x be a rational number written in lowest terms, i.e., $x = \frac{a}{b}$ where a and b are coprime integers. When we refer to "the numerator" or "the denominator" of x we will always mean a and b. If m is an integer coprime to b then by $x \bmod m$ we mean $a\,b^{-1} \bmod m$. The notation $m \mid x$ means that $a\,b^{-1} \equiv 0 \bmod m$, and $m \nmid x$ means that $a\,b^{-1} \not\equiv 0 \bmod m$. If p is a prime and x an integer then the notation $p^r \parallel x$ means that p^r is the highest power of p dividing x.

Let S be a set of integers. Then x is called an *S-integer* if the denominator of x is coprime to every integer in S. When $S = \{m\}$ we just refer to $\{m\}$-integers as m-integers.

If R is a ring then R^* denotes the multiplicative group of invertible elements of R, and R^+ the additive group of R. If K is a field then we denote its algebraic closure by $\overline{K}$. The *projective plane* $\mathbb{P}^2(K)$ over a field K is the set of equivalence

classes of the relation $\sim$ acting on $K^3 \setminus \{(0,0,0)\}$, where $(x_1, y_1, z_1) \sim (x_2, y_2, z_2)$ if and only if there exists $u \in K^*$ such that $x_1 = u\,x_2$, $y_1 = u\,y_2$ and $z_1 = u\,z_2$. We denote the equivalence class containing (x, y, z) by $(x : y : z)$. A point $(x : y : z)$ in the projective plane $\mathbb{P}^2(\mathbb{Q})$ over the rationals is said to be *in lowest terms* if x, y, z are integers and $\gcd(x, y, z) = 1$.

Finally, we denote the cyclic group with n elements by C_n. If g is an element of a group G, then the cyclic subgroup of G generated by g is $\langle g \rangle$.

2.2 The group $\mathbb{Z}^*_{p^r}$

Let p^r be a prime power. We need some results about $\mathbb{Z}^*_{p^r}$, the multiplicative group of invertible integers modulo p^r. Most of these follow from the fact that $\mathbb{Z}^*_{p^r}$ is a cyclic group if p is odd, or the product of two cyclic groups if $p = 2$. (See for example [22, page 91] for a proof.)

Theorem 2.2.1. *Let $r \in \mathbb{N}$. If p is an odd prime then $\mathbb{Z}^*_{p^r}$ is a cyclic group of order $p^{r-1}(p-1)$. For $r \geq 2$, $\mathbb{Z}^*_{2^r}$ is the direct product of two cyclic groups of order 2 and 2^{r-2} respectively.*

So for $r \geq 2$, there is an element h of order 2^{r-2} in $\mathbb{Z}^*_{2^r}$ such that each $x \in \mathbb{Z}^*_{2^r}$ can be written uniquely as $\pm h^s$ for some $s \in \{0, 1, \ldots, 2^{r-2}\}$.

Theorem 2.2.1 has several immediate consequences:

Lemma 2.2.2. *Let p be an odd prime, $r \in \mathbb{N}$ and g a generator of $\mathbb{Z}^*_{p^r}$. Then the element g^s has order*

$$\frac{p^{r-1}(p-1)}{\gcd\left(p^{r-1}(p-1),\, s\right)}$$

in $\mathbb{Z}^*_{p^r}$, and g^s is a quadratic residue modulo p^r if and only if s is even.

Lemma 2.2.3. *Let p be an odd prime, $r \in \mathbb{N}$ and g a generator of $\mathbb{Z}^*_{p^r}$. Then g is also a generator for $\mathbb{Z}^*_{p^k}$ for all $k < r$.*

Let $(-1, h)$ be a generating pair for $\mathbb{Z}^*_{2^r}$. Then it is also a generating pair for $\mathbb{Z}^*_{2^k}$ for all $k < r$.

Theorem 2.2.4. *Let $r \in \mathbb{N}$. If p is an odd prime, then*

$$x^2 \equiv 1 \bmod p^r \;\Leftrightarrow\; x \equiv \pm 1 \bmod p^r.$$

For $p = 2$ we have

$$\begin{aligned} x^2 \equiv 1 \bmod 2^r \;&\Leftrightarrow\; x \equiv \pm 1 \bmod 2^r \quad \textit{or} \quad x \equiv \pm 1 + 2^{r-1} \bmod 2^r \\ &\Leftrightarrow\; x \equiv \pm 1 \bmod 2^{r-1}. \end{aligned}$$

So 1 has two square roots in $\mathbb{Z}^*_{p^r}$ if p is an odd prime, and four square roots in $\mathbb{Z}^*_{2^r}$ for each $r \geq 3$ (namely, ± 1 and $\pm 1 + 2^{r-1}$).

Theorem 2.2.5. *Let x and z be integers, let p be a prime not dividing either, and let $a, b \in \mathbb{N}$. If p is odd then*

$$x \equiv z \bmod p^a \;\Leftrightarrow\; x^{p^b} \equiv z^{p^b} \bmod p^{a+b}.$$

For $p = 2$ we have

$$\begin{aligned} x^{2^b} \equiv z^{2^b} \bmod 2^{a+b} \;&\Leftrightarrow\; x^2 \equiv z^2 \bmod 2^{a+1} \\ &\Leftrightarrow\; x \equiv \pm z \bmod 2^a. \end{aligned}$$

Proof: Let $y = x\, z^{-1} \bmod p^{a+b}$. If p is odd, let g be a generator of $Z^*_{p^{a+b}}$; say $y \equiv g^s \bmod p^{a+b}$. Then, since g is also a generator for $\mathbb{Z}^*_{p^a}$,

$$\begin{aligned} y \equiv 1 \bmod p^a \;&\Leftrightarrow\; g^s \equiv 1 \bmod p^a \\ &\Leftrightarrow\; p^{a-1}(p-1) \mid s \\ &\Leftrightarrow\; p^{a+b-1}(p-1) \mid p^b\, s \\ &\Leftrightarrow\; (g^s)^{p^b} \equiv 1 \bmod p^{a+b} \\ &\Leftrightarrow\; y^{p^b} \equiv 1 \bmod p^{a+b}. \end{aligned}$$

For the $p = 2$ case, let $(-1, h)$ be a generating pair for $\mathbb{Z}^*_{2^{a+b}}$ (and hence for $\mathbb{Z}^*_{2^{a+1}}$). So y can be written uniquely as $y \equiv (-1)^t\, h^s \bmod 2^{a+b}$ for some $t \in \{0, 1\}$ and $s \in \{0, 1, \ldots, 2^{a+b-2} - 1\}$. It follows that

$$y^2 \equiv 1 \bmod 2^{a+1} \Leftrightarrow (-1)^{2t}\, h^{2s} \equiv 1 \bmod 2^{a+1} \Leftrightarrow h^{2s} \equiv 1 \bmod 2^{a+1}.$$

Since h has order 2^{a-1} in $\mathbb{Z}^*_{2^{a+1}}$ and order 2^{a+b-2} in $\mathbb{Z}^*_{2^{a+b}}$, the right hand side holds if and only if

$$\begin{aligned} 2^{a-1} \mid 2s &\Leftrightarrow 2^{a+b-2} \mid 2^b\, s \\ &\Leftrightarrow (h^s)^{2^b} \equiv 1 \bmod 2^{a+b} \\ &\Leftrightarrow \left((-1)^t\, h^s\right)^{2^b} \equiv 1 \bmod 2^{a+b} \\ &\Leftrightarrow y^{2^b} \equiv 1 \bmod 2^{a+b}. \end{aligned}$$

The result now follows from Theorem 2.2.4. □

So for odd p there are precisely p pth roots of x^p modulo p^r, and they are all the same modulo p^{r-1}. For $p = 2$ and $r \geq 3$ there are four square roots of x^2, and they are not all the same modulo 2^{r-1}.

As a consequence of Theorem 2.2.5 we have

Theorem 2.2.6. *Let p be an odd prime, let $a \in \mathbb{N}$, and let y be an integer with order n in $\mathbb{Z}^*_{p^a}$. Then for any $b \in \mathbb{N}$, y^{p^b} has order n in $\mathbb{Z}^*_{p^{a+b}}$.*

Proof: By Theorem 2.2.5 with $x = y^t$ and $z = 1$,

$$\begin{aligned} y^t \equiv 1 \bmod p^a &\Leftrightarrow (y^t)^{p^b} \equiv 1^{p^b} \bmod p^{a+b} \\ &\Leftrightarrow (y^{p^b})^t \equiv 1 \bmod p^{a+b}. \end{aligned}$$

The result follows. □

Theorem 2.2.7. *Let p be a prime. Then for any integers x and $n \geq 0$,*

$$x^{p^{r-1+n}} \equiv x^{p^{r-1}} \bmod p^r.$$

Proof: The result holds trivially for $n = 0$, so we may assume it holds for $n = 0, 1, \ldots, m-1$ for some $m \geq 1$. We prove it then holds for m.

Since $\mathbb{Z}^*_{p^r}$ has $p^r - p^{r-1}$ elements,

$$x^{p^r} \equiv x^{p^{r-1}} \bmod p^r.$$

By the induction hypothesis,

$$\begin{aligned} x^{p^{r-1+m}} &= x^{p^{r-1+(m-1)}} \cdot x^p \\ &\equiv x^{p^{r-1}} \cdot x^p \bmod p^r \\ &\equiv x^{p^r} \bmod p^r \\ &\equiv x^{p^{r-1}} \bmod p^r. \end{aligned}$$

The result now follows by induction. □

Finally, we have

Theorem 2.2.8. *If p is an odd prime, $r \geq 2$ and $B_0, \ldots, B_{p-1}$ are p p-integers such that*

$$B_i \not\equiv B_j \bmod p^r \quad \text{for any } i, j$$

but

$$B_i \equiv B \bmod p^{r-1} \quad \text{for } i = 0, \ldots, p-1$$

(i.e., the B_i are distinct modulo p^r but the same modulo p^{r-1}) then

$$\prod_{i=0}^{p-1} B_i \equiv B^p \bmod p^r.$$

Proof: Let $B_i \equiv B + s_i\, p^{r-1} \bmod p^r$ for $i = 0, 1, \ldots, p-1$, where $s_i \in \{0, 1, \ldots, p-1\}$. Since the B_i are distinct modulo p^r, the s_i are distinct modulo p; i.e., $\{s_0, \ldots, s_{p-1}\} = \{0, \ldots, p-1\}$. Hence $\sum_{i=0}^{p-1} s_i = 1 + 2 + \ldots + (p-1) = \frac{1}{2}p(p-1)$, which is divisible by p since p is odd.

Using the binomial theorem, we get

$$\begin{aligned} \prod_{i=0}^{p-1} B_i &= \prod_{i=0}^{p-1} \left(B + s_i\, p^{r-1}\right) \\ &\equiv B^p + B^{p-1} \left(\sum_{i=0}^{p-1} s_i\right) p^{r-1} \bmod p^r \\ &\equiv B^p \bmod p^r. \end{aligned}$$

□

Chapter 3

Introduction to elliptic curves

In this chapter we present the basic concepts from the theory of elliptic curves required for this thesis, mostly following [2] and [17]. For a fuller treatment see [24] or [25], and for an elementary introduction see [4].

3.1 Definitions

Let K be a field, $\overline{K}$ its algebraic closure and K^* its multiplicative group. An *elliptic curve E over K* is defined as the set of all solutions in the projective plane $\mathbb{P}^2(\overline{K})$ of a homogeneous *Weierstrass equation* of the form

$$E : Y^2\,Z + a_1\,X\,Y\,Z + a_3\,Y\,Z^2 = X^3 + a_2\,X^2\,Z + a_4\,X\,Z^2 + a_6\,Z^3, \qquad (3.1)$$

where the coefficients a_i are in K. The curve E is said to be *non-singular* if, when it is written in the form $F(X, Y, Z) = 0$, at least one of the three partial derivatives $\frac{\partial F}{\partial X}$, $\frac{\partial F}{\partial Y}$, $\frac{\partial F}{\partial Z}$ is non-zero at each projective point $(X : Y : Z)$ on the curve. If all three partial derivatives vanish at some point P, then P is called a *singular point* and E is said to be *singular*. (This definition (also used in [4]) is slightly different from the usual one, which requires elliptic curves to be non-singular; this is more convenient for our purposes because we will usually be considering singular and non-singular curves simultaneously.)

We use the notation E/K to indicate that E is defined over the field K, i.e., that all its coefficients a_i are in K. (Of course then E/K' for every extension

field K' of K.)

There is exactly one point in E with Z-coordinate equal to 0, namely $(0:1:0)$. We call this point the *point at infinity* and denote it by $\mathcal{O}$. The points in $E \setminus \{\mathcal{O}\}$ are called *finite points*.

A point $(X:Y:Z)$ on the curve is *K-rational* if $(X,Y,Z) = \theta\,(X',Y',Z')$ for some $\theta \in \overline{K}$ and $(X',Y',Z') \in K^3 \setminus \{(0,0,0)\}$; i.e., up to projective equivalence, the coordinates of the point are in K. The set of K-rational points on E is denoted by $E(K)$, and it includes the point at infinity $\mathcal{O} = (0:1:0)$. (So E is $E(\overline{K})$.)

The set of non-singular K-rational points on E (including the point at infinity $\mathcal{O}$) is called the *non-singular* part of $E(K)$ and denoted by $E_{ns}(K)$.

We can also write the Weierstrass equation for an elliptic curve using non-homogeneous (affine) coordinates $x = \frac{X}{Z}$, $y = \frac{Y}{Z}$ as

$$E : y^2 + a_1\,x\,y + a_3\,y = x^3 + a_2\,x^2 + a_4\,x + a_6. \tag{3.2}$$

The K-rational points in the affine case are the solutions to E in $K \times K$, together with the point at infinity $\mathcal{O}$. We denote the x- and y-coordinates of a point P by $x(P)$ and $y(P)$ respectively.

Given an elliptic curve defined by equation (3.2) or (3.1), it is useful to define the following constants:

$$\left.\begin{aligned} b_2 &= a_1^2 + 4a_2 \\ b_4 &= a_1\,a_3 + 2a_4 \\ b_6 &= a_3^2 + 4a_6 \\ b_8 &= a_1^2\,a_6 - a_1\,a_3\,a_4 + 4a_2\,a_6 + a_2\,a_3^2 - a_4^2. \end{aligned}\right\} \tag{3.3}$$

Note that

$$b_4^2 + 4\,b_8 = b_2\,b_6. \tag{3.4}$$

The *discriminant* of the curve is defined as

$$\Delta = -b_2^2\,b_8 - 8b_4^3 - 27b_6^2 + 9b_2\,b_4\,b_6. \tag{3.5}$$

The curve is non-singular if and only if $\Delta \neq 0$.

3.2 The group law

The use of elliptic curves in cryptography depends on the fact that the set of non-singular K-rational points on an elliptic curve E over K forms an abelian group:

Theorem 3.2.1. *Let E be an elliptic curve over a field K. Then $E_{ns}(K)$ forms an abelian group with the point at infinity $\mathcal{O}$ as the zero.*

The group law is given by the following algebraic formulae: let

$$E : y^2 + a_1\,x\,y + a_3\,y = x^3 + a_2\,x^2 + a_4\,x + a_6,$$

be an elliptic curve, and let $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$ be finite points on E. Then

$$-P_1 = (x_1, -y_1 - a_1\,x_1 - a_3), \tag{3.6}$$

i.e., the inverse $-P_1$ of P_1 in E is the other point with the same x-coordinate.

If $P_2 = -P_1$ then $P_1 + P_2 = \mathcal{O}$. Otherwise, set

$$\lambda = \begin{cases} \dfrac{y_2 - y_1}{x_2 - x_1} & \text{if } P_2 \neq P_1 \\ \dfrac{3x_1^2 + 2a_2\,x_1 + a_4 - a_1\,y_1}{2y_1 + a_1\,x_1 + a_3} & \text{if } P_2 = P_1, \end{cases} \tag{3.7a}$$

and set

$$\mu = y_1 - \lambda\,x_1. \tag{3.7b}$$

If

$$P_3 = (x_3, y_3) = P_1 + P_2 \neq \mathcal{O}$$

(i.e., if $P_2 \neq -P_1$) then x_3 and y_3 are given by the formulae

$$\left.\begin{aligned} x_3 &= \lambda^2 + a_1\,\lambda - a_2 - x_1 - x_2 \\ y_3 &= -(\lambda + a_1)x_3 - \mu - a_3. \end{aligned}\right\} \tag{3.7c}$$

Remark: Note that if P_1 is a singular point then λ is not defined when $P_2 = P_1$, so the group law does not work. If P_1 is non-singular then λ is defined, since $2y + a_1\,x_1 + a_3 = 0$ only if $P_1 = -P_1$ and we have excluded the case $P_2 = -P_1$.

The above addition law has a geometric interpretation: if $P_1, P_2 \in E_{ns}(K)$ with $P_2 \neq -P_1$ then a straight line between P_1 and P_2 (or if $P_1 = P_2$, the tangent to the curve at P_1) will cut the curve in one more point; this point is $-(P_1 + P_2)$. Note that the tangent line is well-defined at every non-singular point.

Definition: For a positive integer n we let $[n]$ denote the *multiplication-by-n* map from $E(\overline{K})$ to itself. This map takes a point P to $P + \ldots + P$ (P added to itself $n-1$ times). The notation $[n]$ is extended to $n \leq 0$ by defining $[0]P = \mathcal{O}$ and $[-n]P = -([n]P)$. If $[n]P = \mathcal{O}$ for some minimal positive integer n, then n is called the *order* of P in $E(K)$.

Definition: For a non-negative integer n, the set of *n-torsion points* of E, denoted by $E[n]$, is defined by

$$E[n] = \{P \in E(\overline{K}) \mid [n]P = \mathcal{O}\}.$$

(Notice that $E[n]$ is defined over $E(\overline{K})$, not $E(K)$.) It is easy to see that $E[n]$ is a subgroup of E. By definition, $\mathcal{O} \in E[n]$ for all n.

3.3 Birational equivalence

Definition: Two elliptic curves E/K and $\bar{E}/K$ given by the equations

$$E : y^2 + a_1\,x\,y + a_3\,y = x^3 + a_2\,x^2 + a_4\,x + a_6$$

and

$$\bar{E} : \bar{y}^2 + \bar{a}_1\,\bar{x}\,\bar{y} + \bar{a}_3\,\bar{y} = \bar{x}^3 + \bar{a}_2\,\bar{x}^2 + \bar{a}_4\,\bar{x} + \bar{a}_6$$

are said to be *birationally equivalent* over K, denoted $E/K \cong \bar{E}/K$, if and only if there exist constants $r, s, t \in K$ and $u \in K^*$ such that the change of variables

$$x = u^2\,\bar{x} + r \quad \text{and} \quad y = u^3\,\bar{y} + u^2\,s\,\bar{x} + t \tag{3.8}$$

transforms E into $\bar{E}$. The change of variables (3.8) is called an *admissible change of variables*.

Clearly, this transformation is reversible, and its inverse

$$\bar{x} = u^{-2}(x - r) \quad \text{and} \quad \bar{y} = u^{-3}(y - sx - t + rs)$$

also defines an admissible change of variables that transforms $\bar{E}$ into E. So (3.8) defines a bijection between E and $\bar{E}$, and the relationship of birational equivalence over K is an equivalence relation. In fact this bijection restricted to $E(K)$ is a group isomorphism between $E(K)$ and $\bar{E}(K)$. (The converse is not true; $E(K)$ and $\bar{E}(K)$ can be isomorphic groups without E and $\bar{E}$ being birationally equivalent curves over K.) Notice that birational equivalence is defined relative to the field K: two curves might be birationally equivalent over an extension field of K but not over K.

Theorem 3.3.1. *Let E be an elliptic curve over K given by equation* (3.2)*. If* $\mathrm{char}(K) \neq 2$*, then the admissible change of variables given by*

$$x = \bar{x} \quad \textit{and} \quad y = u^3\,\bar{y} - \frac{a_1}{2}\,\bar{x} - \frac{a_3}{2},$$

where $u = \pm 1$, transforms E into the birationally equivalent curve

$$\bar{E} : \bar{y}^2 = \bar{x}^3 + \frac{b_2}{4}\,\bar{x}^2 + \frac{b_4}{2}\,\bar{x} + \frac{b_6}{4}. \tag{3.9}$$

We refer to equation (3.9) as the "b-form" of the curve E.

Theorem 3.3.2. *Let E be an elliptic curve over K given by equation* (3.2)*. If* $\mathrm{char}(K) \neq 2, 3$*, then the admissible change of variables given by*

$$x = \bar{x} - \frac{b_2}{12} \quad \textit{and} \quad y = \bar{y} - \frac{a_1}{2}\left(\bar{x} - \frac{b_2}{12}\right) - \frac{a_3}{2}$$

transforms E into the birationally equivalent curve

$$\bar{E} : \bar{y}^2 = \bar{x}^3 + c_4\,\bar{x} + c_6 \tag{3.10}$$

where

$$c_4 = -\frac{b_2^2}{48} + \frac{b_4}{2}, \quad \textit{and} \quad c_6 = \frac{b_2^3}{864} - \frac{b_2\,b_4}{24} + \frac{b_6}{4}.$$

Equation (3.10) is called the *short Weierstrass form* or the "c-form" of the curve E.

The following result relates the notion of birational equivalence of elliptic curves to the coefficients of their defining equations.

Theorem 3.3.3. *Two elliptic curves E/K and $\bar{E}/K$ of the form* (3.2) *are birationally equivalent over K if and only if there exist $u, r, s, t \in K$, $u \neq 0$ that satisfy*

$$\left.\begin{aligned} u\,\bar{a}_1 &= a_1 + 2s \\ u^2\,\bar{a}_2 &= a_2 - sa_1 + 3r - s^2 \\ u^3\,\bar{a}_3 &= a_3 + ra_1 + 2t \\ u^4\,\bar{a}_4 &= a_4 - sa_3 + 2ra_2 - (t + rs)a_1 + 3r^2 - 2st \\ u^6\,\bar{a}_6 &= a_6 + ra_4 + r^2a_2 + r^3 - ta_3 - t^2 - rta_1. \end{aligned}\right\} \tag{3.11}$$

Equivalently, by writing both curves in the b-form and using the fact that birational equivalence over a field K is an equivalence relation (and proving the characteristic 2 case separately), we can prove

Theorem 3.3.4. *Two elliptic curves E/K and $\bar{E}/K$ of the form* (3.2) *are birationally equivalent over K under the admissible change of variables* (3.8) *for some $s, t \in \mathbb{Q}$ if and only if*

$$\left.\begin{aligned} u^2\,\bar{b}_2 &= b_2 + 12r \\ u^4\,\bar{b}_4 &= b_4 + r\,b_2 + 6r^2 \\ u^6\,\bar{b}_6 &= b_6 + 2r\,b_4 + r^2\,b_2 + 4r^3 \\ u^8\,\bar{b}_8 &= b_8 + 3r\,b_6 + 3r^2\,b_4 + r^3\,b_2 + 3r^4. \end{aligned}\right\} \tag{3.12}$$

We have the following corollary:

Theorem 3.3.5. *If* $\mathrm{char}(K) \neq 2$, *then two elliptic curves E/K and $\bar{E}/K$ have the same b-form, i.e.,*

$$b_i = \bar{b}_i \quad \text{for } i = 2, 4, 6, 8$$

if and only if E and $\bar{E}$ are birationally equivalent over K under an admissible change of variables (3.8) *with $u = \pm 1$, $r = 0$ and $s, t \in K$.*

It turns out that if E is transformed to a birationally equivalent curve by (3.8) then the new discriminant depends only on the old discriminant and u, not on r, s or t:

Theorem 3.3.6. *Let E and $\bar{E}$ be elliptic curves which are birationally equivalent under the admissible change of variables* (3.8). *Then*

$$\Delta = u^{12}\,\bar{\Delta}.$$

3.4 Singular elliptic curves

Let E be a singular elliptic curve over K, given by the equation

$$F(x,y) = y^2 + a_1\,x\,y + a_3\,y - x^3 - a_2\,x^2 - a_4\,x - a_6 = 0.$$

Then it can be shown that E has precisely one singular point, $Q = (x_0, y_0)$. We make the change of variables $x = \bar{x} + x_0,\ y = \bar{y} + y_0$ to move Q to the origin. Since $\frac{\partial \bar{F}}{\partial \bar{x}}$ and $\frac{\partial \bar{F}}{\partial \bar{y}}$ both vanish at $(0,0)$ we have $\bar{a}_3 = \bar{a}_4 = \bar{a}_6 = 0$, so $\bar{E}$ has equation

$$\bar{F}(x,y) = \bar{y}^2 + \bar{a}_1\,\bar{x}\,\bar{y} - \bar{a}_2\,\bar{x}^2 - \bar{x}^3 = 0.$$

Let $\bar{y}^2 + \bar{a}_1\,\bar{x}\,y - \bar{a}_2\,\bar{x}^2 = (\bar{y} - \alpha\,\bar{x})\ (\bar{y} - \beta\,\bar{x})$, where α and β are either in K or in the quadratic extension $K\big(\sqrt{D}\big)$ for some $D \in \mathbb{Q}$. Then $\bar{E}$ is

$$\bar{E} : \bar{x}^3 = (\bar{y} - \alpha\,\bar{x})\ (\bar{y} - \beta\,\bar{x})\,.$$

The singular point Q of E is called a *node* if $\alpha \neq \beta$, and a cusp if $\alpha = \beta$. (So Q is a node if there are two distinct tangent lines at Q, or a cusp if there is only one.)

The structure of the group $E_{ns}(K) = E(K)\backslash\{Q\}$ of non-singular K-rational points is given by the following theorem:

Theorem 3.4.1. *Let E/K be a singular elliptic curve with singular point Q.*

1. *If Q is a cusp, then $E_{ns}(K)$ is isomorphic to K^+.*
2. *If Q is a node, and $\alpha, \beta \in K$, then $E_{ns}(K)$ is isomorphic to K^*.*
3. *If Q is a node, and $\alpha, \beta \notin K$, then $E_{ns}(K)$ is isomorphic to a subgroup of $K\big(\sqrt{D}\big)^*$ for some $D \in K$.*

(Explicit formulae for the isomorphisms in question are given in [17, page 56] and [24, page 106].) So elliptic curves are much simpler in the singular case.

The singular point satisfies the following condition.

Theorem 3.4.2. *Let E be an elliptic curve*

$$E : y^2 + a_1\,x\,y + a_3\,y = x^3 + a_2\,x^2 + a_4\,x + a_6,$$

and let $Q = (x_0, y_0)$ be a point on E. If $x_0 = 0$, then Q is a singular point if and only if $a_3 = a_4 = 0$. Otherwise, Q is a singular point if and only if

$$2y_0 + a_1\,x_0 + a_3 = 0 \qquad \textit{and} \qquad 2x_0^3 - b_4\,x_0 - b_6 = 0.$$

Proof: By definition, Q is a singular point if and only if

$$2y_0 + a_1\,x_0 + a_3 = 0 \tag{3.13}$$

and

$$-a_1\,y_0 + 3x_0^2 + 2a_2\,x_0 + a_4 = 0. \tag{3.14}$$

But if $y_0 = -\frac{1}{2}\,(a_1\,x_0 + a_3)$ then (3.14) becomes

$$a_1\left(\frac{a_1\,x_0 + a_3}{2}\right) + 3x_0^2 + 2a_2\,x_0 + a_4 = 0.$$

This is equivalent to

$$6x_0^2 + \left(a_1^2 + 4a_2\right)\,x_0 + (a_1\,a_3 + 2a_4) = 0,$$

or in other words,

$$6x_0^2 + b_2\,x_0 + b_4 = 0. \tag{3.15}$$

Also, if $y_0 = -\frac{1}{2}\,(a_1\,x_0 + a_3)$ then the elliptic curve equation gives

$$\left(\frac{a_1\,x_0 + a_3}{2}\right)^2 - a_1\,x_0\left(\frac{a_1\,x_0 + a_3}{2}\right) - a_3\left(\frac{a_1\,x_0 + a_3}{2}\right) = x_0^3 + a_2\,x_0^2 + a_4\,x_0 + a_6.$$

This is equivalent to

$$4x_0^3 + \left(a_1^2 + 4a_2\right)\,x_0^2 + (4a_4 - 2a_1a_3)\,x_0 + \left(a_3^2 + 4a_6\right) = 0,$$

or

$$4x_0^3 + b_2\,x_0^2 + b_4\,x_0 + b_6 = 0. \tag{3.16}$$

It follows from (3.13) and (3.15) that if $x_0 = 0$ then Q is singular if and only if $a_3 = b_4 = 0$; that is, if and only if $a_3 = a_4 = 0$.

Multiplying (3.15) by x_0 and subtracting (3.16) gives

$$2x_0^3 - b_4\, x_0 - b_6 = 0. \tag{3.17}$$

This means that if (3.13) holds and $x_0 \neq 0$ then (3.14) is equivalent to (3.17), and the result follows. □

3.5 Some useful elliptic curve identities

Specialising the group law to the case where we are starting with a point $Q = (x_0, y_0)$ and repeatedly adding the point $P = (0, 0)$, we get a result we will need in chapter 7.

Theorem 3.5.1. *Let E/K be an elliptic curve with equation*

$$E : y^2 + a_1\, x\, y + a_3\, y = x^3 + a_2\, x^2 + a_4\, x,$$

and let $P = (0, 0)$ and $Q = (x_0, y_0)$ be non-singular K-rational points on $E(K)$. Denote the point $Q + [n]P$ by (x_n, y_n) whenever $Q + [n]P \neq \mathcal{O}$. Then for all n such that $Q + [n]P \neq -P$, $\mathcal{O}$ or P we have

$$\begin{aligned} x_n^2 \cdot x_{n+1} &= a_4\, x_n - a_3\, y_n \\ x_{n+1} \cdot x_n^2 \cdot x_{n-1} &= -b_8 - b_6\, x_n \\ x_n^2\, (x_{n+1} + x_{n-1}) &= b_4\, x_n + b_6. \end{aligned}$$

Proof: Let $n \in \mathbb{Z}$ such that $Q + [n]P \neq -P$, $\mathcal{O}$ or P, so x_n is defined and non-zero. We first add the points $P = (0, 0)$ and $Q + [n]P = (x_n, y_n)$ to get $Q + [n + 1]P = (x_{n+1}, y_{n+1})$. By the addition formula $\lambda = \frac{y_n}{x_n}$, and so

$$x_{n+1} = \left(\frac{y_n}{x_n}\right)^2 + a_1 \left(\frac{y_n}{x_n}\right) - a_2 - x_n.$$

Similarly, we add the points $-P = (0, -a_3)$ and $Q + [n]P = (x_n, y_n)$ to get $Q + [n - 1]P = (x_{n-1}, y_{n-1})$. By the addition formula $\lambda = \frac{y_n + a_3}{x_n}$, and so

$$x_{n-1} = \left(\frac{y_n + a_3}{x_n}\right)^2 + a_1 \left(\frac{y_n + a_3}{x_n}\right) - a_2 - x_n.$$

Multiplying by x_n^2 and using the elliptic curve formula to replace $y_n^2 + a_1\,x_n\,y_n + a_3\,y_n$ by $x_n^3 + a_2\,x_n^2 + a_4\,x_n$ yields

$$x_{n+1} \cdot x_n^2 = a_4\,x_n - a_3\,y_n \tag{3.18}$$

and

$$x_{n-1} \cdot x_n^2 = (a_4 + a_1\,a_3)\,x_n + a_3\,(y_n + a_3). \tag{3.19}$$

Adding (3.18) and (3.19) gives

$$\begin{aligned} x_n^2\,(x_{n-1} + x_{n+1}) &= (2a_4 + a_1\,a_3)\,x_n + a_3^2 \\ &= b_4\,x_n + b_6. \end{aligned}$$

Multiplying (3.18) and (3.19) together and using the elliptic curve formula to write $y_n^2 + a_1\,x_n\,y_n + a_3\,y_n$ in terms of x_n gives

$$x_{n-1} \cdot x_n^4 \cdot x_{n+1} = -a_3^2\,x_n^3 + \left(a_4\,(a_4 + a_1\,a_3) - a_3^2\,a_2\right)\,x_n^2,$$

or in other words,

$$x_{n-1} \cdot x_n^2 \cdot x_{n+1} = -b_6\,x_n - b_8.$$

This completes the proof. □

3.6 Elliptic curves over finite fields

In this thesis the field K will always be either a finite field $\mathbb{F}_q$ or the rational numbers $\mathbb{Q}$. In this section we give some special properties of elliptic curves over a finite field $\mathbb{F}_q$, and in the next we give some properties of elliptic curves over $\mathbb{Q}$.

3.6.1 The number of points in $E(\mathbb{F}_q)$

Let E be an elliptic curve over $K = \mathbb{F}_q$. Since K is finite the number of K-rational points on E is finite, and it will be denoted by $\#E(\mathbb{F}_q)$. Hasse's Theorem says that $\#E(\mathbb{F}_q)$ is close to $q + 1$:

Theorem 3.6.1. *(Hasse's Theorem)*

Let E be an elliptic curve over a finite field $\mathbb{F}_q$. Then

$$|\#E(\mathbb{F}_q) - (q+1)| \leq 2\sqrt{q}.$$

This makes intuitive sense. Substituting each of the q possible x-coordinates in $\mathbb{F}_q$ into the elliptic curve equation gives a quadratic equation in y over $\mathbb{F}_q$. We would expect this to have a solution for y about half the time, and then to have two solutions, y and $-y - a_1\,x - a_3$. Adding the point at infinity, we get $q+1$ expected rational points.

It turns out that if p is a prime then every number of points in the Hasse range does actually occur:

Theorem 3.6.2. *If p is a prime, then there exists at least one elliptic curve $E/\mathbb{F}_p$ with $\#E(\mathbb{F}_p) = p + 1 + t$ for every t satisfying $|t| \leq 2\sqrt{p}$.*

3.6.2 Singular curves over $\mathbb{F}_q$

For finite fields Theorem 3.4.1 specialises to

Theorem 3.6.3. *Let E be a singular elliptic curve over the finite field $\mathbb{F}_q$ with singular point Q.*

1. *If Q is a cusp, then $E_{ns}(\mathbb{F}_q)$ is isomorphic to $\mathbb{F}_q^+$.*
2. *If Q is a node, then $E_{ns}(\mathbb{F}_q)$ is isomorphic either to $\mathbb{F}_q^*$ or to the subgroup of order $q+1$ in $\mathbb{F}_{q^2}^*$ (a cyclic group of order q^2-1).*

So if Q is a cusp then $\#E_{ns}(\mathbb{F}_q) = q$. If Q is a node then $\#E_{ns}(\mathbb{F}_q)$ is either $q-1$ or $q+1$.

Note that the above bijections reduce the discrete logarithm problem in $E(\mathbb{F}_q)$ to that in $\mathbb{F}_q^+$ if Q is a cusp, or in $\mathbb{F}_q^*$ or $\mathbb{F}_{q^2}^*$ if Q is a node. This is why singular elliptic curves are not used in cryptography.

3.7 Elliptic curves over the rationals $\mathbb{Q}$

Now let E be an elliptic curve over the field of rational numbers $K = \mathbb{Q}$.

The affine point (x, y) is said to be an *integer point* if $x, y \in \mathbb{Z}$, or a *rational point* if $x, y \in \mathbb{Q}$.

3.7.1 Points of finite order in $E(\mathbb{Q})$

Clearly, the rational points of finite order form a subgroup of $E(\mathbb{Q})$; Mazur's Theorem gives the structure of this group.

Theorem 3.7.1. *(Mazur's Theorem)*

Let E be a non-singular elliptic curve over $\mathbb{Q}$. Then the group of rational points on E of finite order is isomorphic either to C_n with $n = 1, 2, \ldots, 10$ or 12, or to $C_n \times C_2$ with $n = 2$, 4, 6 or 8.

So if a rational point P has finite order N in E, then $N \in \{1, 2, \ldots, 10\} \cup 12$.

For singular curves, by Theorem 3.4.1 the non-singular rational points form a group isomorphic either to $\mathbb{Q}^+$ (the additive group of the rationals) or to a subgroup of $\mathbb{Q}\big(\sqrt{D}\big)^*$ where $D \in \mathbb{Q}$ (the multiplicative group of a quadratic extension of $\mathbb{Q}$). Since every element of finite order in $\mathbb{Q}\big(\sqrt{D}\big)^*$ has order 1, 2, 3 or 4 (see Proposition 4.2 in [29]) and no element of $\mathbb{Q}^+$ has finite order except 0, we have

Theorem 3.7.2. *Let E be a singular elliptic curve over $\mathbb{Q}$, and let P be a point of finite order in $E_{ns}(\mathbb{Q})$. Then $P = \mathcal{O}$ or P has order 2, 3 or 4.*

3.7.2 X, Y, Z coordinates of points in $E(\mathbb{Q})$

If an elliptic curve E over $\mathbb{Q}$ has integral coefficients, then it is easy to prove that for any rational point (x, y) on E the denominator of x will be a square Z^2 and the denominator of y will be the cube Z^3. (See for example [18] for a different proof.) If the a_i are not necessarily integers, we have the following generalisation:

Theorem 3.7.3. *Let E be an elliptic curve over $\mathbb{Q}$ with equation*

$$E : y^2 + a_1 x\, y + a_3\, y = x^3 + a_2\, x^2 + a_4\, x + a_6$$

and let $P = (x, y)$ be a finite rational point on E. Let X, Y, Z be integers with $|Z|$ minimal such that

$$x = \frac{X}{Z^2} \quad \text{and} \quad y = \frac{Y}{Z^3}.$$

Then $\gcd(X, Z)$ and $\gcd(Y, Z)$ are divisible only by primes which also divide the denominator of some coefficient a_i.

Proof: Since (x, y) lies on E, we have

$$\frac{Y^2}{Z^6} + a_1 \left(\frac{XY}{Z^5}\right) + a_3 \left(\frac{Y}{Z^3}\right) = \frac{X^3}{Z^6} + a_2 \left(\frac{X^2}{Z^4}\right) + a_4 \left(\frac{X}{Z^2}\right) + a_6,$$

which becomes on multiplying by Z^6

$$Y^2 + a_1\, X\, Y\, Z + a_3\, Y\, Z^3 = X^3 + a_2\, X^2\, Z^2 + a_4\, X\, Z^4 + a_6\, Z^6. \tag{3.20}$$

Let p be any prime dividing Z and not dividing the denominator of any a_i; then it is easy to see from (3.20) that either p divides both Y and X or p divides neither. Suppose p divides X and Y. Then by (3.20), $p^3 \mid Y^2$, and hence $p^2 \mid Y$. But then by (3.20), $p^4 \mid X^3$, and hence $p^2 \mid X$. Using (3.20) a final time, we see that this implies $p^5 \mid Y^2$, and hence $p^3 \mid Y$. But if $p^2 \mid X$ and $p^3 \mid Y$, replacing X by $X' = \frac{X}{p^2}$, Y by $Y' = \frac{Y}{p^3}$ and Z by $Z' = \frac{Z}{p}$ yields a smaller integer $|Z'|$ satisfying $(x, y) = \left(\frac{X'}{Z'^2}, \frac{Y'}{Z'^3}\right)$, contradicting the minimality of $|Z|$. It follows that p does not divide X or Y. □

We refer to the integers X, Y and Z as the X-, Y- and Z-*coordinates* of the point P. Note that X,Y and Z are unique up to the sign of Z and Y, and that in homogeneous coordinates, $P = (X\, Z : Y : Z^3)$.

3.8 Weighted coordinates

Theorem 3.7.3 leads to an alternative projective representation which is very natural for elliptic curves. In the *weighted projective* representation (also called the

Jacobian representation) a triple $(X;Y;Z)$ corresponds to the affine coordinates $\left(\frac{X}{Z^2}, \frac{Y}{Z^3}\right)$ whenever $Z \neq 0$, and the point at infinity $\mathcal{O}$ is $(1;1;0)$. This is equivalent to using a weighted projective curve equation of the form

$$E : Y^2 + a_1\, X\, Y\, Z + a_3\, Y\, Z^3 = X^3 + a_2\, X^2\, Z^2 + a_4\, X\, Z^4 + a_6\, Z^6. \qquad (3.21)$$

We identify the points $(X;Y;Z)$ and $(\theta^2\, X; \theta^3\, Y; \theta\, Z)$ for any $\theta \in \overline{K}$.

The point $(X;Y;Z)$ in weighted coordinates corresponds to the point $(X\, Z : Y : Z^3)$ in homogeneous projective coordinates. We will move freely between the different representations, always making it clear which one we are using.

3.8.1 Elliptic curve addition in weighted coordinates

Point addition can be done in weighted projective coordinates using field multiplications and additions only, with no inversions.

Let $P_1 = (X_1;Y_1;Z_1)$ and $P_2 = (X_2;Y_2;Z_2)$ be finite points on an elliptic curve E of the form (3.21). Substituting $x = \frac{X}{Z^2}$ and $y = \frac{Y}{Z^3}$ into the affine group law from section 3.2 and letting

$$\xi = X_2\, Z_1^2 - X_1\, Z_2^2$$

we can derive the following formulae for the group law in weighted coordinates:

$$-P_1 = (X_1; -Y_1 - a_1\, X_1\, Z_1 - a_3\, Z_1^3; Z_1).$$

If $P_2 = -P_1$ then $P_1 + P_2 = \mathcal{O}$.

If $P_2 \neq \pm P_1$ then $P_3 = (\bar{X}_3; \bar{Y}_3; \bar{Z}_3)$, where

$$\begin{aligned}
\bar{Z}_3 &= Z_1\, Z_2\, \xi, \\
\bar{X}_3 &= \left(Y_2\, Z_1^3 - Y_1\, Z_2^3\right)^2 + a_1\, Z_1\, Z_2\, \xi\, \left(Y_2\, Z_1^3 - Y_1\, Z_2^3\right) \\
&\qquad - a_2\, \left(Z_1\, Z_2\, \xi\right)^2 - \xi^2\, \left(X_2\, Z_1^2 + X_1\, Z_2^2\right),
\end{aligned}$$

and

$$\begin{aligned}
\bar{Y}_3 &= \Big(\left(Y_2\, Z_1^3 - Y_1\, Z_2^3\right) - a_1\, Z_1\, Z_2\, \xi\Big)\, \bar{X}_3 \\
&\qquad + Z_1\, Z_2\, \xi\, \left(Y_1\, X_2\, Z_2 - Y_2\, X_1\, Z_1\right) - a_3\, \left(Z_1\, Z_2\, \xi\right)^3.
\end{aligned}$$

The integers $\bar{X}$, $\bar{Y}$ and $\bar{Z}$ may have factors in common; let $\theta = \gcd(\bar{X}, \bar{Y}, \bar{Z})$. Then P_3 can also be written as $P_3 = (X_3; Y_3; Z_3)$, where

$$X_3 = \frac{\bar{X}_3}{\theta^2}, \quad Y_3 = \frac{\bar{Y}_3}{\theta^3} \quad \text{and} \quad Z_3 = \frac{\bar{Z}_3}{\theta}.$$

A doubling formula for the case $P_2 = P_1$ can be similarly derived. Notice that $\bar{X}_3$, $\bar{Y}_3$ and $\bar{Z}_3$ can be calculated with no field inversions, only additions and multiplications. We only need one field inversion at the end (θ^{-1}) to write $P_3 = (X_3; Y_3; Z_3)$ in lowest terms.

3.9 Elliptic curves over $\mathbb{Z}_{p^r}$

Elliptic curves can also be considered over $\mathbb{Z}_n$, the ring of integers modulo a composite number n. (Such curves are used for instance in Lenstra's factoring algorithm [15] and the Goldwasser-Killian primality proving algorithm.) In this section we define an elliptic curve over $\mathbb{Z}_{p^r}$, where p^r is a prime power, and describe some of its properties.

Definition: Let p^r be a prime power. The projective plane over the ring $\mathbb{Z}_{p^r}$ is

$$\mathbb{P}^2(\mathbb{Z}_{p^r}) = \{(X, Y, Z) \mid X, Y, Z \in \mathbb{Z}_{p^r} \text{ and } X, Y, Z \text{ not all divisible by } p\},$$

where we identify any two points (a, b, c) and $(\theta a, \theta b, \theta c)$ where θ is an invertible element of $\mathbb{Z}_{p^r}$. We denote the equivalence class containing (a, b, c) by $(a : b : c)$.

Definition: Let p^r be a prime power. An elliptic curve E over $\mathbb{Z}_{p^r}$ is the set of projective points $(X : Y : Z) \in \mathbb{P}^2(\mathbb{Z}_{p^r})$ satisfying the equation

$$E : Y^2\,Z + a_1\,X\,Y\,Z + a_3\,Y\,Z^2 \equiv X^3 + a_2\,X^2\,Z + a_4\,X\,Z^2 + a_6\,Z^3 \mod p^r \quad (3.22)$$

where the a_i are in $\mathbb{Z}_{p^r}$. The *point at infinity* $\mathcal{O}_{p^r}$ is the projective point $(0 : 1 : 0)$.

The curve $E/\mathbb{Z}_{p^r}$ is said to be *non-singular* over $\mathbb{Z}_{p^r}$ if, when it is written in the form $F(X, Y, Z) = 0$, at least one of the three partial derivatives $\frac{\partial F}{\partial X}$, $\frac{\partial F}{\partial Y}$, $\frac{\partial F}{\partial Z}$ is coprime to p at each weighted projective point $(X; Y; Z)$ satisfying the equation. (This is equivalent to the usual non-singularity condition in homogeneous

coordinates.) If all three partial derivatives are divisible by p at some point P, then the point P and the curve E are said to be *singular* over $\mathbb{Z}_{p^r}$. The set of non-singular points of $E(\mathbb{Z}_{p^r})$ (which includes the point at infinity $\mathcal{O}_{p^r}$) is called the *non-singular* part of $E(\mathbb{Z}_{p^r})$, and is denoted $E_{ns}(\mathbb{Z}_{p^r})$. The curve E is singular over $\mathbb{Z}_{p^r}$ if and only if p divides the discriminant Δ of E.

3.9.1 The group $E(\mathbb{Z}_{p^r})$

It turns out that the set of non-singular points of an elliptic curve E over $\mathbb{Z}_{p^r}$ still forms a group under a certain addition law; for a good description see Washington [32, pages 59–66].

Theorem 3.9.1. *[15]*

Let p^r be a prime power, and let E be an elliptic curve over $\mathbb{Z}_{p^r}$. Then $E_{ns}(\mathbb{Z}_{p^r})$ forms an abelian group with the point at infinity $(0:1:0)$ as the zero.

The addition law for adding two points $P_1 = (X_1 : Y_1 : Z_1)$ and $P_2 = (X_2 : Y_2 : Z_2)$ where Z_1 and Z_2 are both coprime to p is the same as the addition law in $E_{ns}(\mathbb{Q})$ (which is given in weighted projective coordinates in section 3.8) but with all calculations performed modulo p^r instead of in $\mathbb{Q}$. This works because in weighted coordinates the group law does not involve any inversions until we have to find θ^{-1} at the end, and if Z_1 and Z_2 are coprime to p then θ will be coprime to p and the final inversion will be possible.

If Z_1 or Z_2 is divisible by p, then this formula will not work (since $\bar{X}_3$, $\bar{Y}_3$ and $\bar{Z}_3$ are all divisible by p), but one of two different formulae applies. We do not need to describe the group law further, but refer the interested reader to Lenstra's paper [15].

3.9.2 The reduction modulo p^r map

Let p be a fixed prime and E a fixed elliptic curve over $\mathbb{Q}$ with equation

$$E : Y^2\,Z + a_1\,X\,Y\,Z + a_3\,Y\,Z^2 = X^3 + a_2\,X^2\,Z + a_4\,X\,Z^2 + a_6\,Z^3,$$

where the a_i are p-integers.

Since the a_i are p-integers, the curve

$$E : Y^2\,Z + a_1\,X\,Y\,Z + a_3\,Y\,Z^2 \equiv X^3 + a_2\,X^2\,Z + a_4\,X\,Z^2 + a_6\,Z^3 \bmod p^r,$$

where the a_i are now regarded as elements of $\mathbb{Z}_{p^r}$, is an elliptic curve over $\mathbb{Z}_{p^r}$ which we denote again by E or by $E \bmod p^r$.

There is a natural map from $E(\mathbb{Q})$ to $E(\mathbb{Z}_{p^r})$, defined as follows.

Definition: For any $r \in \mathbb{N}$ the *reduction modulo* p^r map from $E(\mathbb{Q})$ to $E(\mathbb{Z}_{p^r})$ is given by

$$P = (X : Y : Z) \mapsto P \bmod p^r = (X \bmod p^r : Y \bmod p^r : Z \bmod p^r).$$

If $P_1 \bmod p^r = P_2 \bmod p^r$ in $E(\mathbb{Z}_{p^r})$ then we write $P_1 \equiv P_2 \bmod p^r$. The point P is said to be *above* $P \bmod p^r$, and $P \bmod p^r$ is *below* P. P is said to be *nonsingular modulo* p if $P \bmod p^r$ is in $E_{ns}(\mathbb{Z}_{p^r})$, and *singular modulo* p otherwise.

Similarly, we have

Definition: For any $r, k \in \mathbb{N}$ with $k \leq r$ the *reduction modulo* p^r map from $E(\mathbb{Z}_{p^r})$ to $E(\mathbb{Z}_{p^k})$ is given by

$$(X : Y : Z) \mapsto (X \bmod p^k : Y \bmod p^k : Z \bmod p^k).$$

The point $P \bmod p^r$ is said to be *above* $P \bmod p^k$, and $P \bmod p^k$ is *below* $P \bmod p^r$.

Let $P_1, P_2 \in E_{ns}(\mathbb{Q})$, and suppose we want to find $(P_1 + P_2) \bmod p^r$ in $E_{ns}(\mathbb{Z}_{p^r})$. Since the addition laws in $E_{ns}(\mathbb{Z}_{p^r})$ and $E_{ns}(\mathbb{Q})$ are given by the same polynomial equations, it follows that working over $\mathbb{Q}$ and then reducing modulo p^r gives exactly the same result as working over $\mathbb{Z}_{p^r}$. In other words,

Theorem 3.9.2. *Let $r, k \in \mathbb{N}$ with $k \leq r$. The reduction modulo p^r map from $E_{ns}(\mathbb{Q})$ or $E_{ns}(\mathbb{Z}_{p^r})$ to $E_{ns}(\mathbb{Z}_{p^k})$ is a group homomorphism.*

It is frequently more convenient to write E in weighted projective coordinates as

$$E : Y^2 + a_1XYZ + a_3YZ^3 = X^3 + a_2X^2Z^2 + a_4XZ^4 + a_6Z^6.$$

Since the a_i are p-integers, by Theorem 3.7.3 every rational point in $E(\mathbb{Q})$ can be written as $(x, y) = \left(\frac{X}{Z^2}, \frac{Y}{Z^3}\right)$, where p does not divide both Y and Z. So the points in $E(\mathbb{Q})$ and $E(\mathbb{Z}_{p^r})$ which are above the group identity $(0 : 1 : 0)$ of $E(\mathbb{Z}_{p^k})$ for $k \leq r$ are precisely those whose weighted Z-coordinate is divisible by p^k.

3.9.3 The number of points on $E(\mathbb{Z}_{p^r})$

For each $r \in \mathbb{N}$ the group $E_{ns}(\mathbb{Z}_{p^{r+1}})$ has p times as many points as $E_{ns}(\mathbb{Z}_{p^r})$:

Theorem 3.9.3. *[15]*

For $r \in \mathbb{N}$,

$$\#E_{ns}(\mathbb{Z}_{p^r}) = p^{r-1}\,\#E_{ns}(\mathbb{Z}_p).$$

Since the reduction modulo p^r map is a group homomorphism, the next result follows.

Theorem 3.9.4. *Let P be a point in $E_{ns}(\mathbb{Z}_{p^r})$. Then there are precisely p distinct points in $E_{ns}(\mathbb{Z}_{p^{r+1}})$ which are above P.*

Let P be a point in $E(\mathbb{Q})$, and for $r \in \mathbb{N}$, let the order of $P \bmod p^r$ be N_r. Since the $p\,N_r$ points above $\langle P \bmod p^r\rangle$ in $E(\mathbb{Z}_{p^{r+1}})$ form a subgroup of $E(\mathbb{Z}_{p^{r+1}})$, and $\langle P \bmod p^{r+1}\rangle$ is a subgroup of this group, N_{r+1} divides $p\,N_r$. But since the reduction modulo p^r map is a homomorphism, N_{r+1} is a multiple of N_r. Hence N_{r+1} is either N_r or $p\,N_r$. Now let $p^{\bar{w}}$ be the highest power of p dividing the weighted Z-coordinate of $[N_1]P \in E(\mathbb{Q})$. So

$$N_1 = N_2 = \ldots = N_{\bar{w}}, \quad \text{and} \quad N_{\bar{w}+1} = p\,N_{\bar{w}}.$$

This means that for $r < \bar{w}$ the subgroup $\langle P \bmod p^{r+1}\rangle$ has order N_1, while for $r = \bar{w}$ the subgroup $\langle P \bmod p^{\bar{w}+1}\rangle$ contains all p points P_i above P and has order $p\,N_1$ (the P_i are the points $P_1,\ [N_{\bar{w}} + 1]\,P_1,\ \ldots,\ [(p-1)N_{\bar{w}} - 1]\,P_1 \in \langle P_i \bmod p^{\bar{w}+1}\rangle$). In fact we have the following result:

Theorem 3.9.5. *Let p be a prime, and let E be an elliptic curve over $\mathbb{Q}$ with equation*

$$E : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$$

where the a_i are p-integers. Let P be a point in $E(\mathbb{Q})$ which is non-singular modulo p, and for $r \in \mathbb{N}$ let $P \bmod p^r$ have order N_r in $E(\mathbb{Z}_{p^r})$. Let $N_1 \geq 4$, and $p^{\bar{w}}$ be the highest power of p dividing the weighted Z-coordinate of $[N_1]P$.

If p is odd, or $p = 2$ and $w \geq 2$, then for $r \in \mathbb{N}$,

$$N_r = \begin{cases} N_1 & \text{for } r \leq w, \\ p\,N_{r-1} & \text{for } r > w. \end{cases}$$

Otherwise, if $p = 2$ and $w = 1$, let $2^{\bar{v}}$ be the highest power of 2 dividing the weighted Z-coordinate of $[N_2]P$. Then

$$N_r = \begin{cases} N_1 & \text{if } r = 1, \\ 2\,N_1 & \text{if } 2 \leq r \leq \bar{v}, \\ 2\,N_{r-1} & \text{if } r > \bar{v}. \end{cases}$$

Theorem 3.9.5 follows easily from an EDS result by Ward [31] (Theorem 4.7.5) which we will prove by elementary means in chapter 5.

If $N_{r+1} = p\,N_r$ and $Q \in E(\mathbb{Q})$ such that $Q \equiv [k]P \bmod p^r$, then the p points $[k + iN_r]P \bmod p^{r+1}$ are all distinct, and must be precisely the p points above Q in $E(\mathbb{Z}_{p^r})$. Hence the following is true:

Theorem 3.9.6. *Let P and Q be points in $E(\mathbb{Q})$ which are non-singular modulo p, and let $p^{\bar{w}}$ and $p^{\bar{v}}$ be the highest powers of p dividing the weighted Z-coordinates of $[N_1]P$ and $[N_2]P$ respectively in $E(\mathbb{Q})$. If either p is odd and $r \geq \bar{w}$, or $p = 2$ and $r \geq \bar{v}$, then*

$$Q \in \langle P \bmod p^{r+1} \rangle \text{ in } E(\mathbb{Z}_{p^{r+1}}) \quad \Leftrightarrow \quad Q \in \langle P \bmod p^r \rangle \text{ in } E(\mathbb{Z}_{p^r}).$$

In other words, the subgroup generated by $P \bmod p^{r+1}$ in $E(\mathbb{Z}_{p^{r+1}})$ lies precisely above the subgroup generated by $P \bmod p^r$ in $E(\mathbb{Z}_{p^r})$.

If $p = 2$ and $\bar{w} \geq 2$ then $\bar{w} = \bar{v}$, and the result is the same as in the odd p case. Notice, however, that if $p = 2$, $w = 1$ and $\bar{v} \geq 3$ then $Q \bmod 4 \in \langle P \bmod 4 \rangle$ in $E(\mathbb{Z}_4)$ does not imply $Q \bmod 2^v \in \langle P \bmod 2^v \rangle$ in $E(\mathbb{Z}_{2^v})$.

3.10 The division polynomials

In this section we define the division polynomials of an elliptic curve and describe some of their properties. (Charlap and Robbins have given elementary proofs of all these results in [4].)

Let E be an elliptic curve over a field K, and let $P = (x, y)$ be a point in $E(K)$. We are interested in the sequence of points $[n]P$ for $n \in \mathbb{Z}$, i.e., the points obtained by adding P to itself repeatedly. We denote $[n]P = (x_n, y_n)$ whenever $[n]P \neq \mathcal{O}$.

It is clear from the algebraic formulae for elliptic curve addition in section 3.2 that the coordinates of the sum $P_1 + P_2$ of two points are rational functions of the coordinates of P_1 and P_2. By repeated application of the group law it follows that for $n \in \mathbb{N}$ the multiplication-by-n map $P \mapsto [n]P$ can be expressed in terms of (albeit complicated) rational functions in x and y. In fact the following is true.

Theorem 3.10.1. *Let E be an elliptic curve over a field K, and let $n \in \mathbb{Z}$. Then there exist polynomials ψ_n, θ_n and ω_n in $K[x, y]$ such that for all non-singular points $P = (x, y)$ in $E(\overline{K})$,*

$$[n]P = \left(\frac{\theta_n(x,y)}{\psi_n(x,y)^2}, \frac{\omega_n(x,y)}{\psi_n(x,y)^3} \right) \qquad \text{if } [n]P \neq \mathcal{O},$$

and

$$\psi_n(x, y) = 0 \;\Leftrightarrow\; [n]P = \mathcal{O}.$$

The polynomial $\psi_n(x, y)$ is called the *nth division polynomial* of the curve E. The nth division polynomial characterises the points in $E[n] \setminus \{\mathcal{O}\}$, in that $(x, y) \in E[n] \setminus \{\mathcal{O}\}$ if and only if $\psi_n(x, y) = 0$.

The sequences (θ_n) and (ω_n) can be expressed in terms of the sequence (ψ_n) as follows:

Theorem 3.10.2. *Let E/K be an elliptic curve with equation*

$$E : y^2 + a_1\, x\, y + a_3\, y = x^3 + a_2\, x^2 + a_4\, x + a_6.$$

Then for all $n \in \mathbb{Z}$,

$$\theta_n = x\, \psi_n^2 - \psi_{n-1}\, \psi_{n+1}$$

and if $\mathrm{char}(K) \neq 2$ *and* $n \neq 0$,

$$\omega_n = \frac{1}{2}\left(\frac{\psi_{2n}}{\psi_n} - \left(a_1\,\theta_n + a_3\,\psi_n^2\right)\,\psi_n\right).$$

Remark: For $n \neq 0$, $\theta_{-n} = \theta_n$ and $\omega_{-n} = \omega_n + (a_1\ \theta_n + a_3\,\psi_n^2)\ \psi_n$.

The division polynomials can be easily computed, using the following result:

Theorem 3.10.3. *Let* E/K *be an elliptic curve with equation*

$$E : y^2 + a_1\,x\,y + a_3\,y = x^3 + a_2\,x^2 + a_4\,x + a_6.$$

Then the division polynomials ψ_n *are given by the following recursion:*

$$\begin{aligned}
\psi_0 &= 0,\\
\psi_1 &= 1,\\
\psi_2 &= 2y + a_1\,x + a_3,\\
\psi_3 &= 3x^4 + b_2\,x^3 + 3b_4\,x^2 + 3b_6\,x + b_8\\
\psi_4 &= \left(2x^6 + b_2\,x^5 + 5b_4\,x^4 + 10b_6\,x^3 + 10b_8\,x^2 + (b_2\,b_8 - b_4\,b_6)\ x + b_4\,b_8 - b_6^2\right)\psi_2\\
&\ \ \vdots\\
\psi_{2k+1} &= \psi_{k+2}\,\psi_k^3 - \psi_{k-1}\,\psi_{k+1}^3 \qquad \textit{for}\ \ k \geq 2,\\
\psi_{2k} &= \left(\frac{\psi_{k+2}\,\psi_{k-1}^2 - \psi_{k-2}\,\psi_{k+1}^2}{\psi_2}\right)\psi_k \qquad \textit{for}\ \ k \geq 3,\ \ \textit{and}\\
\psi_{-n} &= -\psi_n \qquad \textit{for}\ \ n < 0.
\end{aligned}$$

It is easy to show that

Theorem 3.10.4. *Let* E/K *be an elliptic curve with equation*

$$E : y^2 + a_1\,x\,y + a_3\,y = x^3 + a_2\,x^2 + a_4\,x + a_6.$$

Then the coefficients of the division polynomials ψ_n *are in the ring* $R = \mathbb{Z}[a_1, a_2, a_3, a_4, a_6]$ *for all* $n \in \mathbb{Z}$.

In the important special case where $K = \mathbb{Q}$ we have

Theorem 3.10.5. *Let E be an elliptic curve over $\mathbb{Q}$ with coefficients a_i and division polynomials ψ_n for $n \in \mathbb{Z}$, and let p be a prime such that the a_i are p-integers. Then the curve $E \bmod p$ over $\mathbb{F}_p$ with coefficients $a_i \bmod p$ has division polynomials $\psi_n \bmod p$ for $n \in \mathbb{Z}$.*

So for all points $P \in E(\mathbb{Q})$ which are nonsingular modulo p,

$$\psi_n(x, y) \equiv 0 \bmod p \;\Leftrightarrow\; [n]P \equiv \mathcal{O} \bmod p.$$

Since the division polynomials are always evaluated at points on the curve, the computation of ψ_n can be carried out modulo the equation of E. In particular, we can assume that the degree of ψ_n in y never exceeds one (since we can replace y^2 by $x^3 - a_2\,x^2 - a_4\,x - a_6 - a_1\,x\,y - a_3\,y$ whenever it occurs). Because y enters into the recursion for ψ_n only through the polynomial ψ_2 and $\psi_2^2 \bmod E$ does not depend on y, we have

Theorem 3.10.6. *[4]*

Let $R = \mathbb{Z}[a_1, \ldots, a_6]$. Then for $n \in \mathbb{Z}$, ψ_n satisfies

$$\psi_n(x, y)^2 = n^2 \prod_{\substack{(x_i, y_i) \in \\ E[n] - \{\mathcal{O}\}}} (x - x_i)$$

in $R[x, y] / \left(y^2 + a_1\,x\,y + a_3\,y - x^3 - a_2\,x^2 - a_4\,x - a_6\right)$.

Since the birational map is a group homomorphism (so $P \in E[n] \;\Leftrightarrow\; P' \in E'[n]$ for all $n \in \mathbb{N}$) it follows that

Theorem 3.10.7. *Let E/K be an elliptic curve with equation*

$$E : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6.$$

and let E' be a birationally equivalent curve obtained by an admissible change of variables (3.8) *with $u = 1$. Then*

$$\psi'_n(x', y') = \psi_n(x, y) \quad \text{for all } n \in \mathbb{Z},$$

i.e., the values of the division polynomials evaluated at any particular point do not change during this transformation.

It turns out that if $P = (x, y)$ is a singular point, then $\psi_n(x, y) = 0$ for all $|n| > 1$. Of course, Theorem 3.10.1 does not apply to this case.

Theorem 3.10.8. *Let E/K be an elliptic curve with division polynomials ψ_n for $n \in \mathbb{Z}$. Then the point $P = (x, y)$ is singular if and only if $\psi_3(x, y) = \psi_4(x, y) = 0$, and then $\psi_n(x, y) = 0$ in K for all $|n| > 1$.*

Proof: Consider a birationally equivalent curve E' in which $P' = (0, 0)$. Then P' is singular if and only if $a'_3 = a'_4 = 0$. But by Theorem 3.10.3,

$$\psi'_2(0,0) = a'_3, \quad \psi'_3(0,0) = b'_8, \qquad \text{and} \qquad \psi'_4(0,0) = \left(b'_4\, b'_8 - {b'_6}^2\right) \psi'_2(0,0),$$

where $b'_4 = a'_1\, a'_3 + 2a'_4$, $b'_6 = {a'_3}^2$ and $b'_8 = -a'_1\, a'_3\, a'_4 + a'_2\, {a'_3}^2 - {a'_4}^2$. So $a'_3 = a'_4 = 0$ if and only if $\psi'_3(0, 0) = \psi'_4(0, 0) = 0$, and if this is the case then we also have $\psi'_2(0, 0) = 0$. It is easy to prove that then $\psi'_n(0, 0) = 0$ for all $|n| > 1$. Since by Theorem 3.10.7 we have $\psi_n(x, y) = \psi'_n(0, 0)$ for all $n \in \mathbb{Z}$, the result follows. □

So E is singular if and only if all the ψ_n with $|n| > 1$ have some simultaneous root (which will be the singular point).

Theorem 3.10.9. *[4]*

The division polynomials of an elliptic curve satisfy the following equation:

$$\psi_{m+n}\, \psi_{m-n} = \psi_{m+1}\, \psi_{m-1}\, \psi_n^2 - \psi_{n+1}\, \psi_{n-1}\, \psi_m^2 \qquad \text{for } n, m \in \mathbb{Z}. \tag{3.23}$$

Furthermore,

Theorem 3.10.10. *[6]*

The division polynomials of an elliptic curve have the divisibility property that ψ_n divides ψ_m whenever n divides m.

The relation $[nM](x, y) = [n]\,([M](x, y))$ for all $n, M \in \mathbb{Z}$ leads to the following property:

Theorem 3.10.11. *[6]*

Let $P = (x, y)$ be a non-singular rational point on an elliptic curve E/K, and for $n \in \mathbb{Z}$ denote the point $[n]P$ by (x_n, y_n). Then the division polynomials of E satisfy

$$\psi_{nM}(x, y) = \left(\psi_M(x, y)\right)^{n^2} \psi_n(x_M, y_M) \qquad \text{for all } n, M \in \mathbb{Z}.$$

3.10.1 The relationship between division polynomials and the Z-coordinates of rational points

If $E/\mathbb{Q}$ has integer coefficients and $P = (x, y)$ is an integer point then the ring $R = \mathbb{Z}[a_1, \ldots, a_6]$ is just $\mathbb{Z}$, and $\psi_n(x, y)$ is an integer for all $n \in \mathbb{Z}$. Let $n \in \mathbb{Z}$ such that $[n]P \neq \mathcal{O}$. Then we have

$$x([n]P) = \frac{\theta_n(x,y)}{\psi_n(x,y)^2} \quad \text{and} \quad y([n]P) = \frac{\omega_n(x,y)}{\psi_n(x,y)^3}, \tag{3.24}$$

where

$$\theta_n(x,y) = x\,\psi_n(x,y)^2 - \psi_{n-1}(x,y)\,\psi_{n+1}(x,y). \tag{3.25}$$

By Theorem 3.7.3, we also have

$$x([n]P) = \frac{X_n}{Z_n^2} \quad \text{and} \quad y([n]P) = \frac{Y_n}{Z_n^3},$$

where X_n, Y_n and Z_n are integers (unique up to the sign of Z_n and Y_n) with X_n and Y_n coprime to Z_n.

It follows from (3.25) that if $\psi_n(x, y)$ is coprime to $\psi_{n-1}(x, y)$ and $\psi_{n-1}(x, y)$ then $\psi_n(x, y)$ is coprime to $\theta_n(x, y)$, and (3.24) is already in lowest terms. So we have

$$X_n = \theta_n(x,y), \quad Z_n = \pm\psi_n(x,y), \quad \text{and} \quad Y_n = \pm\omega_n(x,y),$$

and we can choose $Z_n = \psi_n(x, y)$ and $Y_n = \omega_n(x, y)$.

We will discuss this further in section 4.5.4. We will show that if the a_i and x, y are integers then adjacent terms of $(\psi_n(x, y))$ are coprime if and only if $\psi_3(x, y)$ and $\psi_4(x, y)$ are coprime. In the special case where $(x, y) = (0, 0)$ this is true if and only if a_3 and a_4 are coprime; then $(\psi_n(0, 0))$ and (Z_n) are the same sequence.

Chapter 4

Elliptic divisibility sequences

In this chapter we give some background on elliptic divisibility sequences (EDSs). In section 4.1 we define elliptic sequences, EDSs and generalised EDSs, and discuss the notion of equivalence of elliptic sequences. In section 4.2 we describe a special class of elliptic sequences called Lucas sequences, and in section 4.3 we give some formulae for computing elliptic sequences. In section 4.4 we consider the existence and uniqueness of EDSs with given initial values, and in section 4.5 we describe the relationship between elliptic sequences and elliptic curves. In section 4.6 we give some basic properties of EDSs. We are most interested in the properties of an EDS when reduced modulo a prime power; in section 4.7 we give the known results and describe how we have extended them. Wherever possible we point out the connections between EDS and elliptic curve results.

4.1 Definitions

We start with some definitions.

4.1.1 Elliptic sequences and EDSs

Definition: An *elliptic sequence* (h_n) is an infinite sequence of rational numbers $\ldots, h_{-2}, h_{-1}, h_0, h_1, h_2, \ldots$ satisfying

$$h_{m+n}\, h_{m-n} = h_{m+1}\, h_{m-1}\, h_n^2 - h_{n+1}\, h_{n-1}\, h_m^2 \quad \text{for all } m, n \in \mathbb{Z}. \tag{4.1}$$

Definition: An *elliptic divisibility sequence* (or *EDS*) is an elliptic sequence (h_n), all of whose terms are integers, with the *divisibility property* that h_n divides h_m whenever n divides m.

Elliptic sequences and EDSs arise from the study of the division polynomials of an elliptic curve. If $P = (x, y)$ is a rational point on an elliptic curve E over $\mathbb{Q}$ then by Theorem 3.10.9 the sequence (h_n) of rational numbers defined by $h_n = \psi_n(x, y)$ for $n \in \mathbb{Z}$ (i.e., the sequence (h_n) of division polynomials evaluated at P) is an elliptic sequence. Furthermore, if P is an integer point and the coefficients a_i of the curve are integers, then by Theorems 3.10.4 and 3.10.10 the values h_n are integers and have the divisibility property that h_n divides h_m whenever n divides m; in other words, (h_n) is an EDS.

Morgan Ward studied elliptic sequences and EDSs in [30], focusing on their properties when reduced modulo a prime p. (The rational case is not essentially more general, as we will explain later.) He was interested in elliptic sequences as a generalisation of Lucas sequences (see section 4.2) and proved in [30] that many of the properties of Lucas sequences carry over to elliptic sequences. Rachel Shipsey continued this study of EDSs in [23], seeing EDSs as a convenient way of studying elliptic curves. In particular, she and Nelson Stephens used some of the theory of EDSs she developed to find elegant alternative attacks on the elliptic curve discrete log problem in the MOV [16] and anomalous curve [26] cases. However, Ward's definition of an EDS is slightly different from Shipsey's (which is the one we are using); Ward starts the sequence at h_0 and only demands that the recursion (4.1) be satisfied for $m \geq n \geq 1$. It turns out that his definition actually is more general, because it allows a (fairly uninteresting) family of EDSs with $h_0 \neq 0$ (see [30]), whereas we have the following theorem:

Theorem 4.1.1. *If (h_n) is a non-trivial elliptic sequence, then $h_0 = 0$, $h_1 = \pm 1$, and*

$$h_{-n} = -h_n \quad \text{for all } n \in \mathbb{Z}.$$

Proof: Setting $m = n = 0$ in (4.1) gives $h_0 = 0$. Therefore setting $n = 0$ in (4.1) gives $h_m^2 = -h_1\, h_{-1}\, h_m^2$ for all $m \in \mathbb{Z}$, and so (unless all terms h_m are zero)

$h_1 h_{-1} = -1$. Now setting $m = 0$ in (4.1) gives $h_n h_{-n} = h_1 h_{-1} h_n^2 = -h_n^2$ for all $n \in \mathbb{Z}$. It follows that $h_{-n} = -h_n$ for all $n \in \mathbb{Z}$, and hence that $h_1 = \pm 1$. □

However, if the sequence $h_0 = 0, h_1, h_2, h_3, \ldots$ is an EDS by Ward's definition, then the sequence $\ldots, -h_3, -h_2, -h_1, 0, h_1, h_2, h_3, \ldots$ is an EDS by Shipsey's definition, and vice versa. So there is a one-to-one correspondence between Ward's EDSs with $h_0 = 0$ and Shipsey's EDSs.

Since the sequence obtained by multiplying each term of (h_n) by -1 is also a solution of (4.1), Theorem 4.1.1 means that we need only consider elliptic sequences with $h_1 = 1$; we do this throughout the thesis.

Examples:

1. The sequence of integers $\mathbb{Z}$ is an EDS.

2. Let (F_n) be the Fibonacci sequence, defined by the initial values $F_0 = 0$, $F_1 = 1$, and the linear recursion

$$F_n = F_{n-1} + F_{n-2} \quad \text{for} \quad n \geq 2.$$

   It is well known that (F_n) is a divisibility sequence. Now for all $n \geq 0$, let

$$h_n = (-1)^{\frac{1}{2}(n-1)(n-2)} F_n, \quad \text{and} \quad h_{-n} = -h_n,$$

   i.e., let (h_n) be the sequence

$$\ldots, 3, 2, -1, -1, 0, 1, 1, -2, -3, 5, 8, -13, -21, \ldots$$

   Then (h_n) is an elliptic divisibility sequence.

3. The sequence

$$\ldots, -2, -1, 0, 1, 2, 1, 1, 7, \frac{27}{2}, 5, -\frac{169}{4}, -\frac{659}{2}, -\frac{4963}{8}, -\frac{5963}{8}, \ldots$$

   is an elliptic sequence.

4. The sequence

$$\ldots, 6, -2, -3, -1, 0, 1, 3, 2, -6, -170, -1044, -712, 181752, \ldots$$

   is an EDS.

5. The sequence

$$\ldots, -2, -1, 0, 1, 2, 4, 1, -56 - 450, -3586, -6736, \ldots$$

is an elliptic sequence and seems to be an integer sequence, but is not an EDS because it does not have the divisibility property (for example, $h_2 = 2$ does not divide $h_4 = 1$).

4.1.2 Generalised EDSs

It is easy to prove that every solution of (4.1) in fact satisfies a more general recurrence relation:

Theorem 4.1.2. *[30]*

If (h_n) is an elliptic sequence then

$$h_{m+n}\, h_{m-n}\, h_t^2 = h_{m+t}\, h_{m-t}\, h_n^2 - h_{n+t}\, h_{n-t}\, h_m^2 \tag{4.2a}$$

for all $m, n, t \in \mathbb{Z}$.

Note that we can write this in a more symmetric form as

$$h_{m+n}\, h_{m-n}\, h_t^2 + h_{n+t}\, h_{n-t}\, h_m^2 + h_{t+m}\, h_{t-m}\, h_n^2 = 0 \tag{4.2b}$$

for all $m, n, t \in \mathbb{Z}$.

We can define a "generalised" elliptic sequence or EDS using this formula, with elliptic sequences and EDSs as a special case with $h_1 = \pm 1$. (Shipsey does this for EDSs in [23].)

Definition: A *generalised elliptic sequence* is an infinite sequence (h_n) of rational numbers satisfying (4.2a). A *generalised elliptic divisibility sequence* is a generalised elliptic sequence (h_n) all of whose terms are integers, with the divisibility property that h_n divides h_m whenever n divides m.

Dividing every term of a generalised elliptic sequence (ℓ_n) by ℓ_1 we get an elliptic sequence (h_n) and, conversely, for any elliptic sequence (h_n) and any rational number c the sequence (ℓ_n) defined by $\ell_n = c\, h_n$ for all $n \in \mathbb{Z}$ is a generalised elliptic sequence. Moreover, if ℓ_1 is an integer then (h_n) is an EDS if and only if (ℓ_n) is a generalised EDS (because of the divisibility property). Hence there is no real advantage in studying the more general sequences over EDSs.

4.1.3 Equivalent elliptic sequences

It is easy to prove the following result.

Theorem 4.1.3. *[30]*

If (h_n) is an elliptic sequence, then for any rational constant θ, the sequence (h'_n) defined by

$$h'_n = \theta^{n^2-1}\, h_n \quad \text{for all } n \in \mathbb{Z}$$

is also an elliptic sequence.

This leads to the following concept of equivalence:

Definition: Two elliptic sequences (h_n) and (h'_n) are said to be *equivalent* if there exists a rational constant θ such that

$$h'_n = \theta^{n^2-1}\, h_n \quad \text{for all } n \in \mathbb{Z}.$$

(This is obviously an equivalence relation in the technical sense.) Again, Ward defines equivalence slightly differently in [30]; in his definition the constant θ does not have to be a rational number. However if $h_2\, h_3 \neq 0$ then since $\theta^3 = \frac{h'_2}{h_2}$ and $\theta^8 = \frac{h'_3}{h_3}$ are rational it follows that $\frac{(\theta^3)^3}{\theta^8} = \theta$ is also rational, so Ward's definition is not really more general unless h_2 or h_3 is zero.

4.2 Lucas sequences and singular sequences

Elliptic sequences are a generalisation of a class of divisibility sequences studied earlier by Edouard Lucas; in fact many of Ward's results about EDSs (in particular Theorems 4.7.1, 4.7.5 and 4.7.6) were prompted by similar results discovered by Lucas for his sequences.

Definition: Let c be a rational number, and let a and b be the roots of the polynomial $x^2 - c\,x + 1$. If $a \neq b$ let (ℓ_n) be the sequence

$$\ell_n = \frac{a^n - b^n}{a - b} \quad \text{for } \ n \in \mathbb{Z}.$$

If $a = b$ we define $\ell_n = n\, a^{n-1}$. Then (ℓ_n) is called a *Lucas sequence* with parameter c.

Note that a and b are in general quadratic irrationalities given by

$$\{a, b\} = \left\{ \frac{c + \sqrt{c^2 - 4}}{2}, \frac{c - \sqrt{c^2 - 4}}{2} \right\},$$

and that

$$a\,b = 1 \quad \text{and} \quad a + b = c.$$

The initial values of (ℓ_n) are

$$\ldots, -c, -1, 0, 1, c, c^2 - 1, c^3 - 2c, \ldots.$$

Furthermore, for $n \geq 1$, we have

$$\begin{aligned} \ell_n &= a^{n-1} + a^{n-2}\,b + a^{n-3}\,b^2 + \ldots + a\,b^{n-2} + b^{n-1} \\ &= a^{n-1} + a^{n-3} + a^{n-5} + \ldots + a^{-(n-3)} + a^{-(n-1)}. \end{aligned}$$

It is easy to prove from this equation that the sequence (ℓ_n) can be generated by the following linear recursion of order two:

$$\ell_n = c\,\ell_{n-1} - \ell_{n-2} \quad \text{for } n \in \mathbb{Z}. \tag{4.3}$$

But Lucas proved that (ℓ_n) also satisfies the elliptic sequence recursion (4.1). Since c is rational, it follows from (4.3) that ℓ_n is rational for $n \in \mathbb{Z}$, and hence that (ℓ_n) is an elliptic sequence. (Lucas sequences are in fact the only elliptic sequences which also satisfy an order 2 linear recursion, as can easily be shown using $h_{-1} = -1$, $h_0 = 0$, $h_1 = 1$.) The Lucas sequence (ℓ_n) is an EDS if and only if c is an integer (this will follow from Theorem 4.4.6). Note that the sequence of integers $\mathbb{Z}$ is a Lucas sequence with $a = b = 1$ (i.e., $c = 2$).

Lucas sequences (with $h_2\,h_3 \neq 0$, i.e., $c \neq -1, 0, 1$) turn out to be a special case of a type of elliptic sequence called a *singular elliptic sequence*:

Definition: Let (h_n) be an elliptic sequence with $h_2\,h_3 \neq 0$. The *discriminant* of (h_n) is

$$\begin{aligned} \Delta(h_2, h_3, h_4) = \frac{1}{h_2^8\,h_3^3}\Big(&-h_4^4 - 3h_2^5\,h_4^3 + \left(-3h_2^{10} - 8h_2^2\,h_3^3\right)\,h_4^2 \\ &+ \left(-h_2^{15} + 20h_2^7\,h_3^3\right)\,h_4 + h_2^{12}\,h_3^3 - 16h_2^4\,h_3^6 \Big). \end{aligned}$$

The sequence (h_n) is said to be *singular* if $\Delta(h_2, h_3, h_4) = 0$, and *singular modulo* p for a prime p if $\Delta(h_2, h_3, h_4) \equiv 0 \bmod p$. Otherwise S is said to be *non-singular*, or *non-singular modulo* p.

(The reason for the name "singular" will become clear in section 4.5.2.) It is easy to prove from the definition that if (h_n) and (h'_n) are equivalent elliptic sequences with $h'_n = \theta^{n^2-1} h_n$ for all $n \in \mathbb{Z}$, then

$$\Delta(h'_2, h'_3, h'_4) = \theta^{12} \Delta(h_2, h_3, h_4),$$

so (h'_n) is singular if and only if (h_n) is.

Ward used diophantine equations to characterise singular EDSs in terms of their initial values as follows:

Theorem 4.2.1. *[30]*

An EDS (h_n) with $h_2 h_3 \neq 0$ is singular if and only if there exist integers r and s such that

$$h_2 = r, \quad h_3 = s(r^2 - s^3), \quad \text{and} \quad h_4 = rs^3(r^2 - 2s^3).$$

He proved further that all Lucas sequences with $h_2 h_3 \neq 0$ are singular:

Theorem 4.2.2. *[30]*

An elliptic sequence (h_n) with $h_2 h_3 \neq 0$ is a Lucas sequence with parameter c if and only if it is a singular solution with $r = c$ and $s = 1$ in Theorem 4.2.1.

For singular EDSs with $s \neq 1$, we still have the following result:

Theorem 4.2.3. *[30]*

Let (h_n) be a singular EDS, and let

$$c = \pm\sqrt{\frac{r^2}{s^3}} \quad \text{and} \quad \theta = \sqrt{s},$$

where r and s are the integers given in Theorem 4.2.1. Let a and b be the roots of the polynomial $x^2 - c\,x + 1$, and let (ℓ_n) be the sequence

$$\ell_n = \frac{a^n - b^n}{a - b} \quad \text{for all } n \in \mathbb{Z}$$

(unless $a = b = \pm 1$, in which case let $\ell_n = n\,a^{n-1}$). Then

$$h_n = \theta^{n^2-1}\,\ell_n \quad \text{for all } n \in \mathbb{Z}.$$

Ward states in [30] that every singular EDS is equivalent to a Lucas sequence, but that is because he allows both c in his definition of Lucas sequences and θ in his definition of equivalence to be irrational. By our definitions, (ℓ_n) is only a Lucas sequence if $c \in \mathbb{Q}$, and (h_n) is only equivalent to (ℓ_n) if $\theta \in \mathbb{Q}$; so our result matches Ward's if and only if s is a perfect square.

4.3 Computing given terms of an elliptic sequence

There are two useful formulae satisfied by elliptic sequences. The "stepping formula" is obtained by setting $n = 2$ in (4.1), and allows us to find each term h_{m+2} from the initial values and the previous four terms (as long as $h_{m-2} \neq 0$):

$$h_{m+2} = \frac{h_{m+1}\, h_{m-1}\, h_2^2 - h_3\, h_1\, h_m^2}{h_{m-2}} \qquad \text{for all } m \in \mathbb{Z}. \tag{4.4}$$

The "doubling formulae" are obtained by setting first $m = k+1$, $n = k$ and then $m = k+1$, $n = k-1$ in (4.1):

$$h_{2k+1}\, h_1 = h_{k+2}\, h_k^3 - h_{k-1}\, h_{k+1}^3 \qquad \text{for all } k \in \mathbb{Z}, \tag{4.5a}$$

$$h_{2k}\, h_2 = h_k \left(h_{k+2}\, h_{k-1}^2 - h_{k-2}\, h_{k+1}^2 \right) \qquad \text{for all } k \in \mathbb{Z}. \tag{4.5b}$$

These formulae allow us to find a given term h_{ak+b} in $\mathrm{O}(\log a + \log b)$ time by a kind of "repeated doubling" if we know the three initial values h_2, h_3, h_4 and four terms surrounding h_k (for details see [23]). Note that we do not have to know the value of k.

4.4 Existence and uniqueness

If h_n is non-zero for all $n \neq 0$ then we can define a unique sequence (h_n) of rational numbers with initial values $h_0 = 0$, $h_1 = 1$, h_2, h_3, h_4 by the recursion (4.4), and if $h_2 \neq 0$ then we can define a unique sequence (h_n) of rational numbers with initial values $h_0 = 0$, $h_1 = 1$, h_2, h_3, h_4 by the recursion (4.5). A surprising result of Ward's [30] is that these two sequences will be same and will be an elliptic sequence. In other words, if (h_n) is a rational sequence with $h_0 = 0$, $h_1 = 1$

and $h_n \neq 0$ for all $n \neq 0$ then either of condition (4.4) or condition (4.5) implies condition (4.1). (Of course, if $h_0 \neq 0$ or $h_1 \neq \pm 1$ then the sequences generated by (4.4) and (4.5) are not elliptic sequences; they are also not necessarily the same.)

Even more surprisingly, if h_2, h_3, h_4 are integers with $h_2 \mid h_4$ then (h_n) will be an EDS — in other words it will consist of integers (despite the fact that computing h_n from (4.4) involves dividing by h_{n-4}), satisfy the recursion (4.1) and have the divisibility property that $h_n \mid h_m$ whenever $n \mid m$. These remarks follow from the existence and uniqueness results (due to Ward) given in this section.

4.4.1 Elliptic sequences with given initial values

We have the following existence theorem for elliptic sequences:

Theorem 4.4.1. *[30]*

For any three rational numbers h_2, h_3 and h_4 there is an elliptic sequence with initial values $h_0 = 0$, $h_1 = 1$, h_2, h_3, h_4, unless h_2 is zero and h_4 is non-zero. Furthermore, if h_2 and h_3 are not both zero then this solution is unique.

The uniqueness part of Theorem 4.4.1 follows easily from the doubling formulae (4.5) if $h_2 \neq 0$, and if $h_2 = 0$ can easily be proved by induction using the elliptic sequence equation (4.1) first with $m = 2k$ and $n = 2$ and then with $m = 2k - 2$ and $n = 3$. Ward proved the existence part of the theorem by finding an elliptic sequence (h_n) with the required initial values, for any choice of h_2, h_3, h_4. The most important case is where neither h_2 nor h_3 is zero:

Theorem 4.4.2. *[30]*

For any rational numbers $h_0 = 0$, $h_1 = 1$, h_2, h_3 and h_4 such that neither h_2 nor h_3 is zero, there exists an elliptic curve $E/\mathbb{Q}$ and a non-singular rational point $P = (x, y)$ on E such that the sequence $(h_n) = (\psi_n(x, y))$ of division polynomials evaluated at P is an elliptic sequence with initial values h_0, h_1, h_2, h_3, h_4.

Ward also showed how E and P can be calculated from h_2, h_3 and h_4 (see Theorem 4.5.1); we will say more about the connection between elliptic sequences and elliptic curves in section 4.5.

The cases where exactly one of h_2 and h_3 is zero are covered by the next two results of Ward's:

Theorem 4.4.3. *[30]*

A rational sequence (h_n) with $h_0 = 0$, $h_1 = 1$, $h_2 = 0$ and $h_3 \neq 0$ is an elliptic sequence if and only if

$$h_n = \begin{cases} 0 & \text{if } n = 2k, \\ (-1)^{\frac{1}{2}k(k-1)}\, h_3^{\frac{1}{2}k(k+1)} & \text{if } n = 2k+1. \end{cases}$$

Note that an elliptic sequence with $h_2 = 0$ must have $h_4 = 0$.

Theorem 4.4.4. *[30]*

A rational sequence (h_n) with $h_0 = 0$, $h_1 = 1$, $h_2 \neq 0$ and $h_3 = 0$ is an elliptic sequence if and only if

$$h_n = \begin{cases} 0 & \text{if } n = 3k, \\ (-h_2)^{\frac{1}{2}k(k-1)}\, h_4^{\frac{1}{2}k(k+1)} & \text{if } n = 3k+1, \\ -(-h_2)^{\frac{1}{2}(k+1)(k+2)}\, h_4^{\frac{1}{2}k(k+1)} & \text{if } n = 3k+2. \end{cases}$$

The case where h_2 and h_3 are both zero is somewhat different:

Theorem 4.4.5. *[30]*

A rational sequence (h_n) with $h_0 = 0$, $h_1 = 1$ and $h_2 = h_3 = 0$ is an elliptic sequence if and only if, for some integer a which is either even or 1,

$$h_1 = 1,\ h_{-1} = -1,\ h_a \neq 0,\ h_{-a} = -h_a \quad \text{and all other} \quad h_n = 0,$$

or for some odd integer $\ell > 3$,

$$h_n = \begin{cases} 0 & \text{if } n \not\equiv \pm 1 \bmod \ell, \\ (-h_{\ell-1})^{\frac{1}{2}k(k-1)}\, (h_{\ell+1})^{\frac{1}{2}k(k+1)} & \text{if } n = k\ell + 1, \\ -(-h_{\ell-1})^{\frac{1}{2}k(k+1)}\, (h_{\ell+1})^{\frac{1}{2}k(k-1)} & \text{if } n = k\ell - 1. \end{cases}$$

So an elliptic sequence (h_n) with $h_2 = h_3 = 0$ is uniquely determined by a and h_a, or by ℓ, $h_{\ell-1}$ and $h_{\ell+1}$, but not by any fixed number of initial values.

Remark: Ward states that every elliptic sequence (h_n) with $h_1 = 1$, $h_2 = 0$ and $h_3 \neq 0$ is equivalent to the Lucas sequence (with $c = 0$)

$$\ldots, 0, -1, 0, 1, 0, -1, \ldots$$

This is true by Ward's definition of equivalence, i.e., if we allow $\theta = h_3^{\frac{1}{8}}$ (which might not be a rational number). It is only true by our definition if h_3 is an 8th power.

4.4.2 Which elliptic sequences are EDSs

Theorem 4.4.1 says that for any rational numbers h_2, h_3, h_4 (except when $h_2 = 0$, $h_4 \neq 0$) there exists an elliptic sequence with initial values $h_0 = 0, h_1 = 1, h_2, h_3, h_4$. The next theorem shows when this sequence is an elliptic divisibility sequence.

Theorem 4.4.6. *[30]*

Let (h_n) be an elliptic sequence in which the initial values $h_0, h_1, \ldots, h_4$ are integers such that $h_0 = 0$, $h_1 = 1$, h_2 and h_3 are not both zero and h_2 divides h_4. Then (h_n) is an elliptic divisibility sequence (i.e., all terms are integers and h_n divides h_m whenever n divides m).

It can easily be proved using the doubling formula that if h_2, h_3 and h_4 are p-integers and $h_2 \not\equiv 0 \bmod p$, then all terms of (h_n) are p-integers:

Theorem 4.4.7. *If (h_n) is an elliptic sequence then any prime appearing in the denominator of some term in (h_n) must also divide the numerator of h_2 or the denominator of h_2, h_3 or h_4.*

If $h_2 = h_3 = 0$ then an elliptic sequence (h_n) is an elliptic divisibility sequence only if $h_{2k} = 0$ for all $k \in \mathbb{Z}$, in which case by Theorem 4.4.5 (h_n) is the trivial sequence or has $h_1 = \pm 1$, $h_{-1} = -h_1$ and all other terms zero. Theorems 4.4.1, 4.4.5 and 4.4.6 therefore lead to the following existence theorem for EDSs:

Theorem 4.4.8. *For any integers h_2, h_3 and h_4 such that h_2 divides h_4, there exists a unique elliptic divisibility sequence (h_n) with initial values $h_0 = 0$, $h_1 = 1$, h_2, h_3, h_4.*

4.4.3 Almost every elliptic sequence is equivalent to an EDS

It turns out that elliptic sequences (h_n) in which h_2 and h_3 are not both zero are not essentially more general than EDSs:

Theorem 4.4.9. *[30]*

Every elliptic sequence (h_n) in which h_2 and h_3 are not both zero is equivalent to an elliptic divisibility sequence.

Proof: If $h_2 \neq 0$ and the lcm of the denominators of h_2, h_3 and h_4 is a, then the sequence (h'_n) defined by $h'_n = (h_2\, a^2)^{n^2-1}\, h_n$ for all $n \in \mathbb{Z}$ is an EDS by Theorem 4.4.6, since h'_2, h'_3 and h'_4 are integers with h'_2 dividing h'_4. Similarly, if $h_2 = 0$ and h_3 has denominator a then the sequence defined by $h'_n = a^{n^2-1}\, h_n$ for all $n \in \mathbb{Z}$ is an EDS. □

If $h_2 = h_3 = 0$ then there do exist elliptic sequences which are not equivalent to an EDS, but these are determined completely by Theorem 4.4.5. Theorem 4.4.9 therefore means that there is no real advantage in studying elliptic sequences over elliptic divisibility sequences.

4.5 The relationship between elliptic sequences and elliptic curves

In this section we go into more detail about the connection between elliptic sequences and elliptic curves.

By Theorem 3.10.9 the division polynomials ψ_n of an elliptic curve $E/\mathbb{Q}$ satisfy the recursion (3.23) in $\mathbb{Q}[x, y]$, and it follows that for any rational point $P = (x_1, y_1)$ on E, if $h_n = \psi_n(x_1, y_1)$ for $n \in \mathbb{Z}$ then the sequence (h_n) is an elliptic sequence.

Ward's Theorem 4.4.2 together with the uniqueness of elliptic sequences supplies the converse, that for every elliptic sequence (h_n) in which neither h_2

nor h_3 is zero there exists an elliptic curve $E/\mathbb{Q}$ and a rational point $P = (x_1, y_1)$ on E such that the sequence of division polynomials of E evaluated at P is (h_n):

Theorem 4.5.1. *[30]*

Let (h_n) be an elliptic sequence in which neither h_2 nor h_3 is zero. Then there exists an elliptic curve

$$E : y^2 = x^3 + c_4\, x + c_6,$$

where $c_4, c_6 \in \mathbb{Q}$, and a non-singular rational point $P = (x_1, y_1)$ on E such that

$$\psi_n(x_1, y_1) = h_n \quad \text{for all } n \in \mathbb{Z},$$

where ψ_n is the nth division polynomial of E.

Specifically, c_4, c_6 and P are given by the following rational functions of h_2, h_3, h_4:

$$\begin{aligned} c_4 = -\frac{1}{2^4\, 3\, h_2^8\, h_3^4}\Big(& h_2^{20} + 4h_2^{15}\, h_4 - 16h_2^{12}\, h_3^3 + 6h_2^{10}\, h_4^2 - 8h_2^7\, h_3^3\, h_4 \\ & + 4h_2^5\, h_4^3 + 16h_2^4\, h_3^6 + 8h_2^2\, h_3^3\, h_4^2 + h_4^4\Big), \\ c_6 = \frac{2}{2^5\, 3^3\, h_2^{12}\, h_3^6}\Big(& h_2^{30} + 6h_2^{25}\, h_4 - 24h_2^{22}\, h_3^3 + 15h_2^{20}\, h_4^2 - 60h_2^{17}\, h_3^3\, h_4 \\ & + 20h_2^{15}\, h_4^3 + 120h_2^{14}\, h_3^6 - 36h_2^{12}\, h_3^3 h_4^2 \\ & + 15h_2^{10}\, h_4^4 - 48h_2^9\, h_3^6 h_4 + 12h_2^7\, h_3^3\, h_4^3 \\ & + 64h_2^6\, h_3^9 + 6h_2^5\, h_4^5 + 48h_2^4\, h_3^6\, h_4^2 + 12h_2^2\, h_3^3\, h_4^4 + h_4^6\Big), \end{aligned}$$

and

$$(x_1, y_1) = \left(\left(\frac{h_4 + h_2^5}{h_2^2\, h_3}\right)^2 + \frac{h_3}{3\, h_2^2},\ \frac{1}{2}\, h_2\right).$$

Remarks:

1. Ward actually formulated his theorem with a curve of the classical form

   $$E : y^2 = 4x^3 - g_2 x - g_3,$$

   but this curve can clearly be transformed to the more modern form $y^2 = x^3 + c_4\, x + c_6$ by dividing the curve equation by 4 and renaming $-\frac{1}{2}y$ as y, $-\frac{g_2}{4}$ as c_4 and $-\frac{g_3}{4}$ as c_6.

2. Ward also proved that if (h_n) is an elliptic sequence with $h_3 = 0$ then (h_n) is equal to the sequence of division polynomials of some elliptic curve $E/\mathbb{Q}$ of the form $y^2 = x^3 + c_4\,x + c_6$ evaluated at some rational point P if and only if $h_4 = -h_2^5$. (Then P has order 3 in $E(\mathbb{Q})$.)

3. Note that even if E is a singular curve, P is not a singular point. This is because a singular point on E would have to have y-coordinate 0, while P has y-coordinate $\frac{1}{2}\,h_2 \neq 0$.

Rachel Shipsey [23] found an alternative and much simpler formula for an elliptic curve E, containing the non-singular point $P = (0, 0)$, such that the sequence $(\psi_n(0, 0))$ of division polynomials evaluated at P is (h_n). Her formula can be obtained from Ward's by a linear transformation in which the point P is moved to the origin. In both Ward's and Shipsey's formulae the factor 2 appears in the denominator of some coefficient, so the curve cannot be reduced modulo 2, and in Ward's formula this is also true of the factor 3. Here we prove a similar result, using the division polynomial results from section 3.10. We look for a slightly more general formula in which the coefficients can be chosen to have denominators coprime to any given prime not dividing $h_2\,h_3$.

Definition: The elliptic curves over $\mathbb{Q}$ in which $(0, 0)$ is a non-singular point and

$$\psi_n(0, 0) = h_n \quad \text{for all } n \in \mathbb{Z}$$

are said to be *associated with* the elliptic sequence (h_n).

Theorem 4.5.2. *Let (h_n) be an elliptic sequence in which neither h_2 nor h_3 is zero. Then the elliptic curves*

$$E : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x$$

associated with (h_n) *are precisely those with*

$$\begin{aligned} a_3 &= h_2 \\ b_8 &= h_3 \\ b_4 &= \frac{h_4 + h_2^5}{h_2\, h_3} \qquad \textit{and} \\ b_2 &= \frac{b_4^2 + 4h_3}{h_2^2} \end{aligned}$$

where

$$b_4 = a_1\, a_3 + 2\, a_4, \quad b_6 = a_3^2 \qquad \textit{and} \qquad b_8 = -a_1\, a_3\, a_4 + a_2\, a_3^2 - a_4^2.$$

Proof: By the uniqueness part of Theorem 4.4.1, we just need to find coefficients $a_i \in \mathbb{Q}$ such that the initial values of $(\psi_n(0,0))$ match the initial values of (h_n), i.e., (using Theorem 3.23),

$$\begin{aligned} \psi_2(0,0) &= a_3 = h_2 \\ \psi_3(0,0) &= b_8 = h_3 \qquad \text{and} \\ \psi_4(0,0) &= (b_4\, b_8 - b_6^2)\, h_2 = h_4. \end{aligned}$$

So we need $a_3 = h_2$ and $b_8 = h_3$. Since $b_6 = a_3^2 = h_2^4$, we then have

$$b_4 = \frac{h_4 + b_6^2\, h_2}{b_8\, h_2} = \frac{h_4 + h_2^5}{h_2\, h_3}.$$

Since $b_4^2 + 4b_8 = b_2\, b_6$ by (3.4), it follows that

$$b_2 = \frac{b_4^2 + 4h_3}{h_2^2}.$$

Finally, we note that (x, y) is a singular point on E if and only if the partial derivatives

$$3x^2 + 2a_2x - a_1y + a_4 \qquad \text{and} \qquad -2y - a_1x - a_3$$

are both zero. So $P = (0,0)$ is a singular point on E if and only if

$$a_4 = a_3 = 0.$$

But $a_3 = h_2 \neq 0$, so P is non-singular. This completes the proof. □

Equivalently, we have

Theorem 4.5.3. *Let (h_n) be an elliptic sequence in which neither h_2 nor h_3 is zero. Then the (possibly singular) elliptic curves*

$$E : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x$$

associated with (h_n) are precisely those with

$$\begin{aligned} &a_4 \text{ arbitrary} \\ &a_3 = h_2 \\ &a_1 = \frac{h_4 + h_2^5 - 2\,h_2\,h_3\,a_4}{h_2^2\,h_3} \qquad \text{and} \\ &a_2 = \frac{h_2\,h_3^2 + (h_4 + h_2^5)a_4 - h_2\,h_3\,a_4^2}{h_2^3\,h_3}. \end{aligned}$$

Proof: By Theorem 4.5.2, $a_3 = h_2$, so for any choice of a_4 we can write

$$\begin{aligned} a_1 &= \frac{b_4 - 2a_4}{a_3} \\ &= \frac{1}{h_2}\left(\frac{h_4 + h_2^5}{h_2\,h_3} - 2a_4\right) \end{aligned}$$

and

$$\begin{aligned} a_2 &= \frac{b_8 + a_1\,a_3\,a_4 + a_4^2}{a_3^2} \\ &= \frac{1}{h_2^2}\left(h_3 + \frac{1}{h_2}\left(\frac{h_4 + h_2^5}{h_2\,h_3} - 2a_4\right)\,h_2\,a_4 + a_4^2\right). \end{aligned}$$

□

Corollary 4.5.4. *Let (h_n) be an elliptic sequence, and E an associated elliptic curve. Then (h_n) is an EDS if and only if E has integer values of a_3, b_8 and $b_4\,b_8$.*

Proof: By Theorem 4.4.6, (h_n) is an EDS if and only if h_2, h_3 and $\frac{h_4}{h_2}$ are integers. Since

$$h_2 = a_3, \quad h_3 = b_8, \qquad \text{and} \qquad \frac{h_4}{h_2} = \left(b_4\,b_8 - a_3^4\right),$$

this is true if and only if a_3, b_8 and $b_4\,b_8$ are integers. □

Remarks:

1. We could have let P be a general point (x_1, y_1) instead of $(0, 0)$, but since by Theorem 3.10.7 the linear transformation

$$x = x' + x_1, \quad y = y' + y_1$$

moves (x_1, y_1) to the origin without affecting the values of $h_n = \psi_n(x_1, y_1) = \psi'_n(0, 0)$, the general case is not essentially different from the case $P = (0, 0)$.

2. Our curves in Theorem 4.5.3 are birationally equivalent over $\mathbb{Q}$ to Ward's in Theorem 4.5.1, with his point (x_1, y_1) mapped to our point $(0, 0)$. The change of variables is

$$x = u^2\, x' + r, \qquad y = u^3\, y' + u^2\, s\, x' + t,$$

where

$$u = 1, \quad r = \frac{b_2}{12} = x_1, \quad s = \frac{a_1}{2} \quad \text{and} \quad t = \frac{a_3}{2} = y_1.$$

3. Recall that by Theorem 3.10.8 the point $P = (0, 0)$ is singular if and only if $\psi_n(0, 0) = 0$ for all $|n| > 1$.

4. We cannot necessarily choose a_4 so that the a_i are all integers. For instance, if the h_n are integers, h_2 and h_3 are coprime and a_4 is an integer, then a_1 is an integer only if h_3 divides $h_4 + h_2^5$.

5. We can also modify the above proof to prove that if $h_3 = 0$ then (h_n) has an associated elliptic curve if and only if $h_4 = -h_2^5$. The coefficients of E satisfy $a_3 = h_2$ and $b_8 = -a_1\, a_3\, a_4 + a_2\, a_3^2 - a_4^2 = 0$.

6. Similarly, we can show that if $h_2 = 0$ then (h_n) has an associated elliptic curve if and only if $-h_3$ is a square. The coefficients of E satisfy $a_3 = 0$ and $b_8 = -a_4^2 = h_3$; a_1 and a_2 are arbitrary.

4.5.1 Equivalent sequences and curves

Ward proved in [30] that birationally equivalent elliptic curves are associated with equivalent elliptic sequences:

Theorem 4.5.5. *Let (h_n) and (h'_n) be elliptic sequences in which $h_2\,h_3$ and $h'_2\,h'_3$ are non-zero, and let E and E' be associated elliptic curves from Theorem 4.5.3. Then (h'_n) is equivalent to (h_n) under*

$$h'_n = \theta^{n^2-1}\,h_n \qquad \text{for all } n \in \mathbb{Z}$$

if and only if E' is birationally equivalent to E, with $(0,0)$ mapping to $(0,0)$, under the admissible change of variables

$$x = u^2x' \qquad \text{and} \qquad y = u^3y' + u^2sx', \tag{4.6}$$

for $u = \frac{1}{\theta}$ and some $s \in \mathbb{Q}$.

Proof: Suppose E and E' are equivalent under the admissible change of variables (4.6), and let $\theta = \frac{1}{u}$. By Theorem 3.3.3, since $r = t = 0$ we have $a'_3 = \theta^3\,a_3$ and $b'_j = \theta^j\,b_j$ for all j. So by Theorem 3.10.3 the initial values of $(h'_n) = (\psi'_n(x'_1, y'_1))$ are

$$\begin{aligned}
h'_2 &= a'_3 = \theta^3\,a_3 = \theta^{2^2-1}\,h_2,\\
h'_3 &= b'_8 = \theta^8\,b_8 = \theta^{3^2-1}\,h_3, \qquad \text{and}\\
h'_4 &= \left(b'_4\,b'_8 - {b'_6}^2\right) h'_2 = \theta^{12}\left(b_4\,b_8 - b_6^2\right)\theta^3\,h_2 = \theta^{4^2-1}\,h_4.
\end{aligned}$$

It follows from the uniqueness part of Theorem 4.4.1 that $h'_n = \theta^{n^2-1}\,h_n$ for all $n \in \mathbb{Z}$, and hence that (h_n) and (h'_n) are equivalent.

For the converse, suppose (h_n) and (h'_n) are equivalent under $h'_n = \theta^{n^2-1}\,h_n$ for all $n \in \mathbb{Z}$. Then by Theorem 4.5.2,

$$\begin{aligned}
b'_6 &= {h'_2}^2 = \left(\theta^3\,h_2\right)^2 = \theta^6\,b_6,\\
b'_8 &= h'_3 = \theta^8\,h_3 = \theta^8\,b_8,\\
b'_4 &= \frac{h'_4 + {h'_2}^5}{h'_2\,h'_3} = \frac{\theta^{15}\,h_4 + \left(\theta^3\,h_2\right)^5}{\left(\theta^3\,h_2\right)\left(\theta^8\,h_3\right)} = \theta^4\,b_4,
\end{aligned}$$

and

$$b'_2 \;=\; \frac{{b'_4}^2 + 4h'_3}{{h'_2}^2} \;=\; \frac{\theta^4\,{b_4}^2 + 4\theta^8\,h'_3}{{h'_2}^2} \;=\; \theta^8\,b_8.$$

So $\left(\frac{1}{\theta}\right)^j\, b'_j \;=\; b_j$ for $j \;=\; 2, 4, 6, 8$, and it follows by Theorem 3.3.4 that E' is birationally equivalent to E under an admissible change of variables (3.8) with $u = \frac{1}{\theta}$ and $r = 0$. □

Remarks:

1. If $\theta = 1$ (i.e., (h_n) and (h'_n) are the same sequence) then choosing a different value a'_4 for a_4 in Theorem 4.5.3 gives a curve E', also associated with (h_n), which is birationally equivalent to E with $u = 1$, $r = t = 0$ and $s = \frac{a_4 - a'_4}{h_2}$.

   Equivalently, if $u = 1$ then, for any value of s, the associated sequence (h'_n) of E' is the same as (h_n) (changing s corresponds to choosing a different value for a_4 in Theorem 4.5.3). For any other value of u, (h_n) is an equivalent sequence.

2. Recall that (by Theorem 3.10.2) if $P = (0, 0)$ and for all $n \in \mathbb{Z}$ we denote the x-coordinate of $[n]P$ by x_n then we have

   $$x_n = -\frac{h_{n-1}\,h_{n+1}}{h_n^2} \qquad \text{for all } n \in \mathbb{Z}.$$

   Theorem 4.5.5 says

   $$x'_n = -\frac{h'_{n-1}\,h'_{n+1}}{{h'_n}^2} = -\frac{\theta^{(n-1)^2-1}\,h_{n-1}\cdot\theta^{(n+1)^2-1}\,h_{n+1}}{\left(\theta^{n^2-1}\,h_n\right)^2} = -\theta^2\cdot\frac{h_{n-1}\,h_{n+1}}{h_n^2} = \theta^2\,x_n,$$

   as we would expect.

   Finally, we have

Theorem 4.5.6. *Let E be an elliptic curve over $\mathbb{Q}$, let P be a rational point on E, and let $P' = [k]P = (x_k, y_k)$ for some $k \in \mathbb{N}$. For $n \in \mathbb{Z}$ let $h_n = \psi_n(x, y)$ and $h'_n = \psi_n(x_k, y_k)$. Then*

$$\frac{h_{kn}}{h_k} = h_k^{n^2-1}\,h'_n \qquad \textit{for all } n \in \mathbb{Z},$$

i.e., (h'_n) is equivalent to the elliptic sequence obtained by taking every kth term of (h_n) and dividing by h_k.

Proof: It follows from Theorem 3.10.11 that

$$\frac{\psi_{kn}(x,y)}{\psi_k(x,y)} = \Big(\psi_k(x,y)\Big)^{n^2-1}\,\psi_n(x_k,y_k) \quad \text{for all } n \in \mathbb{Z},$$

i.e., that

$$\frac{h_{kn}}{h_k} = h_k^{n^2-1}\,h'_n \quad \text{for all } n \in \mathbb{Z}.$$

Hence, by Theorem 4.1.3, the sequence obtained by taking every kth term of (h_n) and dividing by h_k is an elliptic sequence, equivalent to (h'_n). □

Remark: If $P = (0,0)$ this is the same as considering a birational transformation in which the point $[k]P$ is moved to the origin (i.e., $u = 1$, $s = 0$, $r = x_k$ and $t = y_k$), since then by Theorem 3.10.7

$$\psi'_n(0,0) = \psi_n(x_k,y_k) \quad \text{for all } n \in \mathbb{Z},$$

i.e., the values of the division polynomials evaluated at $[k]P$ are preserved by this transformation.

4.5.2 Singular sequences and curves

Ward found conditions on the initial values of (h_n) for which the associated elliptic curves are singular:

Theorem 4.5.7. *[30]*

Let (h_n) be an elliptic sequence in which $h_2\,h_3 \neq 0$, and let E be an associated elliptic curve from Theorem 4.5.3. Then the discriminant of the curve E is equal to the discriminant of the sequence (h_n), so E is a singular curve if and only if (h_n) is a singular sequence.

Proof: The result is proved simply by substituting the expressions for the b_i in terms of h_2, h_3, h_4 given in Theorem 4.5.2 into the discriminant

$$\Delta = -b_2^2\,b_8 - 8b_4^3 - 27b_6^2 + 9b_2\,b_4\,b_6$$

of the curve E, and rearranging till we get

$$\Delta = \frac{1}{h_2^8\,h_3^3}\Big(-h_4^4 - 3h_2^5\,h_4^3 + \left(-3h_2^{10} - 8h_2^2\,h_3^3\right)\,h_4^2 + \left(-h_2^{15} + 20h_2^7\,h_3^3\right)\,h_4 + h_2^{12}\,h_3^3 - 16h_2^4\,h_3^6\Big),$$

which is the discriminant of the sequence (h_n). $\square$

Example: If (h_n) is the sequence of integers $\mathbb{Z}$ (which we know from section 4.2 is a singular elliptic sequence), then the associated elliptic curves in Theorem 4.5.2 have

$$\begin{aligned} b_8 &= 3, \\ b_6 &= 2^2 = 4, \\ b_4 &= \frac{4 + 2^5}{(2)(3)} = 6 \quad \text{and} \\ b_2 &= \frac{6^2 + 4(3)}{2^2} = 12. \end{aligned}$$

So the discriminant of E is $\Delta = -b_2^2\, b_8 - 8b_4^3 - 27b_6^2 + 9b_2\, b_4\, b_6 = -(12^2)(3) - 8(6^3) - 27(4^2) + 9(12)(6)(4) = 0$, and hence E is singular.

4.5.3 Elliptic sequences and curves reduced modulo a prime power

In this thesis we are interested in the properties of an EDS (h_n) reduced modulo a prime power p^r. Since these are of course related to the properties of an associated elliptic curve reduced modulo p^r, we now prove that if $p \nmid h_2\, h_3$ then we can always find such a reduced curve.

Theorem 4.5.8. *Let (h_n) be an elliptic sequence with $h_2\, h_3 \neq 0$, and let $E/\mathbb{Q}$ be an associated elliptic curve. Let p be a prime such that the coefficients a_i of E are p-integers. Then the h_n are p-integers, and the curve $E \bmod p$ is singular if and only if the sequence (h_n) is singular modulo p. The point $P = (0, 0)$ is singular modulo p if and only if $h_3 \equiv h_4 \equiv 0 \bmod p$, and then $h_n \equiv 0 \bmod p$ for all $|n| > 1$.*

Proof: We have $h_n = \psi_n(0, 0)$ for all $n \in \mathbb{Z}$. The coefficients a_i of E are p-integers, so we may consider the curve $E \bmod p$. By Theorem 4.5.7, the discriminant of (h_n) is equal to the discriminant of E, so $E \bmod p$ is singular over $\mathbb{Z}_p$ if and only if $p \mid \Delta(h_2, h_3, h_4)$, i.e., if and only if (h_n) is singular modulo p.

Note that, by Theorem 3.10.4, the h_n are p-integers since the a_i are. Also, by Theorem 3.10.8 with $K = \mathbb{F}_p$, the point $P = (0, 0)$ is singular modulo p if and only if $h_3 \equiv h_4 \equiv 0 \bmod p$, and if this is true then $h_n \equiv 0 \bmod p$ whenever $|n| > 1$. □

Remark: If (h_n) is an elliptic sequence in which $h_2\, h_3 \neq 0$ and p is a prime which is coprime to h_2 and divides both h_3 and h_4, then the coefficients a_i of the associated elliptic curve E won't all be p-integers (for example, if a_4 is a p-integer then $a_1 = \frac{h_4 + h_2^5 - 2\, h_2\, h_3\, a_4}{h_2^2\, h_3}$ is not a p-integer).

As a consequence of Theorem 4.5.8, we have the following result.

Theorem 4.5.9. *Let (h_n) be an elliptic sequence, and let p be a prime such that h_2, h_3 and h_4 are p-integers and h_2 and h_3 are not divisible by p. Then (h_n) has an associated elliptic curve $E/\mathbb{Q}$ whose coefficients a_i are p-integers. The curve $E \bmod p$ is singular if and only if the sequence (h_n) is singular modulo p, but in either case the point $P = (0, 0)$ is non-singular modulo p.*

Proof: Let

$$E : y^2 + a_1\, x\, y + a_3\, y = x^3 + a_2\, x^2 + a_4\, x$$

be an associated elliptic curve from Theorem 4.5.3, containing the point $P = (0, 0)$ and chosen so that a_4 is a p-integer. Then the coefficients a_i of E are p-integers, and since h_3 is coprime to p the result now follows from Theorem 4.5.8. □

If p is a prime and (h_n) is an elliptic sequence in which h_2, h_3 and h_4 are p-integers and $p \nmid h_2\, h_3$, then by Theorem 4.4.7 every term of (h_n) is a p-integer, so we may consider the sequence $(h_n \bmod p^r)$. Theorem 4.5.9 has several consequences for $(h_n \bmod p^r)$, of which we briefly mention two. Firstly, by Theorem 3.10.5, since $h_n = \psi_n(0, 0)$ for all $n \in \mathbb{Z}$ and P is non-singular modulo p, we have

$$h_n \equiv 0 \bmod p^r \;\Leftrightarrow\; [n]P \equiv \mathcal{O} \bmod p^r \;\Leftrightarrow\; n \equiv 0 \bmod N_r,$$

where N_r is the order of $P = (0, 0)$ in $E \bmod p^r$. So the multiples of p^r in (h_n) are regularly spaced. We describe the pattern of zeroes modulo a prime power in an EDS more fully in section 4.7.

Secondly, if we denote $[n]P$ by (x_n, y_n) for all $n \in \mathbb{Z}$, then by Theorem 3.10.2,

$$x_n = -\frac{h_{n-1}\, h_{n+1}}{h_n^2} \quad \text{for all } n \in \mathbb{Z}.$$

Since $P \bmod p^r$ has order N_r in $E(\mathbb{Z}_{p^r})$, the sequence $(x_n \bmod p^r)$ is periodic with period N_r, i.e.,

$$x_{n+N_r} \equiv x_n \bmod p^r \quad \text{for all } n \in \mathbb{Z}.$$

It follows that

$$\frac{h_{N_r+n-1}\, h_{N_r+n+1}}{h_{N_r+n}^2} \equiv \frac{h_{n-1}\, h_{n+1}}{h_n^2} \bmod p^r \quad \text{for all } n \in \mathbb{Z}.$$

It is therefore not surprising to find that the reduced sequence $(h_n \bmod p^r)$ is periodic, and that it has certain symmetries; we develop this theory in chapter 5.

Note that the coefficients a_i of E might not be $\{h_2, h_3\}$-integers, so the consequences of the elliptic curve representation of elliptic sequences do not necessarily apply to primes dividing h_2 or h_3.

4.5.4 Shipsey's Z-sequence

Finally, we explain how the "Z- sequence" of weighted Z-coordinates of the sequence of points $[n]P$ considered by Shipsey in [23] relates to the sequence of division polynomials evaluated at $(0, 0)$.

Let

$$E : y^2 + a_1 x\, y + a_3\, y = x^3 + a_2\, x^2 + a_4\, x$$

be an elliptic curve with *integer* coefficients a_i, and let $P = (0, 0)$. For $n \in \mathbb{Z}$ let $h_n = \psi_n(0, 0)$, where ψ_n is the nth division polynomial of E. Since the a_i are integers, it follows from Theorems 3.10.9, 3.10.4 and 3.10.10 on the division polynomials that (h_n) is an EDS.

The fact that the a_i are integers also means that, by Theorem 3.7.3, the multiples $[n]P \neq \mathcal{O}$ of P can be written as

$$[n]P = (x_n, y_n) = \left(\frac{X_n}{Z_n^2}, \frac{Y_n}{Z_n^3}\right) \quad \text{for } n \in \mathbb{Z},$$

where X_n, Y_n, Z_n are integers (unique up to the choice of sign of Z_n and Y_n) and X_n and Y_n are coprime to Z_n. (If $[n]P = \mathcal{O}$ then we take $Z_n = 0$.)

Shipsey studied the sequence (Z_n) in her thesis [23]. She proved by an inductive argument that if a_3 and a_4 are coprime then (Z_n) is an EDS in which Z_3 and Z_4 are coprime, and $X_n = -Z_{n-1}\,Z_{n+1}$ for $n \in \mathbb{Z}$. She also proved the converse, that if (Z_n) is any EDS in which h_3 and h_4 are coprime then there exists an elliptic curve E over $\mathbb{Q}$ (the a_i are not necessarily integers) containing the point $P = (0, 0)$, such that the Z-coordinates of $[n]P$ form the EDS (Z_n).

This is explained by the fact that in the case Shipsey was looking at (where the a_i are integers with a_3 and a_4 coprime), the sequence (Z_n) of Z-coordinates of the points $[n](0, 0)$ is the same as the sequence (h_n) of division polynomials evaluated at P. In order to prove this we need to use the result that every two adjacent terms of an EDS (h_n) are coprime if and only if h_3 and h_4 are coprime (see Theorem 4.7.1); this is proved in [31] without using the connection between elliptic sequences and elliptic curves.

Theorem 4.5.10. *Let the a_i be integers. If a_3 and a_4 are coprime then for all $n \in \mathbb{Z}$ we can choose the sign of Z_n so that $Z_n = h_n$. Otherwise $|Z_n| \neq h_n$ at least for $n = 3$.*

Proof: Since the a_i are integers, (h_n) is an EDS. By Theorem 3.10.2,

$$x_n = -\frac{h_{n-1}\,h_{n+1}}{h_n^2} \quad \text{whenever } h_n \neq 0.$$

Since the h_i are integers, it follows by definition of (Z_n) that, whenever $h_n \neq 0$, $Z_n = \pm h_n$ if and only if h_n is coprime to $h_{n-1}\,h_{n+1}$. By Theorem 4.7.1 this is true if and only if h_3 and h_4 are coprime. But by Theorem 4.5.8 (since the coefficients a_i of E are p-integers for every prime p) h_3 and h_4 are coprime if and only if the point $P = (0, 0)$ is non-singular modulo every prime p, i.e., if and only if a_3 and a_4 are coprime. □

Remarks:

1. If the a_i are integers with a_3 and a_4 *not* coprime, then $(h_n) = (\psi_n(0, 0))$ is still an EDS, but $(Z_n) \neq (\pm h_n)$.

2. Shipsey proves in [23] that for any elliptic curve E with integer coefficients, it is always possible to find a birationally equivalent curve E' whose coefficients a'_i are integers with a'_3 and a'_4 coprime and $a'_6 = 0$. (Hence $Z'_n = \psi'_n(0,0)$ for all $n \in \mathbb{Z}$.) Specifically, E' is the curve obtained by moving the point $[M]P$ to the origin for some integer M (i.e., by the admissible change of variables with $u = 1$, $s = 0$, $r = x_M$, $t = y_M$). Let $h'_n = \psi'_n(0,0) = \psi_n(x_M, y_M)$ for all $n \in \mathbb{Z}$. We proved in Theorem 4.5.6 that the sequence (h'_n) is related to the sequence $(h_n) = (\psi_n(0,0))$ by

$$\frac{h_{nM}}{h_M} = {h_M}^{n^2-1}\, h'_n \quad \text{for all } n \in \mathbb{Z},$$

i.e., that (h'_n) is equivalent to the elliptic sequence obtained by taking every Mth term of (h_n) and dividing by h_M. Shipsey's result means that (h'_n) has h'_3 and h'_4 coprime, and is therefore equal to the sequence of Z-coordinates of the points $[n](0,0)$ in E'. It follows that every elliptic curve E is birationally equivalent to one E' for which the sequence (Z'_n) is an EDS and equal to the sequence $(\psi'_n(0,0))$.

3. Finally, it should be noted that Shipsey actually proves that the sequence (Z_n) satisfies the recursion

$$Z_{m+2}\, Z_{m-2} = Z_{m+1}\, Z_{m-1}\, Z_2^2 - Z_1\, Z_3\, Z_m^2 \quad \text{for all } m \in \mathbb{Z},$$

and concludes from this that (Z_n) is an EDS. Since this conclusion is based on Ward's existence result Theorem 4.4.8, her proof of the connection between elliptic sequences and elliptic curves relies implicitly on Ward's.

4.6 Basic properties of elliptic divisibility sequences

In this section we describe some basic properties of EDSs, and in the next we focus on properties of EDSs reduced modulo a prime power. Wherever possible we explain how the properties of an EDS relate to those of the associated elliptic curves.

An interesting result from [31] is that if (h_n) is an EDS, then the sequence obtained from (h_n) by taking every kth term from h_0 and dividing by h_k is also an EDS:

Theorem 4.6.1. *[31]*

Let (h_n) be an elliptic sequence and h_k any non-zero term of (h_n). If

$$\ell_n = \frac{h_{nk}}{h_k} \quad \text{for all } n \in \mathbb{Z}$$

then (ℓ_n) is also an elliptic sequence. Furthermore, if (h_n) is an EDS then (ℓ_n) is an EDS, and if p is a prime dividing both ℓ_3 and ℓ_4 then p divides both h_3 and h_4.

This is proved by elementary means in [31], but also follows from the connection with elliptic curves (see Theorem 4.5.6). (The significance of p dividing both h_3 and h_4 will become clear in the next section.)

In fact we can prove that the sequence obtained from an elliptic sequence by taking every kth term from *any* term h_t satisfies the following recursion:

Theorem 4.6.2. *Let (h_n) be an elliptic sequence and fix two integers k and t such that h_k is a non-zero term of (h_n). If*

$$\ell_s = h_{t+sk} \quad \text{for all } s \in \mathbb{Z},$$

then (ℓ_s) satisfies the recursion

$$\ell_{s+2}\,\ell_{s-2} = \lambda_1\,\ell_{s+1}\,\ell_{s-1} + \lambda_2\,\ell_s^2 \quad \text{for all } s \in \mathbb{Z}, \tag{4.7}$$

where

$$\lambda_1 = \left(\frac{h_{2k}}{h_k}\right)^2 \qquad \text{and} \qquad \lambda_2 = -\frac{h_{3k}}{h_k}.$$

Proof: By Theorem 4.1.2 with $m = t + sk$, $n = 2k$ and $t = k$, we have for any $s \in \mathbb{Z}$,

$$h_{t+(s+2)k}\,h_{t+(s-2)k}\,h_k^2 = h_{t+(s+1)k}\,h_{t+(s-1)k}\,h_{2k}^2 - h_{3k}\,h_k\,h_{t+sk}^2.$$

So (dividing by $h_k^2 \neq 0$)

$$h_{t+(s+2)k}\,h_{t+(s-2)k} = h_{t+(s+1)k}\,h_{t+(s-1)k}\left(\frac{h_{2k}}{h_k}\right)^2 + \left(-\frac{h_{3k}}{h_k}\right)h_{t+sk}^2.$$

In other words,

$$\ell_{s+2}\,\ell_{s-2} = \lambda_1\,\ell_{s+1}\,\ell_{s-1} + \lambda_2\,\ell_s^2.$$

□

A rational sequence satisfying the recursion (4.7) is called a *Somos 4 sequence*; we will return to Somos 4 sequences in chapter 6. Note that the coefficients λ_1, λ_2 do not depend on t, only on k.

If (h_n) is an elliptic sequence in which $h_1, \ldots, h_{12}$ are non-zero, then it turns out that all terms are non-zero except h_0:

Theorem 4.6.3. *Let (h_n) be an elliptic sequence, in which $h_N = 0$ for some minimal positive index N. Then $N \in \{1, 2, \ldots, 10\} \cup \{12\}$.*

Proof: If $N \geq 3$, let E be an associated elliptic curve from Theorem 4.5.3. Then the point $(0, 0)$ has finite order N in $E(\mathbb{Q})$, so if E is non-singular the result follows from Mazur's Theorem 3.7.1. If E is singular then it follows from Theorem 3.7.2. □

If (h_n) is an EDS in which h_3 and h_4 are coprime then the result can be improved:

Theorem 4.6.4. *[30]*

Let (h_n) be an EDS with $\gcd(h_3, h_4) = 1$. If $h_N = 0$ for some minimal positive index N, then $N \leq 5$ and (h_n) is periodic with period either N or $2N$.

Remarks:

1. If the EDS (h_n) is periodic with period $2N$ then $h_{t+N} = -h_t$ for all $t \in \mathbb{Z}$.

2. Since $h_5 = h_4\,h_2^3 - h_3$, it follows that $h_5 = 0$ if and only if $h_3 = h_4\,h_2^3$. Since h_3 is coprime to h_4, and hence to h_2, it follows that $h_5 = 0$ if and only if either $h_2 = h_3 = h_4 = 1$ or two of h_2, h_3, h_4 are -1 and the other is 1.

3. If h_3 and h_4 are not coprime then (h_n) is not periodic (since if $p \mid \gcd(h_3, h_4)$ then it will follow from Theorem 4.7.1 that p divides all terms from h_3 onwards, but $p \nmid h_1$). Since clearly any periodic EDS must have $h_N = h_0 = 0$ for some $N > 0$, it follows that Theorem 4.6.4 covers all periodic EDSs.

4. So periodic EDSs can have only the periods $1, 2, 3, 4, 5, 6, 8$ or 10, and Ward showed that each of these periods does actually occur.

5. Hence the surprising result that if we define a sequence (h_n) recursively by the stepping formula (4.4) and the initial values satisfy $h_0 = 0$, $h_1 = 1$, $h_2, h_3, h_4 \neq 0$, $\gcd(h_3, h_4) = 1$ and $h_4 h_2^3 - h_3 \neq 0$, then the recursion never produces a zero term after h_0.

4.7 EDSs reduced modulo prime powers

When an EDS (h_n) is considered modulo a prime power p^r it shows some interesting structure. In this section we give the known results on the symmetry and periodicity of the reduced sequence $(h_n \bmod p^r)$, and describe how we have extended them in this thesis. Proofs of the new results are left to the next chapter.

4.7.1 The pattern of zeroes

Let (h_n) be an EDS, let p be a prime and let $r \in \mathbb{N}$. It is obvious from the divisibility property of EDSs that if p^r divides h_M for some minimal positive index M then p^r divides h_{sM} for all $s \in \mathbb{Z}$. The next result says firstly that p^r *does* divide some such term h_M, and secondly that (unless p divides both h_3 and h_4) the h_{sM} are the only terms divisible by p^r. That is, the zeroes modulo p^r are regularly spaced in (h_n).

Definition: Let (h_n) be an elliptic sequence, and let p^r be a prime power such that h_n is a p-integer for every $n \in \mathbb{Z}$. Then p^r is said to be *regular* in (h_n) if there exists a positive index M such that

$$h_n \equiv 0 \bmod p^r \;\Leftrightarrow\; n \equiv 0 \bmod M,$$

and *irregular* otherwise. M is called the *gap* of p^r in (h_n).

Our definition of the gap matches Ward's definition of the *rank of apparition* for regular prime powers (see [30] for further details). For a fixed prime p we will usually denote the gap of p^r by N_r for all $r \in \mathbb{N}$.

Theorem 4.7.1. *[31]*

Let (h_n) be an EDS and p a prime. If p divides both h_3 and h_4 then p divides all terms h_n with $|n| \geq 3$. Otherwise p^r is regular in (h_n) for every $r \in \mathbb{N}$.

Equivalently (if h_3 and h_4 are coprime), we have

Theorem 4.7.2. *[31]*

If (h_n) is an EDS in which the initial values h_3 and h_4 are coprime, then

$$\gcd(h_n, h_m) = h_{\gcd(n,m)}$$

for all $n, m \in \mathbb{Z}$.

Remark: Ward proves Theorem 4.7.1 for $r = 1$ in [30] by simple use of the elliptic sequence equation (4.1), and then proves it for all $r \in \mathbb{N}$ by using Theorem 4.6.1 with $k = N_{r-1}$. Of course the $p \nmid \gcd(h_3, h_4)$ case also follows directly from Theorem 4.5.9: if $E/\mathbb{Q}$ is an elliptic curve associated with the EDS (h_n) then the gap N_r of p^r in (h_n) is equal to the order of P mod p^r in $E(\mathbb{Z}_{p^r})$.

Ward [30] proved that if (h_n) is any EDS and p any prime, then p divides one of the first $2p + 1$ terms of (h_n) after h_0. Shipsey [23] pointed out that for $p > 3$ this bound can be improved using the Hasse-Weil Theorem (Theorem 3.6.1) in an associated elliptic curve reduced modulo p. (For $p = 2$ or 3, $N_1 \leq 2\,p + 1$ implies $N_1 \leq p + 1 + 2\sqrt{p}$, so the Hasse bound is not an improvement.)

Theorem 4.7.3. *[23]*

Let (h_n) be an EDS and p a regular prime with gap N_1 in (h_n). Then

$$N_1 \leq p + 1 + 2\sqrt{p}.$$

If an EDS (h_n) is singular modulo a prime p, then we have a much stronger result. Since in this case the associated elliptic curves are singular when reduced modulo p, it follows from Theorem 3.6.3 that there are only a few possible values for the gap of p in (h_n):

Theorem 4.7.4. *[30]*

Let (h_n) be an EDS and let p be a prime not dividing $h_2\, h_3$ but dividing the discriminant $\Delta(h_2, h_3, h_4)$ of (h_n). Then p is regular with gap N_1, where either $N_1 = p$ or N_1 divides $p+1$ or $p-1$.

Unlike in the singular case, for a general elliptic sequence (h_n) there is no simple way to predict the gap of a given prime p in (h_n) (this is of course related to the problem of point-counting in the associated elliptic curves). In fact, it follows from Theorem 3.6.2 that for every prime p and every integer t satisfying $|t| \leq 2\sqrt{p}$ there exists at least one elliptic sequence (h_n) in which p has gap N_1 dividing $p+1+t$. However, it turns out that if $p \nmid h_2\, h_3$ then for $r \in \mathbb{N}$ the gap N_r of p^r is predictable once the gap N_1 of p is known:

Theorem 4.7.5. *Let (h_n) be an EDS in which h_0 is the only zero term, and let p be a regular prime with gap $N_1 \geq 4$ in (h_n). Let p^w be the highest power of p dividing h_{N_1}.*

If p is odd, or $p = 2$ and $w \geq 2$, then for any $r \in \mathbb{N}$, p^r has gap

$$N_r = \begin{cases} N_1 & \text{if } r \leq w \\ pN_{r-1} & \text{if } r > w. \end{cases}$$

If $p = 2$ and $w = 1$ then for some $v \geq 2$,

$$N_r = \begin{cases} N_1 & \text{if } r = 1, \\ 2N_1 & \text{if } 2 \leq r \leq v, \\ 2N_{r-1} & \text{if } r > v. \end{cases}$$

So if p is odd and $r \geq k \geq w$, then $N_r = p^{r-k}\, N_k$. Note that $p^{w+1} \mid h_{pN_1}$.

Ward proved Theorem 4.7.5 for primes $p > 3$ in [31] using an elliptic curve argument, but in section 5.3 we give an elementary proof which also holds for

$p = 2$ and $p = 3$. (If $p = 2$ or 3 the coefficients c_4 and c_6 in Ward's associated elliptic curve (in Theorem 4.5.1) might not be p-integers, which is why his proof doesn't work for these primes.)

The weaker result for $p = 2$ is necessary, as shown by the following example:

Example: Let c be any integer, and let (h_n) be the EDS with initial values

$$h_0 = 0,\ h_1 = 1,\ h_2 = 1,\ h_3 = 1,\ h_4 = 2c.$$

So 2 is a regular prime with gap $N_1 = 4$ in (h_n), and $2^{w-1} \parallel c$. Let 2^r have gap N_r for all $r \in \mathbb{N}$.

Using the elliptic sequence formula, we find

$$h_8 = -4c\left(4c^2 - 3c + 1\right).$$

Note that $4c^2 - 3c + 1$ has opposite parity to c. So if $w \geq 2$ (i.e., if c is even) then $2^{w+1} \parallel h_8$, so $N_{w+1} = 2N_1$ and $v = w$, as expected.

Otherwise, if $w = 1$ (i.e., c is odd) then $4c^2 - 3c + 1$ is even, and $h_8 = h_{2N_1}$ is divisible by $8 = 2^3$. So $v \geq 3$. Now choose any $v \geq 3$, and let c be a solution to the quadratic congruence $4c^2 - 3c + 1 \equiv 0 \bmod 2^{v-2}$. Then 2^v divides h_{2N_1}. This means that if $2 \parallel h_{N_1}$ then h_{N_2} can be divisible by an arbitrarily high power of 2.

The restriction $N_1 \geq 4$ in Theorem 4.7.5 is also necessary, as shown by the following example from [31]:

Example: Let c be any square-free integer greater than 1, and let (h_n) be the EDS with initial values

$$h_2 = 1,\ h_3 = c,\ h_4 = c + 1.$$

By the elliptic sequence formula, $h_5 = 1 + c - c^3$ and $h_6 = -c^2(c^2 + c - 1)$. So if p is any odd prime factor of c then $N_1 = 3$ but $N_2 = 6$, which is less than pN_1.

Remarks:

1. As mentioned in section 3.9.3, Theorem 4.7.5 leads immediately to a result (Theorem 3.9.5) about the order of a point in an elliptic curve over $\mathbb{Z}_{p^r}$.

2. Let (h_n) be an EDS, let p be a regular prime with gap $N_1 \geq 4$ in (h_n) and let $r \in \mathbb{N}$. Let (ℓ_n) be the sequence defined by

$$\ell_n = \frac{h_{nN_r}}{h_{N_r}} \quad \text{for all } n \in \mathbb{Z}.$$

   By Theorem 4.6.1 (ℓ_n) is an EDS. Theorem 4.7.5 says that p has gap p in (ℓ_n), and if p is odd, or $p = 2$ and $r \geq 2$, then $p^2 \nmid \ell_p$. (In fact, this is how Ward proves the result for $p > 3$ in [31], using Theorem 4.7.7.)

3. Note that w can be arbitrarily large (for instance, since we can choose h_2, h_3 and h_4, we can have $p^w \mid h_4$ for any w), but is usually 1. (Ward states this in [31] and it was true of the EDSs we generated.)

   If $h_{N_1} \bmod p^2$ were random for random initial values h_2, h_3, h_4, then a heuristic argument would suggest that we would have $w > 1$ (i.e., $p^2 \mid h_{N_1}$) for about $\frac{1}{p}$ of all EDSs (h_n).

4. Proving results about w for general p in a specific EDS (h_n) seems to be difficult. For example, proving that all primes have $w = 1$ in the EDS $0, 1, 1, -2, -3, 5, 8, \ldots$ would lead to a proof that there are no powers in the Fibonacci sequence (F_n) after F_{12}.

4.7.2 Symmetry

Ward discovered the following pattern in the reduced sequence $(h_n \bmod p)$ (his "symmetry formula"):

Theorem 4.7.6. *[30]*

Let (h_n) be an elliptic divisibility sequence and let p be a regular prime with gap $N_1 \geq 4$ in (h_n). Then there exist integers b_1 and c_1 such that

$$h_{t+sN_1} \equiv c_1^{st} \, (-b_1)^{s^2} \, h_t \bmod p \quad \textit{for all } s, t \in \mathbb{Z}. \tag{4.8}$$

Specifically,

$$b_1 = \frac{(h_{N_1-1})^2\, h_2}{h_{N_1-2}} \bmod p \qquad \textit{and} \qquad c_1 = \frac{h_{N_1-1}\, h_2}{h_{N_1-2}} \bmod p.$$

Furthermore, $b_1^2 \equiv c_1^{N_1} \equiv -h_{N_1-1}\, h_{N_1+1} \bmod p$.

In chapter 5 we consider the symmetry of the sequence $(h_n \bmod p^r)$ for $r \in \mathbb{N}$. We prove that the same symmetry formula holds in $(h_n \bmod p^r)$ if N_1 is replaced by N_r and b_1 and c_1 by new constants b_r and c_r. Ward's proof of Theorem 4.7.6 relies on the fact that for every EDS and regular prime p with gap $N_1 \geq 4$ there is an associated elliptic curve over $\mathbb{F}_p$ in which the point $P = (0, 0)$ has order N_1, and he mentioned in [30] that he was interested in finding an elementary proof (meaning one that does not rely on this connection). Our generalisation is proved by elementary means, using only the elliptic sequence formula (4.1).

Remark: Ward actually stated Theorem 4.7.6 for $s, t \geq 0$ only, but it is easy to modify his proof to work for all $s, t \in \mathbb{Z}$. Also, his proof only works for primes $p > 3$, but he proves the result for $p = 3$ separately by checking it for each of the 21 possibilities for $(h_n \bmod 3)$, and for $p = 2$ it holds trivially.

Another way to look at Theorem 4.7.6 is to fix an index t such that $h_t \not\equiv 0 \bmod p$, and consider the sequence (ℓ_n) obtained from (h_n) by taking every N_1th term from h_t and dividing by h_t, i.e.,

$$\ell_s = \frac{h_{t+sN_1}}{h_t} \text{ for all } s \in \mathbb{Z}.$$

It follows easily from Theorem 4.6.2 that (ℓ_n) is a Somos 4 sequence; Theorem 4.7.6 says that, modulo p, this sequence has the simple form

$$\ell_s \equiv \left(c_1^t\right)^s\, (-b_1)^{s^2} \bmod p \quad \text{ for all } s \in \mathbb{Z},$$

for some constants b_1 and c_1 which can be calculated from $(h_n \bmod p)$. In fact, since $(c_1^t)^s\, (-b_1)^{s^2}$ can be rewritten as $(-c_1^t\, b_1)^{\frac{1}{2}s(s+1)}\, (-c_1^{-t}\, b_1)^{\frac{1}{2}s(s-1)}$ and

$$\ell_1 \equiv -c_1^t\, b_1 \bmod p \quad \text{ and } \quad \ell_{-1} \equiv -c_1^{-t}\, b_1 \bmod p, \tag{4.9}$$

we can write this even more simply:

$$\ell_s \equiv {\ell_1}^{\frac{1}{2}s(s+1)}\, {\ell_{-1}}^{\frac{1}{2}s(s-1)} \bmod p \quad \text{for all } s \in \mathbb{Z}.$$

So in a sense, what we mean by finding symmetries in $(h_n \bmod p)$ is finding a simple expression for $\frac{h_{t+sN_1}}{h_t} \bmod p$ that holds for all $s \in \mathbb{Z}$.

In chapter 5 we show that

$$\frac{h_{t+sN_r}}{h_t} \equiv \left(\frac{h_{t+N_r}}{h_t}\right)^{\frac{1}{2}s(s+1)} \left(\frac{h_{t-N_r}}{h_t}\right)^{\frac{1}{2}s(s-1)} \bmod p^{3r},$$

if $t \not\equiv 0 \bmod N_1$, not just modulo p^r, generalising (4.9). We use this to calculate c_r and b_r in terms of c_w and b_w. So for all $r \in \mathbb{N}$, $\frac{h_{t+sN_r}}{h_t} \bmod p^r$ can be calculated from the sequence $(h_n \bmod p^w)$.

Finally, we consider the sequence obtained from (h_n) by taking every N_rth term from h_0 and dividing by h_{N_r} (assuming $h_{N_r} \neq 0$). By Theorem 4.6.1 this sequence is an EDS. Using the theory of division polynomials, Ward found the following formula for $\frac{h_{sN_r}}{h_{N_r}} \bmod p^{2r}$ in the case where $p \neq 2$ or 3:

Theorem 4.7.7. *Let (h_n) be an EDS in which h_0 is the only zero term, and let $p > 3$ be a regular prime with gap $N_1 \geq 4$ in (h_n). For $r \in \mathbb{N}$ let p^r have gap N_r in (h_n). Then $-h_{N_r+1}\, h_{N_r-1}$ is a quadratic residue modulo p^{2r}, and*

$$\frac{h_{sN_r}}{h_{N_r}} \equiv s\, {\xi_r}^{s^2-1} \bmod p^{2r} \quad \textit{for all } s \in \mathbb{Z},$$

where $\xi_r^2 \equiv -h_{N_r+1}\, h_{N_r-1} \bmod p^{2r}$.

So the EDS (ℓ_s) defined by $\ell_s = \frac{h_{sN_r}}{h_{N_r}}$ for all $s \in \mathbb{Z}$ has an equivalent sequence which is congruent to the sequence of integers $\mathbb{Z}$ modulo p^{2r-1}. In chapter 5 (Theorem 5.2.5) we give an elementary proof that works for all primes p, although for $p = 2$ we only get a modulo 2^{2r-1} formula.

Rachel Shipsey gave independently an elementary proof of the following formula for $\frac{h_{sN_1}}{h_{N_1}}$ modulo p:

Theorem 4.7.8. *[23]*

Let (h_n) be an elliptic divisibility sequence and let p be a prime with gap $N_1 \geq 4$ in (h_n). If $h_{N_1} \neq 0$, then

$$\frac{h_{sN_1}}{h_{N_1}} \equiv s\,(-b_1)^{s^2-1} \bmod p \quad \text{for all } s \in \mathbb{Z}. \tag{4.10}$$

(Shipsey states this for $\gcd(h_3, h_4) = 1$ and $p > 3$ only, but this is an unnecessary assumption.) We prove in chapter 5 (Theorem 5.2.4) that Shipsey's formula (4.10) holds modulo p^r for any $r \in \mathbb{N}$ if N_1 is replaced by N_r and b_1 and c_1 by b_r and c_r.

Notice that Ward's Theorem 4.7.7 gives two possibilities for $\frac{h_{sN_r}}{h_{N_r}} \bmod p^{2r}$ if s is even, since $(-h_{N_r+1}\,h_{N_r-1})^{s^2-1}$ has two square roots in $\mathbb{Z}^*_{p^{2r}}$, reducing to $b_r^{s^2-1}$ and $-b_r^{s^2-1}$ respectively modulo p^r. Shipsey's Theorem 4.7.8 and its generalisation Theorem 5.2.4 specify which of these possibilities is the right one.

4.7.3 Periodicity

Ward used his modulo p symmetry formula (4.8) to prove that for any regular prime $p \nmid h_2\,h_3$ the reduced sequence $(h_n \bmod p)$ is periodic. The period is τ_1 times the gap N_1 for some integer τ_1 dividing $p-1$, which can be explicitly calculated from $(h_n \bmod p)$ once we know N_1.

Theorem 4.7.9. *[30]*

Let (h_n) be an elliptic divisibility sequence and p a regular prime with gap $N_1 \geq 4$ in (h_n). If τ_1 is the least positive integer such that

$$c_1^{\tau_1} \equiv 1 \bmod p \quad \text{and} \quad (-b_1)^{\tau_1^2} \equiv 1 \bmod p,$$

then the sequence $(h_n \bmod p)$ is periodic with period $\tau_1\,N_1$. Furthermore, τ_1 divides $p-1$.

Ward gave an explicit formula for τ_1 if p is an odd prime:

Theorem 4.7.10. *[30]*

Let (h_n) be an elliptic divisibility sequence and p an odd prime with gap $N_1 \geq 4$ in (h_n). Let ϵ and κ be the orders in $\mathbb{Z}^*_p$ of

$$h_2\,(h_{N_1-2})^{-1} \equiv c^2\,b^{-1} \bmod p \quad \text{and} \quad h_{N_1-1} \equiv b\,c^{-1} \bmod p$$

respectively, and let the constant α *be given by*

$$\alpha = \begin{cases} 1 & \text{if } \epsilon \text{ and } \kappa \text{ are both odd,} \\ 0 & \text{if } \epsilon \text{ and } \kappa \text{ are divisible by different powers of 2, or} \\ -1 & \text{otherwise.} \end{cases}$$

Then

$$\tau_1 = 2^\alpha \operatorname{lcm}(\epsilon, \kappa).$$

(The restriction to odd primes is necessary: if $p = 2$ then $\tau_1 \mid (p-1)$ implies $\tau_1 = 1$, but Theorem 4.7.10 would give $\tau_1 = 2$.)

In section 5.6 we look at the periodicity of the sequence $(h_n \bmod p^r)$ for $r \in \mathbb{N}$. We define a quantity τ_r, generalising τ, and use our extension of Ward's symmetry formula (4.8) to the modulo p^r case to prove that for any regular prime p with gap $N_1 \geq 4$ the sequence $(h_n \bmod p^r)$ is periodic with period $\pi_r = \tau_r N_r$.

Shipsey [23] made the following conjecture about the period of $(h_n \bmod p^2)$:

Conjecture 4.7.11. *Let* (h_n) *be an EDS and* p *an odd prime such that* $p \nmid h_2 h_3$ *and* $p^2 \nmid h_{N_1}$. *Then the period of* $(h_n \bmod p^2)$ *is* p *times the period of* $(h_n \bmod p)$.

In chapter 5 we show the following:

Theorem 4.7.12. *Let* (h_n) *be an elliptic divisibility sequence, let* p *be a regular odd prime with gap* $N_1 \geq 4$ *in* (h_n), *and let* p^w *be the highest power of* p *dividing* N_1. *For* $r \in \mathbb{N}$ *let* π_r *be the period of* $(h_n \bmod p^r)$.

Then there exists an integer $u \leq w$ *such that*

$$\pi_r = \begin{cases} \pi_1 & \text{for } r \leq u, \text{ and} \\ p^{r-u}\,\pi_1 & \text{for } r \geq u. \end{cases}$$

So as r increases from 1 the period of $(h_n \bmod p^r)$ remains the same until r reaches some value u, and then increases by a factor of p each time. This confirms Conjecture 4.7.11 for the $r = 2$, $w = 1$ case. We also find a simple formula for τ_r in terms of the orders of b_r and c_r in $\mathbb{Z}_{p^r}^*$, and prove it is equivalent to Theorem 4.7.10 for $r = 1$.

4.7.4 Regular primes with $N_1 = 2$ or 3

If p is a regular prime with gap $N_1 = 2$ or 3 then the sequence $(h_n \bmod p)$ is described completely by the following result of Ward's:

Theorem 4.7.13. *[30]*

Let (h_n) be an EDS and p a regular prime. If p has gap $N_1 = 2$ then

$$h_n \equiv \begin{cases} 0 \bmod p & \text{if } n = 2k, \\ (-1)^{\frac{1}{2}k(k-1)}\, h_3^{\frac{1}{2}k(k+1)} \bmod p & \text{if } n = 2k+1. \end{cases}$$

If p has gap $N_1 = 3$ then

$$h_n \equiv \begin{cases} 0 \bmod p & \text{if } n = 3k, \\ (-h_2)^{\frac{1}{2}k(k-1)}\, {h_4}^{\frac{1}{2}k(k+1)} \bmod p & \text{if } n = 3k+1, \text{or} \\ -(-h_2)^{\frac{1}{2}(k+1)(k+2)}\, {h_4}^{\frac{1}{2}k(k+1)} \bmod p & \text{if } n = 3k+2. \end{cases}$$

It follows that if p divides $h_2\, h_3$ but not $\gcd(h_3, h_4)$ then $(h_n \bmod p)$ is still periodic, and its period depends on the orders of its initial values in $\mathbb{Z}_p^*$. We have not proved any results about $(h_n \bmod p^r)$ where p has gap 2 or 3 and $r \geq 2$.

4.7.5 Irregular primes

Finally, in section 5.7 we consider irregular primes. If h_n is an EDS and p is a prime which divides both h_3 and h_4, then by Theorem 4.7.1 p divides every term h_n for $|n| \geq 3$. We prove that if $p \nmid h_2$ then as n increases the power of p dividing h_n grows at least as fast as $\frac{1}{25}\, n^2$.

Chapter 5

Symmetry and periodicity of elliptic divisibility sequences modulo prime powers

In this chapter we give some new symmetry results for elliptic divisibility sequences reduced modulo a prime power p^r, extending Ward's Theorems 4.7.6 and 4.7.7 and Shipsey's Theorem 4.7.8. We then use them to find the period of such sequences, confirming a conjecture by Shipsey [23] for the $r = 2$ case.

Throughout this chapter (h_n) is an EDS and p is a regular prime with gap $N_1 \geq 4$ in (h_n). The highest power of p dividing h_{N_1} is p^w, and for each $r \in \mathbb{N}$ the gap of p^r in (h_n) is N_r.

Since by Theorem 4.4.9 every elliptic sequence (h_n) in which h_2 and h_3 are not both zero is equivalent to an EDS, these results extend easily to elliptic sequences.

5.1 Symmetries in EDSs modulo p^r

In this section we give an elementary proof of Ward's symmetry formula for $(h_n \bmod p)$ (Theorem 4.7.6), at the same time generalising it to the modulo p^r case.

Definition: The constants b_r and c_r are defined by

$$b_r = -\left(\frac{h_{N_r-1}}{h_{-1}}\right)^2 \frac{h_{-2}}{h_{N_r-2}} \bmod p^r, \qquad \text{and} \qquad c_r = \frac{h_{N_r-1}}{h_{-1}} \frac{h_{-2}}{h_{N_r-2}} \bmod p^r.$$

Note that $p \nmid h_2$ implies $p \nmid h_{N_r-2}$. So the constants b_r and c_r are defined, and coprime to p.

Lemma 5.1.1. *For all integers t,*

$$h_{t+N_r} \equiv c_r^t \, (-b_r) \, h_t \bmod p^r. \tag{5.1}$$

Proof: We first show that equation (5.1) holds for four initial values of $t \in \{-3, -2, -1, 0\}$; the rest will follow easily by induction.

For $t = 0$ equation (5.1) holds trivially, and for $t = -2, -1$ it follows from the definition of b_r and c_r. To prove it for $t = 3$ we note that $h_{N_r-3} \bmod p^r$ can be written in terms of the next two terms, using (4.4) with $m = N_r - 2$:

$$h_{N_r-3} \, h_{N_r-1} \, h_2^2 - h_1 \, h_3 \, h_{N_r-2}^2 = h_{N_r-4} \, h_{N_r} \equiv 0 \bmod p^r.$$

Substituting for h_{N_r-1} and h_{N_r-2} modulo p^r in terms of c_r and b_r and simplifying shows that the formula holds for $t = -3$ too.

Now assume (5.1) holds for $t = -3, -2, -1, 0, 1, \ldots, \tau - 1$ for some $\tau \geq 1$. We prove it holds for $t = \tau$.

By (4.4) with $m = N_r + \tau - 2$ we have

$$h_{N_r+\tau} \, h_{N_r+\tau-4} = h_{N_r+\tau-1} \, h_{N_r+\tau-3} \, h_2^2 - h_1 \, h_3 \, h_{N_r+\tau-2}^2.$$

Since (5.1) holds for the four values of $t \in \{\tau - 4, \ldots, \tau - 1\}$ by the inductive hypothesis, this gives

$$\begin{aligned} h_{N_r+\tau} \left(c_r^{\tau-4} \, (-b_r) \, h_{\tau-4}\right) &\equiv \left(c_r^{\tau-1} \, (-b_r) \, h_{\tau-1}\right) \left(c_r^{\tau-3} \, (-b_r) \, h_{\tau-3}\right) \, h_2^2 \\ &\qquad - h_1 \, h_3 \left(c_r^{\tau-2} \, (-b_r) \, h_{\tau-2}\right)^2 \bmod p^r \\ &\equiv c_r^{2\tau-4} \, (-b_r)^2 \, \left(h_{\tau-1} \, h_{\tau-3} \, h_2^2 - h_1 \, h_3 \, h_{\tau-2}^2\right) \bmod p^r \\ &\equiv c_r^{2\tau-4} \, (-b_r)^2 \, \left(h_\tau \, h_{\tau-4}\right) \bmod p^r. \end{aligned}$$

If $h_{\tau-4} \not\equiv 0 \bmod p$, then dividing both sides by $c_r^{\tau-4}\,(-b_r)\,h_{\tau-4}$, we get

$$h_{N_r+\tau} \equiv c_r^{\tau}\,(-b_r)\,h_\tau \bmod p^r,$$

as required.

If $\tau \equiv 4 \bmod N_1$, then (5.1) holds for the six values of $t \in \{\tau-6, \ldots, \tau_r-1\}$ by the inductive hypothesis. Setting $m = N_r + \tau - 3$ and $n = 3$ in the elliptic sequence formula (4.1) we have

$$h_{N_r+\tau}\,h_{N_r+\tau-6} = h_{N_r+\tau-1}\,h_{N_r+\tau-5}\,h_3^2 - h_2\,h_4\,h_{N_r+\tau-3}^2,$$

and a similar argument shows that (5.1) holds for $t = \tau$. It follows by induction that (5.1) holds for $t \geq 1$. The $t < -3$ case can now be proved in a similar way. □

Lemma 5.1.2. *The constants c_r and b_r are related by*

$$c_r{}^{N_r} \equiv b_r^2 \bmod p^r.$$

Proof: Setting first $t = 1 - N_r$ and then $t = -1$ in (5.1) we get

$$\begin{aligned} h_1 = h_{(1-N_r)+N_r} &\equiv c_r{}^{1-N_r}\,b_r\,h_{-1+N_r} \bmod p^r \\ &\equiv c_r{}^{1-N_r}\,b_r\,\left(c_r{}^{-1}\,b_r\,h_1\right) \equiv c_r{}^{-N_r}\,b_r^2\,h_1 \bmod p^r. \end{aligned}$$

It follows (since $h_1 = 1 \not\equiv 0 \bmod p$) that $c_r{}^{N_r} \equiv b_r^2 \bmod p^r$. □

The rest of the theorem is now proved by writing $h_{t+sN_r} = h_{(N_r+t)+(s-1)N_r}$ and using an easy induction on s together with Lemma 5.1.2:

Theorem 5.1.3. *Let (h_n) be an elliptic divisibility sequence and let p be a regular prime with gap $N_1 \geq 4$ in (h_n). Let $r \in \mathbb{N}$, and let the gap of p^r in (h_n) be N_r. Then for all $s, t \in \mathbb{Z}$,*

$$h_{t+sN_r} \equiv c_r{}^{st}\,(-b_r)^{s^2}\,h_t \bmod p^r.^1 \tag{5.2}$$

[1] Alex Dent has asked why we don't just define b_r as $-b_r$ to get rid of the minus sign. The answer is that this is how Ward defined b_1 in [30].

Proof: The result holds trivially for $s = 0$, and we proved in Lemma 5.1.1 that it holds for $s = 1$. Assume it holds for $s = 0, 1, \ldots, \sigma - 1$ for some integer $\sigma \geq 2$. Then for all $t \in \mathbb{Z}$,

$$\begin{aligned} h_{t+\sigma N_r} &= h_{(N_r+t)+(\sigma-1)N_r} \\ &\equiv {c_r}^{(\sigma-1)(N_r+t)} \, (-b_r)^{(\sigma-1)^2} \, h_{N_r+t} \bmod p^r \\ &\equiv \left({c_r}^{N_r}\right)^{\sigma-1} \, {c_r}^{(\sigma-1)t} \, (-b_r)^{\sigma^2-2\sigma+1} \, \left({c_r}^{t} \, (-b_r) \, h_t\right) \bmod p^r \\ &\equiv \left(b_r^2\right)^{\sigma-1} \, {c_r}^{\sigma t} \, (-b_r)^{\sigma^2-2\sigma+2} \, h_t \bmod p^r \\ &\equiv {c_r}^{\sigma t} \, (-b_r)^{\sigma^2} \, h_t \bmod p^r, \end{aligned}$$

where we have used the induction hypothesis and the fact that $c^{N_r} \equiv b_r^2 \bmod p^r$ by Lemma 5.1.2. It follows by induction that (5.2) holds for all $s \geq 0$. For $s < 0$, we have $h_{t+sN_r} = -h_{-t+(-s)N_r} \equiv -{c_r}^{(-s)(-t)} \, (-b_r)^{(-s)^2} \, h_{-t} \equiv {c_r}^{st} \, (-b_r)^{s^2} \, h_t \bmod p^r$, so (5.2) also holds for $s < 0$. □

We will use the following corollary later:

Corollary 5.1.4. *If $N_1 \geq 4$, then for all integers t,*

$$h_{t+N_r} \, h_{t-N_r} \equiv b_r^2 \, h_t^2 \bmod p^r.$$

The following example of Ward's (see [31]) shows that the restriction $N_1 \geq 4$ in Theorem 5.1.3 is necessary:

Example: Let $\alpha > 1$ be a squarefree integer, let (h_n) be the EDS with initial values $h_0 = 0$, $h_1 = 1$, $h_2 = 1$, $h_3 = \alpha$, $h_4 = 1+\alpha$, and let p be any prime factor of α. Then $h_5 = 1 + \alpha - \alpha^3$, $h_6 = -\alpha^2 \, (\alpha^2 + \alpha - 1)$, and the sequence $(h_n \bmod p^2)$ is

$$\ldots, 0, \ 1, \ 1, \ \alpha, \ 1+\alpha, \ 1+\alpha, \ 0, \ldots$$

So we have $N_1 = 3$, $N_2 = 6$ and

$$c_2 = \frac{h_5 \, h_2}{h_4} \equiv \frac{1+\alpha}{1+\alpha} \equiv 1 \bmod p^2, \quad b_2 = \frac{h_5^2 \, h_2}{h_4} \equiv \frac{(1+\alpha)^2}{1+\alpha} \equiv 1 + \alpha \bmod p^2.$$

If the symmetry formula held we would have $\frac{h_{N_2-t}}{h_t} \equiv c_2^{-t} \, b_2 \equiv 1 + \alpha \bmod p^2$ for all $t \not\equiv 0 \bmod N_1$. But instead we find for $t = 1, 2, 3, 4$ that $\frac{h_{N_2-t}}{h_t} \bmod p^2$ takes

the values

$$1+\alpha,\ 1+\alpha,\ 1,\ \frac{1}{1+\alpha}.$$

Furthermore, $c_2^{N_2} \equiv 1^6 \not\equiv (1+\alpha)^2 \equiv b_2^2 \bmod p^2$.

The symmetry formula can be written in another way as follows. (We will refer to this form in chapter 8.)

Corollary 5.1.5. *Let (h_n) be an elliptic divisibility sequence and let p be a regular prime with gap $N_1 \geq 4$ in (h_n). Let $r \in \mathbb{N}$, and let the gap of p^r in (h_n) be N_r. Then for all $s, t \in \mathbb{Z}$,*

$$h_{t+sN_r} \equiv {c_r}^{st} \left({c_r}^{N_r}\right)^{\frac{1}{2}s(s-1)} (-b_r)^s\, h_t \bmod p^r.$$

Proof: Rearranging the symmetry formula (5.2) and using Lemma 5.1.2, we have

$$\begin{aligned} h_{t+sN_r} &\equiv {c_r}^{st}\,(-b_r)^{s^2}\, h_t \bmod p^r \\ &\equiv {c_r}^{st}\left(b_r^2\right)^{\frac{1}{2}s(s-1)} (-b_r)^s\, h_t \bmod p^r \\ &\equiv {c_r}^{st}\left({c_r}^{N_r}\right)^{\frac{1}{2}s(s-1)} (-b_r)^s\, h_t \bmod p^r. \end{aligned}$$

The result follows. □

Remark: We could also have proved Theorem 5.1.3 using the periodicity of the sequence $(x_n \bmod p^r)$, where for $n \in \mathbb{Z}$ we have $(x_n, y_n) = [n](0,0)$ in an associated elliptic curve $E/\mathbb{Q}$.

Another way to look at Theorem 5.1.3 is to fix an index t such that $h_t \neq 0$, and consider the sequence (ℓ_s) obtained from (h_n) by taking every N_rth term from h_t and dividing by h_t, i.e.,

$$\ell_s = \frac{h_{t+sN_r}}{h_t} \quad \text{for all } s \in \mathbb{Z}.$$

It follows easily from Theorem 4.6.2 that (ℓ_s) is a Somos 4 sequence; Theorem 5.1.3 says that, modulo p^r, this sequence has the simple form

$$\ell_s \equiv \left(c_r^t\right)^s (-b_r)^{s^2} \bmod p^r \quad \text{for all } s \in \mathbb{Z},$$

for some constants b_r and c_r which can be calculated from (h_n). In fact, since $(c_r^t)^s\ (-b_r)^{s^2}$ can be rewritten as $(-c_r^t\, b_r)^{\frac{1}{2}s(s+1)}\ (-{c_r}^{-t}\, b_r)^{\frac{1}{2}s(s-1)}$ and

$$\ell_1 \equiv -c_r^t\, b_r \bmod p^r \quad \text{and} \quad \ell_{-1} \equiv -{c_r}^{-t}\, b_r \bmod p^r,$$

we can write this even more simply:

$$\ell_s \equiv {\ell_1}^{\frac{1}{2}s(s+1)}\, {\ell_{-1}}^{\frac{1}{2}s(s-1)} \bmod p^r \quad \text{for all } s \in \mathbb{Z}.$$

In the next section we will prove that in fact

$$\ell_s \equiv {\ell_1}^{\frac{1}{2}s(s+1)}\, {\ell_{-1}}^{\frac{1}{2}s(s-1)} \bmod p^{3r} \quad \text{for all } s \in \mathbb{Z},$$

i.e., that our simple formula for the general term of (ℓ_s) actually holds modulo p^{3r}.

Remark: Notice that when thinking of the sequence (ℓ_s) we fix t and vary s, while in the proof of Theorem 5.1.3 we fix s and vary t.

5.2 Extending the symmetry formula to higher powers of p

We now find a form of the EDS symmetry formula for $\frac{h_{t+sN_r}}{h_t}$ that holds modulo p^{3r} when $p \nmid h_t$.

Consider the following example.

Example: Let (h_n) be the EDS

$$\ldots -3,\ -4,\ -1,\ -1,\ 0,\ 1,\ 1,\ 4,\ 3,\ -61,\ -280,\ -3931,\ -24603,\ 900364, \ldots.$$

The table shows the sequence (h_n) reduced modulo 5^3 and written in rows of length $N_1 = 6$. (The entry in row s and column t is $h_{t+sN_1} \bmod 5^3$.)

0	1	1	4	3	64
95	69	22	114	94	61
5	49	56	91	108	1
10	86	92	86	79	69
35	36	96	104	68	49

100	74	57	29	4	86
115	69	71	61	58	36
65	26	117	91	119	74
20	111	106	89	28	69
105	114	72	79	124	26

25	24	1	71	78	111
45	56	72	111	44	114

Let (ℓ_n) be the sequence defined by $\ell_n = \frac{h_{t+nN_1}}{h_t}$ for $n \in \mathbb{Z}$. The next table shows $\frac{h_{t+(s-1)N_1}\, h_{t+(s+1)N_1}}{h_{t+sN_1}^2} = \frac{\ell_{s-1}\, \ell_{s+1}}{\ell_s} \bmod 5^3$ for column $t \not\equiv 0 \bmod N_1$ and row s.

·	·	·	·	·	·
·	84	59	34	59	84
·	84	59	34	59	84
·	84	59	34	59	84
·	84	59	34	59	84

·	84	59	34	59	84
·	84	59	34	59	84
·	84	59	34	59	84
·	84	59	34	59	84
·	84	59	34	59	84

·	84	59	34	59	84
·	84	59	34	59	84

Notice that the entries in each column $t \not\equiv 0 \bmod N_1$ are constant, and that this constant is congruent to 9 mod 5^2 for all columns t.

This observation is the basis of the next proof, since it is fairly easy to prove by induction that if

$$\frac{\ell_{s-1}\,\ell_{s+1}}{\ell_s^2} \bmod p^{3r}$$

is constant over all $s \in \mathbb{Z}$ for given $t \not\equiv 0 \bmod N_r$ then

$$\ell_s \equiv {\ell_1}^{\frac{1}{2}s(s+1)}\,{\ell_{-1}}^{\frac{1}{2}s(s-1)} \bmod p^{3r} \quad \text{for all } s \in \mathbb{Z}.$$

Theorem 5.2.1. *Let (h_n) be an EDS, let p be a regular prime with gap $N_1 \geq 4$, and let p^r have gap N_r for all $r \in \mathbb{N}$. Then for all integers s and $t \not\equiv 0 \bmod N_1$,*

$$\frac{h_{t+sN_r}}{h_t} \equiv \left(\frac{h_{t+N_r}}{h_t}\right)^{\frac{1}{2}s(s+1)} \left(\frac{h_{t-N_r}}{h_t}\right)^{\frac{1}{2}s(s-1)} \bmod p^{3r}. \tag{5.3a}$$

In other words, if (ℓ_n) is the sequence defined by $\ell_n = \frac{h_{t+nN_r}}{h_t}$ for all $n \in \mathbb{Z}$, then

$$\ell_s \equiv {\ell_1}^{\frac{1}{2}s(s+1)}\,{\ell_{-1}}^{\frac{1}{2}s(s-1)} \bmod p^{3r} \quad \textit{for all } s \in \mathbb{Z}. \tag{5.3b}$$

Proof: Fix $t \not\equiv 0 \bmod N_1$. We first prove that $\frac{\ell_{s+1}\,\ell_{s-1}}{\ell_s^2} = \frac{h_{t+(s+1)N_r}\,h_{t+(s-1)N_r}}{h_{t+sN_r}^2} \bmod p^{3r}$ is constant over all $s \in \mathbb{Z}$. Then we use induction on s to prove the symmetry formula.

Setting $m = t + sN_r$, $n = N_r$ in the elliptic sequence formula (4.2a) and dividing by $(h_{t+sN_r})^2$, we get

$$\frac{h_{t+(s+1)N_r}\,h_{t+(s-1)N_r}}{(h_{t+sN_r})^2} = \left(\frac{h_{(t+1)+sN_r}\,h_{(t-1)+sN_r}}{(h_{t+sN_r})^2}\right) h_{N_r}^2 + h_{1-N_r}\,h_{1+N_r}.$$

But by Theorem 5.1.3, the bracketed quantity is

$$\begin{aligned}
&\frac{h_{(t+1)+sN_r}\,h_{(t-1)+sN_r}}{(h_{t+sN_r})^2} \\
&\quad\equiv \frac{\left({c_r}^{s(t+1)}\,(-b_r)^{s^2}\,h_{t+1}\right)\left({c_r}^{s(t-1)}\,(-b_r)^{s^2}\,h_{t-1}\right)}{\left({c_r}^{st}\,(-b_r)^{s^2}\,h_t\right)^2} \bmod p^r \\
&\quad\equiv \frac{h_{t+1}\,h_{t-1}}{h_t^2} \bmod p^r,
\end{aligned}$$

and so (since p^{2r} divides $h_{N_r}^2$),

$$\left(\frac{h_{(t+1)+sN_r}\, h_{(t-1)+sN_r}}{(h_{t+sN_r})^2}\right) h_{N_r}^2 \equiv \left(\frac{h_{t+1}\, h_{t-1}}{h_t^2}\right) h_{N_r}^2 \bmod p^{3r}.$$

It follows (using the elliptic sequence formula (4.1) with $m = t$ and $n = N_r$) that

$$\begin{aligned}\frac{h_{t+(s+1)N_r}\, h_{t+(s-1)N_r}}{(h_{t+sN_r})^2} &\equiv \left(\frac{h_{t+1}\, h_{t-1}}{h_t^2}\right) h_{N_r}^2 + h_{1-N_r}\, h_{1+N_r} \bmod p^{3r} \\ &\equiv \frac{h_{t-N_r}\, h_{t+N_r}}{h_t^2} \bmod p^{3r},\end{aligned}$$

which is constant over all $s \in \mathbb{Z}$. In other words,

$$\frac{\ell_{s-1}\, \ell_{s+1}}{\ell_s^2} \equiv \ell_1\, \ell_{-1} \bmod p^{3r} \quad \text{for all } s \in \mathbb{Z}.$$

We now use induction to prove that formula (5.3) holds for all $s \in \mathbb{Z}$. Note that it holds trivially for $s = 0$ and $s = 1$, so we may assume it holds for $s = 0, 1, \ldots, \sigma$ for some $\sigma \geq 1$; we prove it holds for $s = \sigma + 1$. We have

$$\ell_{\sigma+1} \equiv \frac{\ell_\sigma^2}{\ell_{\sigma-1}} \cdot \ell_1\, \ell_{-1} \bmod p^{3r}.$$

But by the inductive hypothesis

$$\ell_\sigma \equiv {\ell_1}^{\frac{1}{2}\sigma(\sigma+1)}\, {\ell_{-1}}^{\frac{1}{2}\sigma(\sigma-1)} \bmod p^{3r}$$

and

$$\ell_{\sigma-1} \equiv {\ell_1}^{\frac{1}{2}(\sigma-1)\sigma}\, {\ell_{-1}}^{\frac{1}{2}(\sigma-1)(\sigma-2)}\, h_t \bmod p^{3r}.$$

So

$$\begin{aligned}\ell_{\sigma+1} &\equiv \frac{\left({\ell_1}^{\frac{1}{2}\sigma(\sigma+1)}\, {\ell_{-1}}^{\frac{1}{2}\sigma(\sigma-1)}\right)^2}{{\ell_1}^{\frac{1}{2}(\sigma-1)\sigma}\, {\ell_{-1}}^{\frac{1}{2}(\sigma-1)(\sigma-2)}} \cdot \ell_1\, \ell_{-1} \bmod p^{3r} \\ &\equiv {\ell_1}^{\frac{1}{2}(\sigma+1)(\sigma+2)}\, {\ell_{-1}}^{\frac{1}{2}(\sigma+1)\sigma} \bmod p^{3r}.\end{aligned}$$

which is (5.3). It follows by induction that (5.3) holds for all $s \geq 0$.

For $s < 0$, we note that

$$\begin{aligned} h_{t+sN_r} &= -h_{-t+(-s)N_r} \\ &\equiv -\left(\frac{h_{-t+N_r}}{h_{-t}}\right)^{\frac{1}{2}(-s)(-s+1)} \left(\frac{h_{-t-N_r}}{h_{-t}}\right)^{\frac{1}{2}(-s)(-s-1)} h_{-t} \bmod p^{3r} \\ &\equiv \left(\frac{h_{t-N_r}}{h_t}\right)^{\frac{1}{2}s(s-1)} \left(\frac{h_{t+N_r}}{h_t}\right)^{\frac{1}{2}s(s+1)} h_t \bmod p^{3r}. \end{aligned}$$

So (5.3) also holds for $s < 0$. □

Remarks:

1. Note that, unlike Theorem 5.1.3, Theorem 5.2.1 gives no information for $s = \pm 1$.

2. The symmetry formula (5.3) can also be written

$$\ell_s \equiv {\ell_1}^{-s} \, (\ell_{-1}\,\ell_1)^{\frac{1}{2}s(s-1)} \bmod p^{3r},$$

and it is perhaps intuitively easier to see how this form of the formula arises from the fact that $\frac{\ell_{s+1}\,\ell_{s-1}}{\ell_s^2} \bmod p^{3r}$ is constant. Starting with

$$\frac{\ell_s}{\ell_{s-1}} \equiv \frac{\ell_{s-1}}{\ell_{s-2}} \cdot (\ell_1\,\ell_{-1}) \bmod p^{3r}$$

and replacing $\frac{\ell_{s-j}}{\ell_{s-j-1}}$ by $\frac{\ell_{s-j-1}}{\ell_{s-j-2}} \cdot (\ell_{-1}\,\ell_1)$ modulo p^{3r} for $j = 1, 2, \ldots, s-2$, we get

$$\frac{\ell_s}{\ell_{s-1}} \equiv \frac{\ell_1}{\ell_0} \, (\ell_{-1}\,\ell_1)^{s-1} \bmod p^{3r}$$

for all $s \in \mathbb{Z}$. Then writing ℓ_s as

$$\ell_s = \frac{\ell_s}{\ell_{s-1}} \cdot \frac{\ell_{s-1}}{\ell_{s-2}} \cdots \frac{\ell_2}{\ell_1} \cdot \frac{\ell_1}{\ell_0}$$

and noting that $\ell_0 = 1$, we get

$$\begin{aligned} \ell_s &\equiv \left(\frac{\ell_1}{\ell_0}\right)^s \cdot (\ell_{-1}\,\ell_1)^{(s-1)+(s-2)+\ldots+2+1} \bmod p^{3r} \\ &\equiv \ell_1^s \, (\ell_{-1}\,\ell_1)^{\frac{1}{2}s(s-1)} \bmod p^{3r}, \end{aligned}$$

which is the desired formula.

3. For all $s \in \mathbb{Z}$, the quantity $\frac{\ell_{s-1}\,\ell_{s+1}}{\ell_s^2} \bmod p^{2r}$ is constant for all choices of $t \not\equiv 0 \bmod N_1$, since (by the elliptic sequence equation (4.1) with $m = t + sN_r$ and $n = N_r$)

$$\begin{aligned}\frac{h_{t+(s-1)N_r}\,h_{t+(s+1)N_r}}{h_{t+sN_r}^2} &= \frac{h_{t+sN_r-1}\,h_{t+sN_r+1}\,h_{N_r}^2 - h_{N_r-1}\,h_{N_r+1}\,h_{t+sN_r}^2}{h_{t+sN_r}^2} \\ &\equiv -h_{N_r-1}\,h_{N_r+1} \bmod p^{2r}.\end{aligned}$$

4. Formula (5.3) reduces to our modulo p^r symmetry formula (5.2), since

$$\frac{h_{t+N_r}}{h_t} \equiv c_r^t\,(-b_r) \bmod p^r \quad \text{and} \quad \frac{h_{t-N_r}}{h_t} \equiv c^{-t}\,(-b_r)^{(-1)^2} \bmod p^r$$

for all $t \not\equiv 0 \bmod N_1$, so

$$\begin{aligned}\frac{h_{t+sN_r}}{h_t} &\equiv \left(\frac{h_{t+N_r}}{h_t}\right)^{\frac{1}{2}s(s+1)} \left(\frac{h_{t-N_r}}{h_t}\right)^{\frac{1}{2}s(s-1)} \bmod p^r \\ &\equiv \left(c_r^t\,(-b_r)\right)^{\frac{1}{2}s^2+\frac{1}{2}s} \left(c^{-t}\,(-b_r)\right)^{\frac{1}{2}s^2-\frac{1}{2}s} \bmod p^r \\ &\equiv {c_r}^{st}\,(-b_r)^{s^2} \bmod p^r.\end{aligned}$$

5. We would like to have a formula for $\frac{h_{t+sN_r}}{h_t}$ that holds for higher powers of p than p^{3r}. (Equivalently, we would like to have a symmetry formula for the columns of $(h_n \bmod p^r)$ when it is written in rows of length N_j for any $j < r$, and in particular for $j = 1$.) However, modulo p^{3r+1} it is no longer true that $\frac{\ell_{s-1}\,\ell_{s+1}}{\ell_s^2}$ is constant over all s for given t, so we cannot obtain a formula of the form

$$\ell_s \equiv f(t)\,g(t)^s\,h(t)^{s^2}\,h_t \bmod p^{3r+1}$$

where $f(t)$, $g(t)$ and $h(t)$ depend on t but not on s.

Finally, we notice another two symmetries in $(h_n \bmod p^{3r})$ which are consequences of Theorem 5.2.1:

Corollary 5.2.2. *Let (h_n) be an EDS, let p be a regular prime with gap $N_1 \geq 4$, and let p^r have gap N_r for all $r \in \mathbb{N}$. Then for all integers s and $t \not\equiv 0 \bmod N_1$,*

$$\frac{h_{t+sN_r}\,h_{t-sN_r}}{h_t^2} \equiv \left(\frac{h_{t+N_r}\,h_{t-N_r}}{h_t^2}\right)^{s^2} \bmod p^{3r}.$$

In other words, if (ℓ_s) is the sequence defined by $\ell_s = \frac{h_{t+sN_1}}{h_t}$ for all $s \in \mathbb{Z}$, then

$$\ell_s \, \ell_{-s} \equiv (\ell_1 \, \ell_{-1})^{s^2} \bmod p^{3r}.$$

Corollary 5.2.3. *Let (h_n) be an EDS, let p be a regular prime with gap $N_1 \geq 4$, and let p^r have gap N_r for all $r \in \mathbb{N}$. Then for all integers s and $t \not\equiv 0 \bmod N_1$,*

$$\frac{h_{t+sN_r}}{h_{t-sN_r}} \equiv \left(\frac{h_{t+N_r}}{h_{t-N_r}} \right)^s \bmod p^{3r}.$$

In other words, if (ℓ_s) is the sequence defined by $\ell_s = \frac{h_{t+sN_1}}{h_t}$ for all $s \in \mathbb{Z}$, then

$$\frac{\ell_s}{\ell_{-s}} \equiv \left(\frac{\ell_1}{\ell_{-1}} \right)^s \bmod p^{3r}.$$

We now consider the sequence (ℓ_n) obtained from (h_n) by taking every N_rth term from h_0 and dividing by h_{N_r} (assuming $h_{N_r} \neq 0$). By Theorem 4.6.1, (ℓ_n) is an EDS.

We first prove that Shipsey's formula (4.10) for $\ell_s = \frac{h_{sN_1}}{h_{N_1}} \bmod p$ holds modulo p^r for any $r \in \mathbb{N}$ if N_1 is replaced by N_r and b_1 and c_1 by b_r and c_r. The proof is similar to hers, but with use of our Theorem 5.1.3 replacing use of Ward's Theorem 4.7.6.

Note that $\frac{h_{sN_r}}{h_{N_r}}$ is an integer for all $s \in \mathbb{Z}$ by the divisibility property of EDSs. Of course if $h_{N_r} = 0$ then $h_{sN_r} = 0$ for all $s \in \mathbb{Z}$.

Theorem 5.2.4. *Let (h_n) be an EDS and p a regular prime with gap $N_1 \geq 4$ in (h_n). Let $r \in \mathbb{N}$, let p^r have gap N_r in (h_n), and let $h_{N_r} \neq 0$.*

Then for all integers s,

$$\frac{h_{sN_r}}{h_{N_r}} \equiv s \, (-b_r)^{s^2-1} \bmod p^r. \tag{5.4}$$

Proof: The result holds trivially for $s = 0$ and $s = 1$, so we may suppose it holds for $s = 0, 1, \ldots, \sigma - 1$ for some $\sigma \geq 2$. We prove it then holds for $s = \sigma$.

By the elliptic sequence formula with $m = (\sigma - 1)N_r + 1$ and $n = N_r - 1$, we have

$$\frac{h_{\sigma N_r}}{h_{N_r}} \cdot h_{2+(\sigma-2)N_r} = h_{2+(\sigma-1)N_r} \left(\frac{h_{(\sigma-1)N_r}}{h_{N_r}} \right) h_{N_r-1}^2 - h_{N_r-2} \cdot h_{1+(\sigma-1)N_r}^2.$$

But by the induction hypothesis,

$$\frac{h_{(\sigma-1)N_r}}{h_{N_r}} \equiv (\sigma-1)\,(-b_r)^{(\sigma-1)^2-1} \bmod p^r,$$

and by Theorem 5.1.3, we can write

$$\begin{aligned}
h_{2+(\sigma-2)N_r} &\equiv {c_r}^{2(\sigma-2)}\,(-b_r)^{(\sigma-2)^2}\,h_2 \bmod p^r,\\
h_{2+(\sigma-1)N_r} &\equiv {c_r}^{2(\sigma-1)}\,(-b_r)^{(\sigma-1)^2}\,h_2 \bmod p^r,\\
h_{1+(\sigma-1)N_r} &\equiv {c_r}^{\sigma-1}\,(-b_r)^{(\sigma-1)^2} \bmod p^r,\\
h_{N_r-1} &\equiv {c_r}^{-1}\,b_r \bmod p^r, \qquad \text{and}\\
h_{N_r-2} &\equiv {c_r}^{-2}\,b_r\,h_2 \bmod p^r.
\end{aligned}$$

So we have

$$\begin{aligned}
&\frac{h_{\sigma N_r}}{h_{N_r}} \cdot \left({c_r}^{2(\sigma-2)}\,(-b_r)^{(\sigma-2)^2}\,h_2\right)\\
&\qquad\equiv {c_r}^{2(\sigma-1)}\,(-b_r)^{(\sigma-1)^2}\,h_2 \cdot (\sigma-1)\,(-b_r)^{(\sigma-1)^2-1} \cdot \left({c_r}^{-1}\,b_r\right)^2\\
&\qquad\qquad - {c_r}^{-2}\,b_r\,h_2 \cdot \left({c_r}^{\sigma-1}\,(-b_r)^{(\sigma-1)^2}\right)^2 \bmod p^r,
\end{aligned}$$

from which

$$\begin{aligned}
\frac{h_{\sigma N_r}}{h_{N_r}} &\equiv \frac{{c_r}^{2\sigma-2-2}(-b_r)^{\sigma^2-2\sigma+1+\sigma^2-2\sigma+2}\,h_2\,(\sigma-1) + {c_r}^{-2+2\sigma-2}(-b_r)^{1+2\sigma^2-4\sigma+2}\,h_2}{{c_r}^{2\sigma-4}(-b_r)^{\sigma^2-4\sigma+4}h_2} \bmod p^r\\
&\equiv \frac{{c_r}^{2\sigma-4}(-b_r)^{2\sigma^2-4\sigma+3}\,\sigma\,h_2}{{c_r}^{2\sigma-4}(-b_r)^{\sigma^2-4\sigma+4}\,h_2} \bmod p^r\\
&\equiv \sigma(-b_r)^{\sigma^2-1} \bmod p^r,
\end{aligned}$$

as required. It follows by induction that (5.4) holds for all $s \geq 0$. Since $h_{sN_r} = -h_{(-s)N_r}$ it follows easily that it also holds for all $s < 0$. This completes the proof. □

In the rest of this section we give an elementary proof of Theorem 4.7.7, in the process extending it to the $p = 2$ case. This result was proved for $p > 3$ by Ward in [31] using elliptic curves.

Theorem 5.2.5. *Let (h_n) be an EDS in which h_0 is the only zero term, and let p be a regular prime with gap $N_1 \geq 4$ in (h_n). For $r \in \mathbb{N}$ let p^r have gap N_r in (h_n).*

If p is odd then

$$\frac{h_{sN_r}}{h_{N_r}} \equiv s\,{\xi_r}^{s^2-1} \bmod p^{2r} \qquad \textit{for all } s \in \mathbb{Z}, \tag{5.5a}$$

where ξ_r is the square root of $-h_{N_r+1}\,h_{N_r-1}$ modulo p^{2r} which reduces to $-b_r$ modulo p^r.

For $p = 2$ and $r \geq 2$ we have

$$\frac{h_{sN_r}}{h_{N_r}} \equiv s\,{\xi_r}^{s^2-1} \bmod 2^{2r-1} \qquad \textit{for all } s \in \mathbb{Z}, \tag{5.5b}$$

where ξ_r is one of the square roots of $-h_{N_r+1}\,h_{N_r-1}$ modulo 2^{2r-2} which reduces to $-b_r$ modulo 2^r.

Proof: Note that the formulae hold trivially for $s = 0$ and $s = 1$. We need to prove they hold for $s = 2$; the result for general s then follows easily by induction.

Since by Theorem 5.1.3

$$\begin{aligned} h_{N_r+2}\,h_{N_r-1}^2 + h_{N_r-2}\,h_{N_r+1}^2 & \\ &\equiv \left(-c_r^2\,b_r\,h_2\right)\,\left(-c_r^{-1}\,b_r\,h_1\right)^2 + \left(-c_r^{-2}\,b_r\,h_{-2}\right)\,\left(-c_r\,b_r\,h_1\right)^2 \bmod p^r \\ &\equiv 0 \bmod p^r, \end{aligned}$$

it follows that

$$\left(h_{N_r+2}\,h_{N_r-1}^2 + h_{N_r-2}\,h_{N_r+1}^2\right)^2 \equiv 0 \bmod p^{2r}. \tag{5.6}$$

But by the doubling formula

$$\left(h_{N_r+2}\,h_{N_r-1}^2 - h_{N_r-2}\,h_{N_r+1}^2\right)^2 \equiv h_2^2\,\left(\frac{h_{2N_r}}{h_{N_r}}\right)^2 \bmod p^{2r}. \tag{5.7}$$

Multiplying out and subtracting (5.6) from (5.7) we get

$$-4\,\left(h_{N_r+2}\,h_{N_r-2}\right)\,\left(h_{N_r+1}\,h_{N_r-1}\right)^2 \equiv h_2^2\,\left(\frac{h_{2N_r}}{h_{N_r}}\right)^2 \bmod p^{2r}.$$

But by (4.4) with $m = N_r$

$$h_{N_r+2}\,h_{N_r-2} = h_{N_r+1}\,h_{N_r-1}\,h_2^2 - h_1\,h_3\,h_{N_r}^2 \equiv h_{N_r+1}\,h_{N_r-1}\,h_2^2 \bmod p^{2r},$$

and it follows that

$$\left(\frac{h_{2N_r}}{h_{N_r}}\right)^2 \equiv 4\,\left(-h_{N_r+1}\,h_{N_r-1}\right)^3 \bmod p^{2r}. \tag{5.8}$$

If p is odd then we have, taking square roots,

$$\frac{h_{2N_r}}{h_{N_r}} \equiv 2\,{\xi_r}^3 \bmod p^{2r},$$

where ${\xi_r}^2 \equiv -h_{N_r+1}\,h_{N_r-1} \bmod p^{2r}$. Since we know from Theorem 5.2.4 that $\frac{h_{2N_r}}{h_{N_r}} \equiv 2\,(-b_r)^3 \bmod p^r$ it follows that $\xi_r \equiv -b_r \bmod p^r$.

If $p = 2$ and $r \geq 2$ then (5.8) becomes

$$\left(\frac{h_{2N_r}}{2\,h_{N_r}}\right)^2 \equiv (-h_{N_r+1}\,h_{N_r-1})^3 \bmod 2^{2r-2},$$

from which (taking square roots)

$$\frac{h_{2N_r}}{2\,h_{N_r}} \equiv {\xi_r}^3 \bmod 2^{2r-2},$$

i.e.,

$$\frac{h_{2N_r}}{h_{N_r}} \equiv 2\,{\xi_r}^3 \bmod 2^{2r-1},$$

where ${\xi_r}^2 \equiv -h_{N_r+1}\,h_{N_r-1} \bmod 2^{2r-2}$ and $\xi_r \equiv -b_r \bmod 2^r$.

Now suppose the result holds for $s = 0, 1, \ldots, \sigma$ for some $\sigma \geq 2$; we want to prove it holds for $s = \sigma+1$. By the elliptic sequence formula (4.1) with $m = \sigma\,N_r$ and $n = N_r$,

$$\frac{h_{(\sigma+1)N_r}}{h_{N_r}} \cdot \frac{h_{(\sigma-1)N_r}}{h_{N_r}} = h_{\sigma N_r+1}\,h_{\sigma N_r-1} - h_{N_r+1}\,h_{N_r-1}\left(\frac{h_{\sigma N_r}}{h_{N_r}}\right)^2.$$

Hence by Corollary 5.2.2

$$\frac{h_{(\sigma+1)N_r}}{h_{N_r}} \cdot \frac{h_{(\sigma-1)N_r}}{h_{N_r}} \equiv -\left(-h_{N_r+1}\,h_{N_r-1}\right)^{\sigma^2} - h_{N_r+1}\,h_{N_r-1}\left(\frac{h_{\sigma N_r}}{h_{N_r}}\right)^2 \bmod p^{3r}.$$

If p is odd, then by the induction hypothesis,

$$\begin{aligned}\frac{h_{(\sigma+1)N_r}}{h_{N_r}} \cdot (\sigma-1)\,{\xi_r}^{(\sigma-1)^2-1} &\equiv -\left({\xi_r}^2\right)^{\sigma^2} + {\xi_r}^2\left(\sigma\,{\xi_r}^{\sigma^2-1}\right)^2 \bmod p^{2r}\\ &\equiv \left(\sigma^2-1\right)\,{\xi_r}^{2\sigma^2} \bmod p^{2r},\end{aligned}$$

where ${\xi_r}^2 \equiv -h_{N_r+1}\,h_{N_r-1} \bmod p^{2r}$. Hence (since $\sigma \neq 1$)

$$\begin{aligned}\frac{h_{(\sigma+1)N_r}}{h_{N_r}} &\equiv \frac{(\sigma+1)\;(\sigma-1)}{(\sigma-1)}\,{\xi_r}^{2\sigma^2-(\sigma^2-2\sigma)} \bmod p^{2r}\\ &\equiv (\sigma+1)\,{\xi_r}^{(\sigma+1)^2-1} \bmod p^{2r},\end{aligned}$$

as required. The $p = 2$ case is proved by a similar argument. □

5.3 The behaviour of N_r as r increases

In this section we give an elementary proof of Ward's Theorem 4.7.5 on the gap N_r, simultaneously extending it to the $p = 2$ case. (For the odd p case the proof is the same as Ward's in [31].)

Theorem 5.3.1. *Let (h_n) be an EDS in which h_0 is the only zero term, and let p be a regular prime with gap $N_1 \geq 4$ in (h_n). For $r \in \mathbb{N}$ let p^r have gap N_r in (h_n). Let p^w be the highest power of p dividing h_{N_1}.*

If p is odd, or $p = 2$ and $w \geq 2$, then for any $r \in \mathbb{N}$, p^r has gap

$$N_r = \begin{cases} N_1 & \text{if } r \leq w \\ p\, N_{r-1} & \text{if } r > w. \end{cases}$$

If $p = 2$ and $w = 1$ then for some $v \geq 2$,

$$N_r = \begin{cases} N_1 & \text{if } r = 1, \\ 2N_1 & \text{if } 2 \leq r \leq v, \\ 2N_{r-1} & \text{if } r > v. \end{cases}$$

Proof: We first show that $N_{r+1} = N_r$ or $p\, N_r$ for all $r \geq 1$. By Theorem 5.2.4 with $s = p$

$$\frac{h_{p\, N_r}}{h_{N_r}} \equiv p\, (-b_r)^{p^2-1} \bmod p^r.$$

Hence $h_{p\, N_r}$ is divisible by at least one higher power of p than h_{N_r}. It follows (since p^{r+1} is regular in (h_n) with gap N_{r+1}) that N_{r+1} divides $p\, N_r$, and hence that $N_{r+1} = N_r$ or $p\, N_r$ for all $r \geq 1$.

Now if p is odd, or $p = 2$ and $r \geq 2$, then by Theorem 5.2.5 (since $h_{N_r} \neq 0$) we have

$$\frac{h_{pN_r}}{h_{N_r}} \equiv p\, {\xi_r}^{p^2-1} \bmod p^2.$$

So $\frac{h_{p\, N_r}}{h_{N_r}}$ is divisible by p but *not* by p^2; that is, $h_{p\, N_r}$ is divisible by *exactly* one higher power of p than h_{N_r} is. It follows that if p is odd then

$$N_{r+1} = p\, N_r \quad \text{for all } r \geq w,$$

and if $p = 2$ then

$$N_{r+1} = p\,N_r \quad \text{for all } r \geq v.$$

This completes the proof. □

5.4 Writing b_r and c_r in terms of b_w and c_w

In this section we use Theorem 5.2.1 to find expressions for the constants c_r and b_r in terms of the constants b_w and c_w, and hence in terms of h_2, h_{N_1-1} and h_{N_1-2} modulo p^w, in the case where p is an odd prime. (So for all $r \in \mathbb{N}$, $\frac{h_{t+sN_r}}{h_t} \bmod p^r$ can be calculated from the sequence modulo p^w.)

Of course, if $r \leq w$ (i.e., $p^r \mid h_{N_1}$), then $N_r = N_w = N_1$ and so $c_r \equiv c_w \bmod p^r$ and $b_r \equiv b_w \bmod p^r$ by definition of c_r and b_r.

Theorem 5.4.1. *Let (h_n) be an EDS, and let p be an odd regular prime with gap $N_1 \geq 4$ in (h_n). Let p^w be the highest power of p dividing h_{N_1}. Then for $r \geq w$,*

$$c_r \equiv {c_w}^{p^{r-w}} \bmod p^r \quad \textit{and} \quad b_r \equiv {b_w}^{p^{2(r-w)}} \bmod p^r.$$

Proof: We first show that for odd p and $r \geq w$,

$$c_{r+1} \equiv {c_r}^p \bmod p^{r+1} \quad \text{and} \quad b_{r+1} \equiv {b_r}^{p^2} \bmod p^{r+1}.$$

Since $r \geq w$, $N_{r+1} = pN_r$ and

$$c_{r+1} \equiv \frac{h_{-1+pN_r}\,h_2}{h_{-2+pN_r}} \bmod p^{r+1}.$$

Using Theorem 5.2.1 with $s = p$ and the fact that p is odd we can write this as

$$\begin{aligned} c_{r+1} &\equiv \frac{\left(\frac{h_{1+N_r}}{h_1}\right)^{\frac{1}{2}(-p)(-p+1)} \left(\frac{h_{1-N_r}}{h_1}\right)^{\frac{1}{2}(-p)(-p-1)}}{\left(\frac{h_{2+N_r}}{h_2}\right)^{\frac{1}{2}(-p)(-p+1)} \left(\frac{h_{2-N_r}}{h_2}\right)^{\frac{1}{2}(-p)(-p-1)}} \bmod p^{r+1} \\ &\equiv \left(\frac{\left(\frac{h_{1+N_r}}{h_1}\right)^{\frac{1}{2}(p-1)} \left(\frac{h_{1-N_r}}{h_1}\right)^{\frac{1}{2}(p+1)}}{\left(\frac{h_{2+N_r}}{h_2}\right)^{\frac{1}{2}(p-1)} \left(\frac{h_{2-N_r}}{h_2}\right)^{\frac{1}{2}(p+1)}} \right)^p \bmod p^{r+1}. \end{aligned}$$

But by Theorem 5.1.3 the bracketed quantity is

$$\frac{\left(\frac{h_{1+N_r}}{h_1}\right)^{\frac{1}{2}(p-1)} \left(\frac{h_{1-N_r}}{h_1}\right)^{\frac{1}{2}(p+1)}}{\left(\frac{h_{2+N_r}}{h_2}\right)^{\frac{1}{2}(p-1)} \left(\frac{h_{2-N_r}}{h_2}\right)^{\frac{1}{2}(p+1)}} \equiv \frac{(-c_r\, b_r)^{\frac{1}{2}(p-1)}\ (-{c_r}^{-1}\, b_r)^{\frac{1}{2}(p+1)}}{(-c_r^2\, b_r)^{\frac{1}{2}(p-1)}\ (-{c_r}^{-2}\, b_r)^{\frac{1}{2}(p+1)}} \bmod p^r$$

$$\equiv c_r \bmod p^r,$$

so by Theorem 2.2.5,

$$c_{r+1} \equiv {c_r}^p \bmod p^{r+1}.$$

Similarly, we can write

$$\begin{aligned} b_{r+1} &\equiv c_{r+1}\, h_{-1+pN_r} \bmod p^{r+1} \\ &\equiv {c_r}^p \left(\frac{h_{-1+N_r}}{h_{-1}}\right)^{\frac{1}{2}p(p+1)} \left(\frac{h_{-1-N_r}}{h_{-1}}\right)^{\frac{1}{2}p(p-1)} h_{-1} \bmod p^{r+1} \\ &\equiv \left(-c_r\, {h_{1-N_r}}^{\frac{1}{2}(p+1)}\, {h_{1+N_r}}^{\frac{1}{2}(p-1)}\right)^p \bmod p^{r+1}. \end{aligned}$$

But

$$\begin{aligned} -c_r\, {h_{1-N_r}}^{\frac{1}{2}(p+1)}\, {h_{1+N_r}}^{\frac{1}{2}(p-1)} &\equiv -c_r \left(-{c_r}^{-1}\, b_r\right)^{\frac{1}{2}(p+1)}\ (-c_r\, b_r)^{\frac{1}{2}(p-1)} \bmod p^r \\ &\equiv -(-b_r)^p \bmod p^r \\ &\equiv {b_r}^p \bmod p^r, \end{aligned}$$

and so by Theorem 2.2.5,

$$b_{r+1} \equiv {b_r}^{p^2} \bmod p^{r+1}.$$

It now follows by an easy induction that

$$c_r \equiv {c_w}^{p^{r-w}} \bmod p^r \text{ and } b_r \equiv {b_w}^{p^{2(r-w)}} \bmod p^r$$

for all $r \geq w$. □

We can write b_{r+1} in terms of b_r and c_r in another way as follows. (We will refer to this form in chapter 8.)

Corollary 5.4.2. *If p is an odd prime then for $r \geq w$,*

$$c_{r+1} \equiv {c_r}^p \bmod p^{r+1}$$

and

$$b_{r+1} \equiv \left({c_r}^{\frac{1}{2}(p-1)\,N_r}\, b_r \right)^p \bmod p^{r+1}.$$

Proof: Since $b_r^2 \equiv {c_r}^{N_r} \bmod p^r$,

$$\begin{aligned} b_{r+1} &\equiv {b_r}^{p^2} \bmod p^{r+1} \\ &\equiv \left(b_r^2\right)^{\frac{1}{2}p(p-1)}\, {b_r}^p \bmod p^{r+1} \\ &\equiv \left({c_r}^{N_r}\right)^{\frac{1}{2}p\,(p-1)}\, {b_r}^p \bmod p^{r+1} \\ &\equiv \left({c_r}^{\frac{1}{2}(p-1)\,N_r}\, b_r \right)^p \bmod p^{r+1}. \end{aligned}$$

□

It now follows by an easy induction that

Corollary 5.4.3. *If p is an odd prime then for $r \geq w$,*

$$c_r \equiv {c_w}^{p^{r-w}} \bmod p^r$$

and

$$b_r \equiv \left({c_w}^{\frac{1}{2}N_1\,(p^{r-w}-1)}\, b_w \right)^{p^{r-w}} \bmod p^r.$$

Example: Let (h_n) be the EDS

$$\ldots, -3,\ -4,\ -1,\ -1,\ 0,\ 1,\ 1,\ 4,\ 3,\ -61,\ -280,\ -3931,\ -24603,\ 900364, \ldots$$

and let $p = 5$. Here

$$N_1 = 6, \quad N_2 = 30 \quad \text{and} \quad N_3 = 150,$$

so $w = u = 1$.

We have $c_1 = 3$, $b_1 = 2$,

$$c_2 = (3)^5 \equiv 18 \bmod 25, \quad b_2 = (2)^{25} \equiv 7 \bmod 25,$$

and

$$c_3 = (18)^5 \equiv 68 \bmod 125, \quad b_3 = (7)^{25} \equiv 57 \bmod 125.$$

Remarks:

1. If p is an odd prime with gap $N_1 \geq 4$, then b_r and c_r reduce to b_1 and c_1 respectively modulo p, for all $r \in \mathbb{N}$. (This is because $x^p \equiv x \bmod p$ for all integers x, so by Theorem 5.4.1 b_r and c_r reduce to b_w and c_w respectively modulo p. But since $N_w = N_1$, by definition b_w and c_w reduce to b_1 and c_1 modulo p.)

2. Note however that b_r and c_r (where $r \geq w+1$) do not necessarily reduce to b_w and c_w modulo p^w unless $w = 1$:

$$b_r \equiv {b_w}^{p^{2(r-w)}} \equiv b_w \bmod p^w \;\Leftrightarrow\; {b_w}^{p^{2(r-w)}-1} \equiv 1 \bmod p^w,$$

which is true if and only if the order m_w of b_w in $\mathbb{Z}_{p^w}^*$ divides

$$p^{2(r-w)} - 1 = (p-1)\left(p^{2(r-w)-1} + p^{2(r-w)-2} + \ldots + p + 1\right).$$

Since m_w also divides $p^{w-1}(p-1)$ and p^{w-1} is coprime to $p^{2(r-w)-1} + p^{2(r-w)-2} + \ldots + p + 1$, this is true if and only if m_w divides $(p-1)$, i.e., if and only if $b_w \equiv g^{p^{w-1}s} \bmod p^w$ for a generator g of $\mathbb{Z}_{p^w}$. Heuristically (if b_w were a random element of $\mathbb{Z}_{p^w}^*$), we would expect this to happen about $\frac{1}{p^{w-1}}$ of the time.

Similarly, ${c_w}^{p^{r-w}} \equiv c_w \bmod p^w$ if and only if the order of c_w modulo p^w divides $(p-1)$.

Incidentally, if $w = 1$ (which is usually the case) then we have the following two corollaries:

Corollary 5.4.4. *If p is an odd prime and $w = 1$ then*

$$c_r \equiv c_1^{p^{r-1}} \bmod p^r \quad \text{and} \quad b_r \equiv b_1^{p^{r-1}} \bmod p^r.$$

Corollary 5.4.5. *If p is an odd prime and $w = 1$ then*

$$\frac{h_{t+sN_r}}{h_t} \equiv \left(\frac{h_{t+sN_1}}{h_t}\right)^{p^{r-1}} \bmod p^r.$$

Proof: By Theorem 5.1.3,

$$\frac{h_{t+sN_1}}{h_t} \equiv c_1^{st}\,(-b_1)^{s^2} \bmod p.$$

This implies (by Theorem 2.2.5) that

$$\begin{aligned}\left(\frac{h_{t+sN_1}}{h_t}\right)^{p^{r-1}} &\equiv \left(c_1^{st}\,(-b_1)^{s^2}\right)^{p^{r-1}} \bmod p^r \\ &\equiv \left(c_1^{p^{r-1}}\right)^{st} \left((-b_1)^{p^{r-1}}\right)^{s^2} \bmod p^r.\end{aligned}$$

But by Corollary 5.4.5, since $w = 1$, we can write $c_1^{p^{r-1}} \equiv c_r \bmod p^r$ and $(-b_1)^{p^{r-1}} \equiv (-b_r) \bmod p^r$. So

$$\begin{aligned}\left(\frac{h_{t+sN_1}}{h_t}\right)^{p^{r-1}} &\equiv {c_r}^{st}\,(-b_r)^{s^2} \bmod p^r \\ &\equiv \frac{h_{t+sN_r}}{h_t} \bmod p^r.\end{aligned}$$

□

In other words, if $w = 1$ then for all $s \in \mathbb{Z}$ we can find the sth term of the sequence $\left(\frac{h_{t+sN_r}}{h_t} \bmod p^r\right)$ by raising the sth term of the sequence $\left(\frac{h_{t+sN_1}}{h_t} \bmod p\right)$ to the power p^{r-1}.

5.5 The order of b_r and c_r

In this section we consider the order of the constants b_r and c_r in $\mathbb{Z}_{p^r}^*$. We will use these results when we find the period of $(h_n \bmod p^r)$ in the next section.

Definition: Let (h_n) be an elliptic divisibility sequence and p a regular prime with gap $N_1 \geq 3$ in (h_n). For $r \in \mathbb{N}$ let n_r and m_r denote the order of the constants c_r and b_r in $\mathbb{Z}_{p^r}^*$ respectively.

We first show how n_r and m_r are related:

Theorem 5.5.1. *Let (h_n) be an EDS, and let p be a regular prime with gap $N_1 \geq 4$ in (h_n). For $r \in \mathbb{N}$ let p^r have gap N_r in (h_n), and denote the order of ${c_r}^{N_r}$ by $\xi = \frac{n_r}{\gcd(n_r, N_r)}$. Then*

$$m_r = \begin{cases} \xi \text{ or } 2\xi & \text{if } n_r \text{ and } N_r \text{ are both odd, or} \\ \xi & \text{(which is odd) if } N_r \text{ is divisible by a higher power of 2 than } n_r, \text{ or} \\ 2\xi & \text{otherwise.} \end{cases}$$

So m_r divides $2\,n_r$.

Proof: If N_r is even then by Lemma 5.1.2 we have $b_r \equiv {c_r}^{\frac{N_r}{2}} \bmod p^r$, and so b_r has order $m_r = \frac{n_r}{\gcd(n_r, \frac{N_r}{2})}$. If N_r is divisible by a higher power of 2 than n_r is then this is $m_r = \xi$; otherwise $m_r = 2\xi$.

If N_r is odd, then by Lemma 5.1.2,

$$b_r^2 \equiv {c_r}^{N_r} \bmod p^r.$$

The order of ${c_r}^{N_r}$ is $\frac{n_r}{\gcd(n_r, N_r)} = \xi$. Since b_r^2 therefore has order ξ, and b_r^2 also has order $\frac{m_r}{\gcd(m_r, 2)}$, it follows that

$$m_r = \xi \gcd(m_r, 2).$$

Since N_r is odd, ξ has the same parity as n_r. So if n_r is even, then m_r is even and $m_r = 2\,\xi$. If n_r is odd, then m_r could be ξ or $2\,\xi$. □

We next consider how the order of b_r and c_r modulo p^r changes as r increases. This will turn out to be related to the periodicity behaviour of $(h_n \bmod p^r)$ as r increases.

Theorem 5.5.2. *Let (h_n) be an EDS, let p be a regular odd prime with gap $N_1 \geq 4$ in (h_n), and for $r \in \mathbb{N}$ let p^r have gap N_r in (h_n). Then there exist integers u and v with $1 \leq u \leq v \leq w$ such that*

$$n_r = \begin{cases} n_1 & \text{for } 1 \leq r \leq u \\ p\,n_{r-1} & \text{for } u < r \leq w \\ n_w & \text{for } r \geq w, \end{cases}$$

and

$$m_r = \begin{cases} m_1 & \text{for } 1 \le r \le v \\ p\, m_{r-1} & \text{for } v < r \le w \\ \frac{m_{r-1}}{p} & \text{for } w < r \le 2w - v \\ m_1 & \text{for } r \ge 2w - v. \end{cases}$$

Proof: We first consider the case $r \le w$. Since p is odd, $\mathbb{Z}_{p^w}^*$ is cyclic; let g be a generator for $\mathbb{Z}_{p^w}^*$. Then $c_w \equiv \frac{h_{N_1-1}\, h_2}{h_{N_1-2}} \equiv g^s \bmod p^w$ for some $1 \le s \le p^{w-1}(p-1)$. Let p^{u-1} be the highest power of p dividing s. (Clearly $u \le w$.)

Since $N_r = N_1$ for $r \le w$, we have $c_r \equiv g^s \bmod p^r$, and so the order of c_r in $\mathbb{Z}_{p^r}^*$ is

$$n_r = \frac{p^{r-1}\,(p-1)}{\gcd\,(p^{r-1}\,(p-1), s)}$$

for all $r \le w$.

For $r \le u$, $p^{r-1} \mid s$, and so $\gcd\,(p^{r-1}\,(p-1), s) = p^{r-1} \gcd(p-1, s)$. Hence

$$n_r = \frac{p-1}{\gcd(p-1, s)} = n_1.$$

For $u \le r \le w$, $\gcd\,(p^{r-1}\,(p-1), s) = p^{u-1} \gcd(p-1, s)$, so

$$n_r = \frac{p^{r-1}(p-1)}{p^{u-1} \gcd(p-1, s)} = p^{r-u} \cdot \frac{p-1}{\gcd(p-1, s)} = p^{r-u}\, n_1.$$

The proof for m_r in the case $r \le w$ is similar.

By Theorem 5.5.1, $m_r \mid 2\, n_r$ for all $r \in \mathbb{N}$. Since $p \ne 2$ this means that if $p \mid m_r$ then $p \mid n_r$, and it follows that $u \le v$.

Now let $r \ge w$. Then by Theorem 5.4.1 $c_r \equiv {c_w}^{p^{r-w}} \bmod p^r$. Hence by Theorem 2.2.6, $n_r = n_w$. We also have

$$b_r \equiv b_{r-1}^{p^2} \equiv (b_{r-1}^p)^p \bmod p^r,$$

so by Theorem 2.2.6, b_{r-1}^p has order m_{r-1} in $\mathbb{Z}_{p^r}^*$. It follows that b_r has order

$$m_r = \frac{m_{r-1}}{\gcd(m_{r-1}, p)}$$

in $\mathbb{Z}_{p^r}^*$. Thus m_r is equal to $\frac{m_{r-1}}{p}$ if m_{r-1} is divisible by p, or m_{r-1} otherwise. The result follows. □

Corollary 5.5.3. *If p is an odd prime and $N_1 \ge 4$ then n_r and m_r have the same parity as n_1 and m_1 respectively, for all $r \in \mathbb{N}$.*

5.6 The period of $(h_n \bmod p^r)$

In this section we prove that if p is an odd prime and p^w is the highest power of p dividing h_{N_1} then for all $r \geq w$ the sequence $(h_n \bmod p^r)$ is periodic with period $\tau_w N_r = p^{r-w} \tau_w N_1$, where τ_w is a constant that can be found from the sequence $(h_n \bmod p^w)$. This confirms Shipsey's Conjecture 4.7.11 for the $r = 2$ and $w = 1$ case.

We first define a constant τ_r associated with the sequence $(h_n \bmod p^r)$:

Definition: Let (h_n) be an elliptic divisibility sequence and p be a regular prime with gap $N_1 \geq 3$ in (h_n). For $r \in \mathbb{N}$ we define τ_r to be the least positive integer for which

$$c_r{}^{\tau_r} \equiv (-b_r)^{\tau_r^2} \equiv 1 \bmod p^r.$$

Our first result is simply an adaptation of Ward's proof in [30] of the modulo p case, but we include the proof because it's short.

Theorem 5.6.1. *Let (h_n) be an elliptic divisibility sequence, and p a regular prime with gap $N_1 \geq 4$ in (h_n). For $r \in \mathbb{N}$ let the gap of p^r in (h_n) be N_r. Then $(h_n \bmod p^r)$ is periodic with period $\pi_r = \tau_r N_r$. Furthermore, τ_r divides $p^{r-1}(p-1)$.*

Proof: By Theorem 5.1.3,

$$h_{t+\tau_r N_r} \equiv c_r{}^{\tau_r t} (-b_r)^{\tau_r^2} h_t \equiv h_t \bmod p^r$$

for all $t \in \mathbb{Z}$. It follows that $(h_n \bmod p^r)$ is periodic, with period π_r dividing $\tau_r N_r$. Furthermore, since $h_n \equiv 0 \bmod p^r$ if and only if $n \equiv 0 \bmod N_r$, the period must be a multiple of N_r — say $\pi_r = v N_r$. Hence $v \mid \tau_r$.

Now since

$$h_t \equiv h_{vN_r+t} \equiv c_r{}^{vt} (-b_r)^{v^2} h_t \bmod p^r$$

for all $t \in \mathbb{Z}$, we have $c_r{}^{vt}(-b_r)^{v^2} \equiv 1 \bmod p^r$ for all $t \in \mathbb{Z}$. It follows (setting $t = 2$ and $t = 1$ and dividing) that $c_r{}^v \equiv 1 \bmod p^r$ and hence that $(-b_r)^{v^2} \equiv 1 \bmod p^r$. Hence $v \geq \tau_r$, and so $v = \tau_r$.

Finally, note that since $\mathbb{Z}_{p^r}^*$ has $p^{r-1}(p-1)$ elements,

$$h_{p^{r-1}(p-1)N_r+t} \equiv {c_r}^{p^{r-1}(p-1)t}(-b_r)^{\left(p^{r-1}(p-1)\right)^2} h_t \equiv h_t \bmod p^r$$

for all integers t. It follows that the period $\tau_r N_r$ divides $p^{r-1}(p-1)N_r$, and hence that τ_r divides $p^{r-1}(p-1)$. □

So τ_r is the number of zeroes before $(h_n \bmod p^r)$ starts repeating.

An equivalent definition for τ_r is the following:

Theorem 5.6.2. *The constant τ_r is also the least positive integer for which*

$${c_r}^{\tau_r} \equiv 1 \bmod p^r$$

and

$$(-b_r)^{\tau_r} \equiv {c_r}^{-\frac{1}{2}\tau_r(\tau_r-1)N_r} \bmod p^r.$$

Proof: By Lemma 5.1.2, $(-b_r)^2 \equiv {c_r}^{N_r} \bmod p^r$. So

$$\begin{aligned}(-b_r)^{\tau_r}\,{c_r}^{\frac{1}{2}\tau_r(\tau_r-1)N_r} &\equiv (-b_r)^{\tau_r}\,(-b_r^2)^{\frac{1}{2}\tau_r(\tau_r-1)} \bmod p^r \\ &\equiv (-b_r)^{\tau_r^2} \bmod p^r,\end{aligned}$$

and the result now follows from the definition of τ_r. □

It turns out that if p is an odd prime then we can find τ_r from the order of c_r and b_r in $\mathbb{Z}_{p^r}^*$:

Theorem 5.6.3. *Let (h_n) be an elliptic divisibility sequence and p a regular odd prime with gap $N_1 \geq 4$ in (h_n). For $r \in \mathbb{N}$ let the gap of p^r in (h_n) be N_r, and let c_r and b_r have order n_r and m_r respectively in $\mathbb{Z}_{p^r}^*$. Then*

$$\tau_r = \begin{cases} n_r & \text{if} \quad \text{either } n_r \text{ or } m_r \text{ is even, or} \\ 2n_r & \text{if} \quad \text{both } n_r \text{ and } m_r \text{ are odd.} \end{cases}$$

Proof: First note that, since ${c_r}^{\tau_r} \equiv 1 \bmod p^r$, τ_r is a multiple of n_r. By Lemma 5.1.2, ${b_r}^{2n_r} \equiv {c_r}^{N_r n_r} \equiv 1 \bmod p^r$, so by Theorem 2.2.4 (since p is odd)

$${b_r}^{n_r} \equiv \pm 1 \bmod p^r.$$

If n_r is even, then

$$(-b_r)^{n_r^2} \equiv ({b_r}^{n_r})^{n_r} \equiv (\pm 1)^{n_r} \equiv 1 \bmod p^r,$$

so $\tau_r \leq n_r$ and hence $\tau_r = n_r$.

If n_r is odd, then (since ${b_r}^{n_r} \equiv \pm 1 \bmod p^r$)

$$(-b_r)^{n_r^2} \equiv -({b_r}^{n_r})^{n_r} \equiv -{b_r}^{n_r} \bmod p^r.$$

So if ${b_r}^{n_r} \equiv -1 \bmod p^r$ then $(-b_r)^{n_r^2} \equiv 1 \bmod p^r$. Then $\tau_r \leq n_r$ by definition of τ_r, and hence $\tau_r = n_r$. If ${b_r}^{n_r} \equiv 1 \bmod p^r$ then $(-b_r)^{n_r^2} \equiv -1 \bmod p^r$ and $\tau_r \neq n_r$, but

$$(-b_r)^{(2n_r)^2} \equiv ({b_r}^{n_r})^{4n_r} \equiv 1 \bmod p^r,$$

so $\tau_r = 2n_r$. Finally, we note that since ${b_r}^{2n_r} \equiv 1 \bmod p^r$, m_r divides $2\,n_r$. If m_r is even then the odd n_r cannot be a multiple of m_r, so ${b_r}^{n_r} \equiv -1 \bmod p^r$ and we have $\tau_r = n_r$. If m_r is odd then $m_r \mid 2n_r$ implies $m_r \mid n_r$, so ${b_r}^{n_r} \equiv 1 \bmod p^r$ and we have $\tau_r = 2n_r$. □

Note that the restriction to odd primes is necessary: if $p = 2$ then $\tau_1 \mid p-1$ implies $\tau_1 = 1$, but $b_1 = c_1 = 1$, so $n_1 = m_1 = 1$, and Theorem 5.6.3 gives $\tau_1 = 2$.

We need to check that Theorem 5.6.3 agrees with Ward's Theorem 4.7.10:

Corollary 5.6.4. *[30]*

Let (h_n) be an elliptic divisibility sequence, and let p be a regular odd prime with gap $N_1 \geq 4$ in (h_n). For $r \in \mathbb{N}$ let the gap of p^r in (h_n) be N_r.

Let ϵ_r and κ_r be the orders in $\mathbb{Z}_{p^r}^*$ of

$$h_2\,(h_{N_r-2})^{-1} \equiv c_r^2\,{b_r}^{-1} \bmod p^r$$

and

$$h_{N_r-1} \equiv b_r\,{c_r}^{-1} \bmod p^r$$

respectively, and let the constant α_r be given by

$$\alpha_r = \begin{cases} 1 & \text{if } \epsilon_r \text{ and } \kappa_r \text{ are both odd,} \\ 0 & \text{if } \epsilon_r \text{ and } \kappa_r \text{ are divisible by different powers of } 2, \text{ and} \\ -1 & \text{otherwise.} \end{cases}$$

Then

$$\tau_r = 2^{\alpha_r}\ \mathrm{lcm}(\epsilon_r, \kappa_r).$$

Proof: First note that

$$c_r{}^{\mathrm{lcm}(\epsilon_r,\kappa_r)} \equiv \left(c_r^2\, b_r{}^{-1}\right)^{\mathrm{lcm}(\epsilon_r,\kappa_r)}\ \left(b_r\, c_r{}^{-1}\right)^{\mathrm{lcm}(\epsilon_r,\kappa_r)} \equiv 1 \bmod p^r,$$

so

$$n_r \mid \mathrm{lcm}(\epsilon_r, \kappa_r).$$

Recall that $b_r{}^{n_r} \equiv 1$ or $-1 \bmod p^r$ (i.e., that $m_r \mid 2\, n_r$).

First suppose $b_r{}^{n_r} \equiv -1 \bmod p^r$ (i.e., that $m_r \nmid n_r$). Then

$$\left(c_r^2\, b_r{}^{-1}\right)^{n_r} \equiv -1 \bmod p^r \quad \text{and} \quad \left(b_r\, c_r{}^{-1}\right)^{n_r} \equiv -1 \bmod p^r,$$

so ϵ_r and κ_r do not divide n_r but do divide $2n_r$. It follows that

$$\mathrm{lcm}(\epsilon_r, \kappa_r) = 2n_r,$$

and that ϵ_r and κ_r are both even, and both divisible by exactly one higher power of 2 than n_r is; so $\alpha_r = -1$. Since $m_r \nmid n_r$ but $m_r \mid 2n_r$, m_r must be even, and so by Theorem 5.6.3,

$$\tau_r = n_r = 2^{-1}\,(2\, n_r) = 2^{\alpha_r}\,\mathrm{lcm}(\epsilon_r, \kappa_r),$$

as required.

Suppose from now on that $b_r{}^{n_r} \equiv 1 \bmod p^r$ (i.e., that $m_r \mid n_r$). Then

$$\left(c_r^2\, b_r{}^{-1}\right)^{n_r} \equiv 1 \bmod p^r \quad \text{and} \quad \left(b_r\, c_r{}^{-1}\right)^{n_r} \equiv 1 \bmod p^r,$$

so $\epsilon_r \mid n_r$ and $\kappa_r \mid n_r$, and it follows that

$$\mathrm{lcm}(\epsilon_r, \kappa_r) = n_r.$$

If n_r is odd, then ϵ_r and κ_r are both odd, so $\alpha_r = 1$. Since $m_r \mid n_r$, m_r is also odd, so by Theorem 5.6.3

$$\tau_r = 2\, n_r = 2^{\alpha_r}\,\mathrm{lcm}(\epsilon_r, \kappa_r),$$

as required.

If n_r is even, then ${b_r}^{\frac{n_r}{2}} \equiv \pm 1 \bmod p^r$. If ${b_r}^{\frac{n_r}{2}} \equiv 1 \bmod p^r$ then

$$\left({c_r}^2 {b_r}^{-1}\right)^{\frac{n_r}{2}} \equiv {c_r}^{n_r} \cdot \left({b_r}^{\frac{n_r}{2}}\right)^{-1} \equiv 1 \bmod p^r$$

and

$$\left(b_r {c_r}^{-1}\right)^{\frac{n_r}{2}} \equiv {b_r}^{\frac{n_r}{2}} \cdot \left({c_r}^{\frac{n_r}{2}}\right)^{-1} \equiv (1)(-1) \equiv -1 \bmod p^r,$$

so $\epsilon_r \mid \frac{n_r}{2}$ and $\kappa_r \nmid \frac{n_r}{2}$. Otherwise, if ${b_r}^{\frac{n_r}{2}} \equiv -1 \bmod p^r$ then

$$\left(b_r {c_r}^{-1}\right)^{\frac{n_r}{2}} \equiv {b_r}^{\frac{n_r}{2}} \cdot \left({c_r}^{\frac{n_r}{2}}\right)^{-1} \equiv (-1)(-1) \equiv 1 \bmod p^r$$

and

$$\left({c_r}^2 {b_r}^{-1}\right)^{\frac{n_r}{2}} \equiv {c_r}^{n_r} \cdot \left({b_r}^{\frac{n_r}{2}}\right)^{-1} \equiv (1)(-1) \equiv -1 \bmod p^r,$$

so $\kappa_r \mid \frac{n_r}{2}$ and $\epsilon_r \nmid \frac{n_r}{2}$. Since ϵ_r and κ_r both divide n_r, it follows that if n_r is even then ϵ_r and κ_r are divisible by different powers of 2, i.e., that $\alpha_r = 0$. But since n_r is even, by Theorem 5.6.3

$$\tau_r = n_r = 2^0 \, n_r = 2^{\alpha_r} \operatorname{lcm}(\epsilon_r, \kappa_r),$$

as required. □

Theorem 5.6.3 together with Theorem 5.5.2 leads to a result about the behaviour of τ_r as r increases:

Theorem 5.6.5. *Let (h_n) be an elliptic divisibility sequence, let p be a regular odd prime with gap $N_1 \geq 4$ in (h_n), and let p^w be the highest power of p dividing N_1. Then there exists a positive integer $u \leq w$ such that*

$$\tau_r = \begin{cases} \tau_1 & \text{for} \quad r \leq u \\ p^{r-u} \, \tau_1 & \text{for} \quad u \leq r \leq w, \text{ and} \\ \tau_w & \text{for} \quad r \geq w. \end{cases}$$

Proof: By Corollary 5.5.3, n_r and m_r have the same parity as n_1 and m_1 for all $r \in \mathbb{N}$. It follows by Theorem 5.6.3 that if both n_1 and m_1 are odd then $\tau_r = 2n_r$ for all $r \in \mathbb{N}$, and otherwise $\tau_r = n_r$ for all $r \in \mathbb{N}$. The result now follows from Theorem 5.5.2. □

In other words, for $r \geq w$ the reduced sequence $(h_n \bmod p^r)$ repeats after the same number of zeroes no matter what r is.

Since τ_1 divides $p-1$ by Theorem 5.6.1, we have

Corollary 5.6.6. *If p is a regular odd prime with $N_1 \geq 4$, then*

$$\tau_r \mid p^{w-u}(p-1).$$

Finally we prove that as r increases from 1 the period of $(h_n \bmod p^r)$ remains the same until r reaches some value u, and then increases by a factor of p each time. This confirms Shipsey's Conjecture 4.7.11 for the $r = 2$, $w = 1$ case.

Theorem 5.6.7. *Let (h_n) be an elliptic divisibility sequence, let p be a regular odd prime with gap $N_1 \geq 4$ in (h_n), and let p^w be the highest power of p dividing N_1. For $r \in \mathbb{N}$ let π_r be the period of $(h_n \bmod p^r)$.*

Then there exists an integer $u \leq w$ such that

$$\pi_r = \begin{cases} \pi_1 & \text{for} \quad r \leq u, \text{ and} \\ p^{r-u}\,\pi_1 & \text{for} \quad r \geq u. \end{cases}$$

Proof: Since $N_r = N_1$ for $r \leq w$, $N_r = p^{r-w}N_1$ for $r \geq w$ and $\pi_r = N_r\,\tau_r$ for all $r \in \mathbb{N}$, the result follows from Theorem 5.6.5. □

By Theorem 5.5.2, u can be found from n_w, the order of c_w in $\mathbb{Z}^*_{p^w}$, since p^{w-u} is the highest power of p dividing n_w.

We have found many EDSs in which $w > 1$ and $\tau_w > \tau_1$ $(u \neq w)$, and also a few in which $w > 1$ and $\tau_w = \tau_1$ $(u = w)$, as in the following example.

Example: Let (h_n) be the EDS

$$\ldots, -1,\ 0,\ 1,\ 1,\ 3,\ 3,\ -24,\ -99,\ -675,\ -4401,\ 38799\ldots,$$

reduced modulo 5^3 and written in rows of length $N_1 = 7$.

0	1	1	3	3	101	26
75	99	49	97	97	124	24
100	101	76	78	78	76	101
100	24	124	97	97	49	99
75	26	101	3	3	1	1

0	124	124	122	122	24	99
50	26	76	28	28	1	101
25	24	49	47	47	49	24
25	101	1	28	28	76	26
50	99	24	122	122	124	124

0	1	1	3	3	101	26

Here

$$N_1 = N_2 = 7 \quad \text{and} \quad N_3 = (5)(7) = 35$$

(i.e., $w = 2$), and

$$\tau_1 = \tau_2 = \tau_3 = 2.$$

(The sequence $(h_n \bmod 5^r)$ starts repeating after two zeroes for $r = 1, 2, 3$.) So $u = 2 = w$. The period is

$$\pi_1 = \pi_2 = (2)(7) = 14 \quad \text{and} \quad \pi_3 = (2)(35) = 70.$$

If we continued the table we would find $N_4 = (5)(35) = 175$, $\tau_4 = 2$ and $\pi_4 = (2)(175) = 350$.

Remark: David Gale quotes in [13] an experimental result of Raphael Robinson's that says for a particular sequence of rational numbers called Somos(4) the period modulo p^r seems to behave exactly like this. Since Somos(4) and EDSs are both special cases of a class of sequences called Somos 4 sequences, this motivates an attempt to prove a more general result for all Somos 4 sequences. We do this in chapters 8 and 9.

5.7 Irregular primes

In this section we consider the sequence $(h_n \bmod p^r)$ for primes p which divide both h_3 and h_4. By Theorem 4.7.1, p is irregular and divides all terms h_n for $|n| \geq 3$. Let $f(n)$ be the maximum power of p dividing h_n; we are interested in how fast $f(n)$ grows with n. In this section we find a lower bound for $f(n)$ in terms of n.

From the EDS doubling formulae (4.5),

$$h_{2t+1} = h_{t+2}\, h_t^3 - h_{t-1}\, h_{t+1}^3$$
$$h_{2t}\, h_2 = h_t \left(h_{t+2}\, h_{t-1}^2 - h_{t-2}\, h_{t+1}^2 \right)$$

for all $t \in \mathbb{Z}$, so we have

$$f(2t+1) \geq \min\left\{ f(t+2) + 3f(t), f(t-1) + 3f(t+1) \right\} \tag{5.9a}$$

and

$$f(2t) + f(2) \geq f(t) + \min\left\{ f(t+2) + 2f(t-1), f(t-2) + 2f(t+1) \right\} \tag{5.9b}$$

for all $t \in \mathbb{Z}$. So if $f(n)$ were a smooth function (i.e., if $f(t+i)$ were close to $f(t)$ for small i) then we would expect $f(2t)$ and $f(2t+1)$ to be about 4 times $f(t)$ for all $t \in \mathbb{Z}$ — i.e., we would expect $f(n)$ to grow as n^2. This prompts the following proof.

Theorem 5.7.1. *Let (h_n) be an elliptic divisibility sequence, and let p be a prime dividing $\gcd(h_3, h_4)$ but not dividing h_2. For all $n \neq 0$, let $f(n)$ be the maximum power of p dividing h_n. Then for all $|n| \geq 3$,*

$$f(n) \geq \frac{1}{25}\, n^2.$$

Proof: Note that $f(1) = 0$, $f(2) = 0$, $f(3) \geq 1$, $f(4) \geq 1$, and $f(n) = f(-n)$ for all $n \in \mathbb{Z}$. Let a be any positive rational number such that $f(n) \geq an^2$ for $n = 3, 4, \ldots, 9$. So we may suppose that $f(n) \geq an^2$ for $3 \leq n \leq 2t-1$ for some $t \geq 5$, and we prove that then $f(n) \geq an^2$ for $n = 2t$ and $n = 2t+1$ too.

Note that (since $t \geq 4$),

$$f(t+2) + 3f(t) \geq a(t+2)^2 + 3at^2 = 4at^2 + 4at + 4a,$$

and

$$f(t-1) + 3f(t+1) \geq a(t-1)^2 + 3a(t+1)^2 = 4at^2 + 4at + 4a.$$

So by (5.9a),

$$f(2t+1) \geq 4at^2 + 4at + 4a = a(2t+1)^2 + 3a > a(2t+1)^2.$$

Similarly (since $t \geq 5$),

$$f(t+2) + 2f(t-1) \geq a(t+2)^2 + 2a(t-1)^2 = 3at^2 + 6a,$$

and

$$f(t-2) + 2f(t+1) \geq a(t-2)^2 + 2a(t+1)^2 = 3at^2 + 6a,$$

so by (5.9b) (since $f(2) = 0$)

$$f(2t) \geq at^2 + 3at^2 + 6a = a(2t)^2 + 6a > a(2t)^2.$$

It follows by induction that $f(n) \geq an^2$ for all $n \geq 3$, and hence for all $|n| \geq 3$.

It only remains to find an appropriate a, i.e., a positive rational number such that $f(n) \geq an^2$ for $n = 3, 4, \ldots, 9$. Since $f(1) = 0$, $f(2) = 0$, $f(3) \geq 1$ and $f(4) \geq 1$, it is easy to prove by applying the doubling formula repeatedly that

$$f(5) \geq 1, \quad f(6) \geq 2, \quad f(7) \geq 3, \quad f(8) \geq 3, \quad \text{and} \quad f(9) \geq 4.$$

So it is sufficient to find an $a \in \mathbb{Q}$ satisfying

$$9\,a \leq 1, \;\; 16\,a \leq 1, \;\; 25\,a \leq 1, \;\; 36\,a \leq 2, \;\; 49\,a \leq 3, \;\; 64\,a \leq 3, \quad \text{and} \quad 81\,a \leq 4.$$

So we can use $a = \frac{1}{25}$. In other words, we have proved that

$$f(n) \geq \frac{1}{25}\,n^2 \quad \text{for all } |n| \geq 3.$$

So the power of p dividing h_n grows at least quadratically with n. □

Remark: Note that if $p \nmid h_2$ and p^2 does not divide h_3 or h_4, then $h_5 = h_4 h_2^3 - h_3^3$ is divisible by p but not by p^2, i.e., $f(5) = 1$. So $f(5) \geq a(5)^2$ implies $1 \geq 25a$, i.e., $a \leq \frac{1}{25}$. This shows that we cannot find a higher number a than $\frac{1}{25}$ to use in Theorem 5.7.1, although it might be possible to improve the bound by adding a linear and constant term in n.

We have not been able to prove an upper bound on $f(n)$.

Chapter 6

Somos 4 sequences

Elliptic divisibility sequences are a special case of a class of sequences of rational numbers called Somos 4 sequences. In this chapter we define these sequences and describe what is known and what conjectures have been made about their periodicity properties modulo a prime power, and in the next we describe some recent results relating Somos 4 sequences to elliptic curves.

6.1 Somos sequences

Definition: For $k \geq 4$, a *Somos k sequence* is a sequence (h_n) of rational numbers satisfying the recursion

$$h_n\, h_{n-k} = \sum_{i=1}^{\lfloor \frac{k}{2} \rfloor} \lambda_i\, h_{n-i}\, h_{n-k+i} \qquad \text{for } n \in \mathbb{Z}, \tag{6.1}$$

where the *coefficients* $\lambda_1, \lambda_2, \ldots, \lambda_{\lfloor \frac{k}{2} \rfloor}$ and the *initial values* $h_0, h_1, \ldots, h_{k-1}$ are given rational numbers. We denote by Somos(k) the particular Somos k sequence whose coefficients and initial values are all 1.

Since each term h_n of a Somos k sequence is defined in terms of the preceding k terms we must choose the coefficients and k initial values $h_0, h_1, \ldots, h_{k-1}$. We can then extend the sequence to the right using

$$h_n = \frac{\sum_{i=1}^{\lfloor \frac{k}{2} \rfloor} \lambda_i\, h_{n-i}\, h_{n-k+i}}{h_{n-k}} \qquad \text{for } n \geq k$$

(as long as $h_{n-k} \neq 0$) and to the left using

$$h_n = \frac{\sum_{i=1}^{\lfloor \frac{k}{2} \rfloor} \lambda_i \, h_{n+k-i} \, h_{n+i}}{h_{n+k}} \qquad \text{for } n < 0$$

(as long as $h_{n+k} \neq 0$). If some of the coefficients and initial values are negative, then it may happen that some term $h_n = 0$. In this case if $n \geq k$ then the term h_{n+k} is not defined by the recursion, so we shall make the convention that the sequence terminates to the right at h_{n+k-1}. Similarly, if $h_n = 0$ for $n < 0$ then the sequence terminates to the left at h_{n-k+1}. We will sometimes find it more convenient to make (h_n) terminate at the zero term h_n instead of $k-1$ terms later. We call both types of sequences Somos k sequences, and will always make it clear which type we are speaking of. We denote the set of indices over which h_n is defined (i.e., before the sequence terminates) by I. Of course if $h_0, h_1, \ldots, h_{k-1}$ and the λ_i are all positive then the sequence has no zero terms and never terminates; then $I = \mathbb{Z}$. Somos 4 sequences containing a zero term are considered in section 6.4.

If we decrease each index of a Somos k sequence (h_n) by t (which is equivalent to "starting" the sequence at h_t instead of h_0), we get a new sequence which also satisfies the Somos k recursion (6.1) with the same coefficients:

Definition: For an integer t, the *t–translate* of a Somos k sequence (h_n) is the sequence (ℓ_n) defined by

$$\ell_n = h_{n+t} \quad \text{for all } n + t \in I.$$

The set of indices over which (ℓ_n) is defined is denoted by I_t; so $n \in I_t$ if and only if $n + t \in I$.

If k is even (say $k = 2q$) then setting $m = n - q$ we can rewrite the Somos k recursion (6.1) as

$$h_{m+q} \, h_{m-q} = \sum_{j=0}^{q-1} \lambda_{q-j} \, h_{m+j} \, h_{m-j} \qquad \text{for } m \in \mathbb{Z}.$$

If (h_n) is a generalised elliptic sequence then, by Theorem 4.1.2, for any $t \in \{1, 2 \ldots, q-1\}$ (h_n) satisfies

$$h_{m+q}\, h_{m-q} = \left(\frac{h_q^2}{h_t^2}\right) h_{m+t}\, h_{m-t} - \left(\frac{h_{q+t}\, h_{q-t}}{h_t^2}\right) h_m^2 \quad \text{for all } m \in \mathbb{Z}.$$

So (h_n) satisfies $q-1$ different Somos $2q$ recursions (for $t = 1, 2, \ldots, q-1$), in each of which $\lambda_{q-t} = \frac{h_q^2}{h_t^2}$, $\lambda_q = -\frac{h_{q+t}\, h_{q-t}}{h_t^2}$ and the other coefficients are zero. (Of course, since $h_0 = 0$, if we are looking at (h_n) as a Somos $2q$ sequence then we have to terminate it to the left at h_{-2q+1}.)

We are interested in particular in Somos 4 sequences, that is, in sequences of rational numbers satisfying the recursion

$$h_{m+2}\, h_{m-2} = \lambda_1\, h_{m+1}\, h_{m-1} + \lambda_2\, h_m^2 \qquad \text{for all } m \in \mathbb{Z} \tag{6.2}$$

for some rational constants λ_1, λ_2. Generalised elliptic sequences are a special case of Somos 4 sequences in which the coefficients are $\lambda_1 = \left(\frac{h_2}{h_1}\right)^2$ and $\lambda_2 = -\frac{h_3}{h_1}$.

Examples:

1. Somos(4) is

   $\ldots,\ 314,\ 59,\ 23,\ 7,\ 3,\ 2,\ 1,\ 1,\ 1,\ 1,\ 2,\ 3,\ 7,\ 23,\ 59,\ 314,\ 1529,\ 8209,\ \ldots$

   Notice that $h_{-n} = h_{3+n}$ for all $n \in \mathbb{Z}$.

2. The Somos 4 sequence with initial values $2, 3, 5, 7$ and coefficients $\lambda_1 = 1$ and $\lambda_2 = 2$ is

   $$\ldots, \frac{1184}{9}, \frac{128}{3}, \frac{40}{3}, 4, 4, 2, 3, 5, 7, \frac{71}{2}, \frac{551}{6}, \frac{1898}{3}, \frac{101125}{18}, \ldots.$$

   Note that this sequence and its translates do not satisfy the elliptic sequence equation (4.1), since they contain no zero term.

3. All rational sequences generated by a linear recursion of order 2 are Somos 4 sequences. For if (h_n) is a sequence generated by

   $$h_{n+1} = A\, h_n + B\, h_{n-1} \tag{6.3}$$

for some rational constants A and B, then for all $n \in \mathbb{Z}$,

$$\begin{aligned} h_{n+2}\,h_{n-2} &= \left(A\,h_{n+1} + B\,h_n\right)\left(\frac{h_n - A\,h_{n-1}}{B}\right) \\ &= (h_{n+1} - B\,h_{n-1})\,\frac{A}{B}\,h_n - \frac{A^2}{B}\,h_{n+1}\,h_{n-1} + h_n^2 \\ &= (A\,h_n)\,\frac{A}{B}\,h_n - \frac{A^2}{B}\,h_{n+1}\,h_{n-1} + h_n^2 \\ &= -\frac{A^2}{B}\,h_{n+1}\,h_{n-1} + \left(\frac{A^2}{B} + 1\right)\,h_n^2. \end{aligned}$$

So (h_n) is a Somos 4 sequence with $\lambda_1 = -\frac{A^2}{B}$ and $\lambda_2 = -\lambda_1 + 1$. For example, the Fibonacci sequence $F_{n+1} = F_n + F_{n-1}$ is a Somos 4 sequence with coefficients $\lambda_1 = -1$ and $\lambda_2 = 2$, and the Mersenne sequence $M_n = 2^n - 1 = 3\,M_{n-1} - 2\,M_{n-2}$ is a Somos 4 sequence with coefficients $\lambda_1 = \frac{9}{2}$ and $\lambda_2 = -\frac{7}{2}$.

Conversely, a Somos 4 sequence (h_n) satisfies an order 2 linear recurrence $h_{n+1} = A\,h_n + B\,h_{n-1}$ for some $A, B \in \mathbb{Q}$ if and only if the coefficients of (h_n) are given by

$$\lambda_1 = \frac{(h_0\,h_3 - h_1\,h_2)^2}{(h_1\,h_3 - h_2^2)(h_0\,h_2 - h_1^2)} \quad \text{and} \quad \lambda_2 = 1 - \lambda_1.$$

Then A and B are given by

$$A = \frac{h_0\,h_3 - h_1\,h_2}{h_0\,h_2 - h_1^2} \quad \text{and} \quad B = \frac{h_2^2 - h_1\,h_3}{h_0\,h_2 - h_1^2},$$

Remarks:

1. In David Gale's definition of Somos sequences in [12] the coefficients λ_i are integers. In Raphael Robinson's definition of "generalised Somos sequences" in [21] he assumes both the coefficients and the initial values to be positive integers, so that the recursion never produces a zero term. We do not make the same assumption, because we want elliptic sequences to be a special case, and for these $\lambda_2 = -h_1\,h_3$ and $h_0 = 0$.

2. If one of the initial values h_n is zero, then the sequence terminates to the left at h_{n-k+1} and to the right at h_{n+k-1}.

3. Aside from the above termination convention, in a Somos sequence (h_n) there is nothing to distinguish the "start" of the sequence h_0 — for any integer r (as long as $h_r, \ldots, h_{r+k-1}$ are defined and non-zero) the r–translate of (h_n) is also a Somos 4 sequence with the same coefficients. This is in contrast to elliptic sequences, whose translates are usually not elliptic sequences. The difference is that (6.1) is a "local" recursion, relating h_n to the k previous terms, as opposed to a "global" recursion like (4.1), which relates h_n to every other term in the sequence.

4. There are various alternative ways of handling the situation where some term $h_r = 0$ in a Somos k sequence (h_n). Gale [13] makes the convention that the sequence terminates to the right at h_r (if $r > 0$), and we allow it to terminate at either h_r or h_{r+k-1}. Robinson [21] ensures the recursion never produces a zero term by looking only at sequences with positive coefficients and initial values.

   Another option is simply to accept the non-uniqueness; i.e., to allow any choice for h_{r+k}. Then all Somos k sequences would be infinite, and for any rational number c there would be a Somos k sequence (h'_n) with the same initial values as (h_n) and $h'_{r+k} = c$. However, this option is not really very different from the option of terminating the sequence at h_{r+k-1}, in that any infinite sequence of rational numbers generated by the Somos k recursion in this manner can be constructed by "sticking together" a number of overlapping Somos k sequences, all with the same coefficients and each terminating $k-1$ terms after the zero term. Since we are mostly interested in results on periodicity and symmetry, which of course do not hold throughout such an infinite sequence if it contains zeroes and hence arbitrary terms, we prefer the option of terminating the sequence.

5. If $\ldots, h_{-2}, h_{-1}, h_0, h_1, h_2, \ldots$ is a Somos k sequence with coefficients λ_i, then the reversed sequence $\ldots, h_2, h_1, h_0, h_{-1}, h_{-2}, \ldots$ is also a Somos k sequence with the same coefficients. To see this, write $\ell_n = h_{-n}$ for all $n \in \mathbb{Z}$ and

note that

$$\ell_{-n}\,\ell_{-n+k} = \sum_{i=1}^{\lfloor \frac{k}{2} \rfloor} \lambda_i\,\ell_{-n+i}\,\ell_{-n+k-i} \qquad \text{for } n \in \mathbb{Z},$$

or (replacing $-n+k$ by m),

$$\ell_{m-k}\,\ell_{m} = \sum_{i=1}^{\lfloor \frac{k}{2} \rfloor} \lambda_i\,\ell_{m-k+i}\,\ell_{m-i} \qquad \text{for } m \in \mathbb{Z}.$$

6.2 Basic properties

The coefficients of a Somos 4 sequence are related to any six consecutive terms as follows:

Theorem 6.2.1. *Let (h_n) be a Somos 4 sequence with coefficients λ_1, λ_2. Then for any index k with $k-2, k-1, \ldots, k+3 \in I$,*

$$\lambda_1 \left(h_{k-1}\,h_{k+1}^3 - h_k^3\,h_{k+2} \right) = h_{k-2}\,h_{k+1}^2\,h_{k+2} - h_{k-1}\,h_k^2\,h_{k+3}$$

and

$$\lambda_2 \left(h_{k-1}\,h_{k+1}^3 - h_k^3\,h_{k+2} \right) = h_{k-1}^2\,h_{k+1}\,h_{k+3} - h_k\,h_{k+2}^2\,h_{k-2}.$$

Proof: Using the Somos 4 recursion we have

$$h_{k+2}\,h_{k-2} = \lambda_1\,h_{k+1}\,h_{k-1} + \lambda_2\,h_k^2$$

and

$$h_{k+3}\,h_{k-1} = \lambda_1\,h_{k+2}\,h_k + \lambda_2\,h_{k+1}^2.$$

Multiplying the first equation by h_{k+1}^2 and the second by h_k^2 and subtracting, we get

$$h_{k+1}^2\,h_{k+2}\,h_{k-2} - h_k^2\,h_{k+3}\,h_{k-1} = \lambda_1 \left(h_{k+1}^3\,h_{k-1} - h_k^3\,h_{k+2} \right).$$

Similarly,

$$h_k\,h_{k+2}^2\,h_{k-2} - h_{k+1}\,h_{k-1}^2\,h_{k+3} = \lambda_2 \left(h_k^3\,h_{k+2} - h_{k+1}^3\,h_{k-1} \right).$$

□

We will find it useful to work out the next four values $h_5, \ldots, h_8$ of (h_n) in terms of the initial values h_0, h_1, h_2, h_3, and h_4:

Theorem 6.2.2. *Let (h_n) be a Somos 4 sequence with coefficients λ_1, λ_2 and non-zero initial values $h_0, \ldots, h_3$. Then*

$$h_5 = \frac{1}{h_1}\Big(\lambda_1\, h_2\, h_4 + \lambda_2\, h_3^2\Big),$$

$$h_6 = \frac{1}{h_1\, h_2}\Big(\lambda_1\, \lambda_2\, h_3^3 + \lambda_1^2\, h_2\, h_3\, h_4 + \lambda_2\, h_1\, h_4^2\Big),$$

$$h_7 = \frac{1}{h_1^2\, h_2\, h_3}\Big(\lambda_2^3\, h_2\, h_3^4 + \big(\lambda_1\, \lambda_2^2\, h_2^2\, h_3^2\big)\ h_4 + \big(\lambda_1\, \lambda_2\, h_0\, h_3^2\big)\ h_4^2 + \lambda_1\ \big(\lambda_1\, h_0\, h_2 + \lambda_2\, h_1^2\big)\ h_4^3\Big),$$

and

$$h_8 = \frac{1}{h_1^3\, h_2^2\, h_3}\Big(\lambda_1\, \lambda_2^3\, h_0\, h_3^6 + \lambda_1\, \lambda_2^2\, h_3^4\ \big(3\lambda_1\ h_0\, h_2 + 2\lambda_2\, h_1^2\big)\ h_4 + 3\lambda_1^2\, \lambda_2\, h_2\, h_3^2\, \big(\lambda_1\, h_0\, h_2 + \lambda_2\, h_1^2\big)\ h_4^2 + \big(\lambda_1^3\, h_2^2\, (\lambda_1\, h_2\, h_0\ + \lambda_2\, h_1^2) + \lambda_2^3\, h_1^3\, h_3\big)\ h_4^3\Big).$$

Proof: The proof is by direct calculation, using the Somos 4 recursion (6.2) to replace $\lambda_1\, h_1\, h_3 + \lambda_2\, h_2^2$ by $h_0\, h_4$ whenever it occurs. □

6.3 Equivalent Somos 4 sequences

We next define what we mean by equivalence in Somos 4 sequences.

Theorem 6.3.1. *Let (h_n) be a Somos 4 sequence with coefficients λ_1, λ_2, and let a, b and c be rational numbers such that $an^2 + bn + c$ is an integer for all $n \in \mathbb{Z}$ (in other words, such that $2a$, $2b$ and c are integers and $2a$ is odd if and only if $2b$ is odd). Then for any rational number θ, the sequence given by*

$$h'_n = \theta^{a\,n^2+b\,n+c}\, h_n \qquad \textit{for all } n \in I$$

is a Somos 4 sequence with coefficients

$$\lambda'_1 = \theta^{6a}\, \lambda_1 \qquad \textit{and} \qquad \lambda'_2 = \theta^{8a}\, \lambda_2.$$

Proof: Since $an^2 + bn + c$ is an integer for all $n \in \mathbb{Z}$, each term h'_n is a rational number. We have for all $n \in \mathbb{Z}$,

$$h_n\, h_{n-4} = \lambda_1\, h_{n-3}\, h_{n-1} + \lambda_2\, h_{n-2}^2.$$

Hence

$$\begin{aligned} h'_n\, h'_{n-4} &= \left(\theta^{an^2+bn+c}\, h_n\right)\, \left(\theta^{a(n-4)^2+b(n-4)+c}\, h_{n-4}\right) \\ &= \theta^{(2a)n^2+(2b-8a)n+(16a-4b+2c)}\, h_n\, h_{n-4}, \end{aligned}$$

and

$$\begin{aligned} &\lambda'_1\, h'_{n-1}\, h'_{n-3} + \lambda'_2\, {h'_{n-2}}^2 \\ &\quad = \left(\lambda_1\, \theta^{6a}\right)\, \left(\theta^{a(n-1)^2+b(n-1)+c}\, h_{n-1}\right)\, \left(\theta^{a(n-3)^2+b(n-3)+c}\, h_{n-3}\right) \\ &\qquad\qquad + \left(\lambda_2\, \theta^{8a}\right)\, \left(\theta^{a(n-2)^2+b(n-2)+c}\, h_{n-2}\right)^2 \\ &\quad = \theta^{6a+a(n^2-2n+1+n^2-6n+9)+b(n-1+n-3)+2c}\, \lambda_1\, h_{n-1}\, h_{n-3} \\ &\qquad\qquad + \theta^{8a+2(a(n^2-4n+4)+b(n-2)+c)}\, \lambda_2\, h_{n-2}^2 \\ &\quad = \theta^{(2a)n^2+(2b-8a)n+(16a-4b+2c)}\, \left(\lambda_1 h_{n-1} h_{n-3} + \lambda_2 h_{n-2}^2\right) \\ &\quad = h'_n\, h'_{n-4}. \end{aligned}$$

Hence (h'_n) is a Somos 4 sequence with coefficients $\lambda'_1 = \theta^{6a}\, \lambda_1$ and $\lambda'_2 = \theta^{8a}\, \lambda_2$. □

This leads to the following concept of equivalence:

Definition: Two Somos 4 sequences (h_n) and (h'_n) are said to be *equivalent* if, for some $k \in \mathbb{N}$, there exist rational numbers a_i, b_i, c_i, θ_i for $i = 1, \ldots, k$ such that $2a_i$, $2b_i$ and c_i are integers, $2a_i$ and $2b_i$ have the same parity, $\theta_i \neq 0$, and

$$h'_n = \theta_1^{a_1 n^2+b_1 n+c_1} \cdot \theta_2^{a_2 n^2+b_2 n+c_2} \cdots \theta_k^{a_k n^2+b_k n+c_k}\, h_n \quad \text{for all } n \in I.$$

This is obviously an equivalence relation in the technical sense. By Theorem 6.3.1, (h'_n) has coefficients

$$\lambda'_1 = \theta_1^{6a_1}\theta_2^{6a_2}\cdots\theta_k^{6a_k}\lambda_1 \qquad \text{and} \qquad \lambda'_2 = \theta_1^{8a_1}\theta_2^{8a_2}\cdots\theta_k^{8a_k}\lambda_2.$$

We can express this equivalence condition in another way as follows:

Theorem 6.3.2. *Let (h_n) and (h'_n) be Somos 4 sequences. Then (h_n) and (h'_n) are equivalent if and only if there exist $\alpha, \beta, \gamma \in \mathbb{Q}(\sqrt{D})$, where $D \in \mathbb{Q}$, such that α^2, β^2, $\alpha\,\beta$ and γ are non-zero rational numbers and*

$$h'_n = \alpha^{n^2}\,\beta^n\,\gamma\,h_n \qquad \text{for all } n \in I.$$

The coefficients of (h'_n) are $\lambda'_1 = \alpha^6\,\lambda_1$ and $\lambda'_2 = \alpha^8\,\lambda_2$.

Proof: Let (h'_n) be equivalent to (h_n). Then there exist rational numbers a_i, b_i, c_i, θ_i for $i = 1, \ldots, k$ such that $2a_i$, $2b_i$ and c_i are integers, $2a_i$ and $2b_i$ have the same parity, $\theta_i \neq 0$, and for all $n \in I$,

$$\begin{aligned} h'_n &= \theta_1^{a_1 n^2 + b_1 n + c_1} \cdot \theta_2^{a_2 n^2 + b_2 n + c_2} \cdots \theta_k^{a_k n^2 + b_k n + c_k}\, h_n \\ &= \left(\theta_1^{a_1}\,\theta_2^{a_2}\,\cdots\theta_k^{a_k}\right)^{n^2} \left(\theta_1^{b_1}\,\theta_2^{b_2}\,\cdots\theta_k^{b_k}\right)^{n} \left(\theta_1^{c_1}\,\theta_2^{c_2}\,\cdots\theta_k^{c_k}\right) h_n. \end{aligned}$$

Let

$$\alpha = \theta_1^{a_1}\,\theta_2^{a_2}\,\cdots\theta_k^{a_k},\ \beta = \theta_1^{b_1}\,\theta_2^{b_2}\,\cdots\theta_k^{b_k} \quad \text{and } \gamma = \theta_1^{c_1}\,\theta_2^{c_2}\,\cdots\theta_k^{c_k}.$$

Then α and β are either both rational numbers or both non-rational square roots of rational numbers, and we have

$$h'_n = \alpha^{n^2}\,\beta^n\,\gamma\,h_n.$$

The converse follows by writing $h'_n = (\alpha^2)^{\frac{1}{2}n(n-1)}\,(\alpha\,\beta)^n\,\gamma$ and using the definition of equivalence. □

As α^2, β^2 and γ vary over all non-zero rational numbers with $\alpha\,\beta \in \mathbb{Q}$, the sequence (h'_n) varies over all Somos 4 sequences equivalent to (h_n).

It is easy to prove the following:

Theorem 6.3.3. *Let (h_n) and (h'_n) be Somos 4 sequences, and let (ℓ_n) and (ℓ_n) be t–translates of (h_n) and (h'_n) respectively. Then (ℓ_n) and (ℓ_n) are equivalent if and only if (h_n) and (h'_n) are equivalent.*

We will need the following result in chapter 7:

Theorem 6.3.4. *Let (h_n) be a Somos 4 sequence. Then (h_n) is equivalent to the constant sequence $\ldots, 1, 1, 1, \ldots$ if and only if*

$$\frac{h_{-2}\, h_0}{h_{-1}^2} = \frac{h_{-1}\, h_1}{h_0^2} = \frac{h_0\, h_2}{h_1^2}. \tag{6.4}$$

Proof: By Theorem 6.3.2 (h_n) is equivalent to the constant sequence $\ldots, 1, 1, 1, \ldots$ if and only if $h_n = \alpha^{n^2}\, \beta^n\, \gamma$ for $n \in \mathbb{Z}$, where α^2, β^2, $\alpha\,\beta$ and γ are non-zero rational numbers. So if (h_n) is equivalent to $\ldots, 1, 1, 1, \ldots$ then

$$\frac{h_{n+1}\, h_{n-1}}{h_n^2} = \frac{\alpha^{(n+1)^2}\, \beta^{n+1}\, \gamma \cdot \alpha^{(n-1)^2}\, \beta^{n-1}\, \gamma}{\left(\alpha^{n^2}\, \beta^n\, \gamma\right)^2} = \alpha^2$$

for all $n \in \mathbb{Z}$, and in particular (h_n) satisfies (6.4).

For the converse, suppose (h_n) satisfies (6.4), and choose $\alpha, \beta, \gamma \in \mathbb{Q}$ such that

$$\gamma = h_0, \quad \alpha\,\beta = \frac{h_1}{h_0} \quad \text{and} \quad \alpha\,\beta^{-1} = \frac{h_{-1}}{h_0}.$$

We prove by induction that then $h_n = \alpha^{n^2}\, \beta^n\, \gamma$ for all $n \in \mathbb{Z}$. This holds for $n = -1, 0, 1$ by our choice of α, β and γ, and it follows from (6.4) that it holds for $n = -2, 2$. Substituting into $h_2\, h_{-2} = \lambda_1\, h_1\, h_{-1} + \lambda_2\, {h_0}^2$ and simplifying yields

$$\alpha^8 = \alpha^2\, \lambda_1 + \lambda_2.$$

We now use the Somos 4 equation $h_n\, h_{n-4} = \lambda_1\, h_{n-1}\, h_{n-3} + \lambda_2\, {h_{n-2}}^2$ and induction on n to prove that

$$h_n = \alpha^{n^2-8}\, \beta^n\, \gamma\, \left(\alpha^2\, \lambda_1 + \lambda_2\right) = \alpha^{n^2}\, \beta^n\, \gamma$$

for all $n \in \mathbb{Z}$. □

Remarks:

1. Clearly any two equivalent Somos 4 sequences are defined over the same set of indices I (i.e., they have zeroes in the same places).

2. This notion of equivalence is more general than the one we described for elliptic sequences. In fact it is easy to show that if (h_n) satisfies the elliptic

sequence equation (4.1) and (h'_n) is an equivalent Somos 4 sequence given by

$$h'_n = \theta^{an^2+bn+c}\, h_n \quad \text{for all } n \in I,$$

where $\theta \neq \pm 1$, then (h'_n) satisfies (4.1) if and only if $b = 0$ and $c = -a$ (since this is the only choice which makes $h'_1 = \pm 1$ and $h'_{-1} = \mp 1$); then $h'_n = (\theta^a)^{n^2-1}\, h_n$ for all $n \in I$. Otherwise, if $\theta = \pm 1$ then (h'_n) is either $(-h_n)$ or $(\theta^{n^2-1}\, h_n)$. So if two elliptic sequences are equivalent as Somos 4 sequences, then either they are also equivalent as elliptic sequences or one can be obtained from the other by multiplying each term by (-1).

Similarly, if (h_n) satisfies the generalised EDS equation (4.2a) and $\theta \neq \pm 1$ then the equivalent Somos 4 sequence (h'_n) satisfies (4.2a) if and only if $b = 0$.

3. An equivalence

$$h'_n = \theta^{f(n)}\, h_n \quad \text{for all } n \in I,$$

where f is a polynomial of degree greater than 2 such that $f(n)$ is an integer for all $n \in \mathbb{Z}$, does not necessarily produce another Somos 4 sequence. For example, if $f(n) = dn^3 + an^2 + bn + c$ then we find that (h'_n) satisfies the recursion

$$h'_{m+2}\, h'_{m-2} = \lambda'_1\, h'_{m+1}\, h'_{m-1} + \lambda'_2\, {h'_m}^2 \qquad \text{for all } m \in I,$$

but with $\lambda'_1 = \theta^{6a+(18n-36)d}\lambda_1$ and $\lambda'_2 = \theta^{8a+(24n-48)d}\lambda_2$, which depend on n. So (h'_n) is not a Somos 4 sequence unless $d = 0$.

6.4 Somos 4 sequences containing a zero term

In this section we show that Somos 4 sequences containing a zero term are easy to describe in terms of elliptic sequences.

Let (h_n) be a Somos 4 sequence with non-zero coefficients, in which $h_r = 0$ for some index r. If $r \in \{0, 1, 2, 3\}$ then the sequence terminates to the left at h_{r-3} and to the right at h_{r+3}, which is not very interesting, so we assume that the initial values of (h_n) are non-zero.

We first find simple expressions for the coefficients λ_i in terms of the five terms around h_r:

Theorem 6.4.1. *Let (h_n) be a Somos 4 sequence with non-zero coefficients and initial values, in which $h_r = 0$ for some index $r \notin \{0, 1, 2, 3\}$. Then (h_n) has coefficients*

$$\lambda_1 = -\frac{h_{r+2}^2\, h_{r-1}}{h_{r+1}^3} \qquad \textit{and} \qquad \lambda_2 = \frac{h_{r-1}\, h_{r+3}}{h_{r+1}^2}.$$

Proof: Since $h_r = 0$, we have

$$h_{r+2}\, h_{r-2} = \lambda_1\, h_{r+1}\, h_{r-1} + \lambda_2\, h_r^2 = \lambda_1\, h_{r+1}\, h_{r-1}$$

and

$$h_{r-1}\, h_{r+3} = \lambda_1\, h_r\, h_{r+2} + \lambda_2\, h_{r+1}^2 = \lambda_2\, h_{r+1}^2,$$

and it follows that

$$\lambda_1 = -\frac{h_{r+2}\, h_{r-2}}{h_{r-1}\, h_{r+1}} \quad \text{and} \quad \lambda_2 = \frac{h_{r-1}\, h_{r+3}}{h_{r+1}^2}. \tag{6.5}$$

If $r < 0$, then since $h_{r-1}, \ldots, h_{r+4} \in I$ and $h_{r+1}\, h_{r+3} \neq 0$ it follows from Theorem 6.2.1 with $k = r + 1$ that

$$\lambda_1 = -\frac{h_{r+2}^2\, h_{r-1}}{h_{r+1}^3},$$

and we are done. Otherwise, if $r > 0$ then since $h_{r-4}, \ldots, h_{r+1} \in I$ and $h_{r-1}\, h_{r-3} \neq 0$ it follows from Theorem 6.2.1 with $k = r - 2$ that

$$\lambda_1 = -\frac{h_{r-2}^2\, h_{r+1}}{h_{r-1}^3}.$$

But since by (6.5) we also have $\lambda_1 = -\frac{h_{r+2}\, h_{r-2}}{h_{r-1}\, h_{r+1}}$, it follows that

$$h_{r-2} = -\left(-\frac{h_{r-1}}{h_{r+1}}\right)^2 h_{r+2},$$

and hence that

$$\lambda_1 = -\frac{h_{r+2}^2\, h_{r-1}}{h_{r+1}^3}.$$

□

It follows that there is a relationship between h_{r+j} and h_{r-j} for $|j| \leq 3$. (Of course, for $|j| \geq 4$, h_{r+j} and h_{r-j} are not both defined.)

Theorem 6.4.2. *Let (h_n) be a Somos 4 sequence with non-zero coefficients and initial values, in which $h_r = 0$ for some index $r \notin \{0, 1, 2, 3\}$. Then for $j = 1, 2, 3$,*

$$h_{r-j} = -\left(-\frac{h_{r-1}}{h_{r+1}}\right)^j h_{r+j}.$$

Proof: For $j = 1$, the result holds trivially, and for $j = 2$ and $j = 3$ it follows by substituting the expressions for λ_1 and λ_2 from Theorem 6.4.1 into

$$h_{r+2}\, h_{r-2} = \lambda_1\, h_{r+1}\, h_{r-1} + \lambda_2\, h_r^2 = \lambda_1\, h_{r+1}\, h_{r-1}$$

and

$$h_{r+1}\, h_{r-3} = \lambda_1\, h_r\, h_{r-2} + \lambda_2\, h_{r-1}^2 = \lambda_2\, h_{r-1}^2,$$

and simplifying (noting that $h_{r+1}\, h_{r+2} \neq 0$). □

Remark: If $h_r = 0$ were in the initial values (i.e., $0 \leq r \leq 3$) then the sequence would terminate to the left at h_{r-3} and to the right at h_{r+3}, the coefficients would be independent of the values of the terms around h_r, and h_{r+j} would depend on the coefficients.

If $h_r = 0$, $h_{r-1} = -1$ and $h_{r+1} = 1$ then the r–translate of (h_n) satisfies the elliptic sequence equation (4.1):

Lemma 6.4.3. *Let (h_n) be a Somos 4 sequence with non-zero coefficients and initial values, in which for some index $r \notin \{0, 1, 2, 3\}$*

$$h_{r-1} = -1, \;\; h_r = 0 \quad \textit{and} \;\; h_{r+1} = 1,$$

and let (Z_n) be the r–translate of (h_n), i.e.,

$$Z_n = h_{r+n} \qquad \textit{for all } r + n \in I.$$

Then (Z_n) is a segment of an elliptic sequence.

Proof: Note that the r–translate (Z_n) satisfies a Somos 4 recursion with the same coefficients λ_1, λ_2 as (h_n). By Theorem 6.4.1,

$$\lambda_1 = -\frac{h_{r+2}^2\, h_{r-1}}{h_{r+1}^3} = h_{r+2}^2 = Z_2^2 \quad \text{and} \quad \lambda_2 = \frac{h_{r-1}h_{r+3}}{h_{r+1}^2} = -h_{r+1}\, h_{r+3} = -Z_1\, Z_3.$$

So (Z_n) satisfies

$$Z_{m+2}\, Z_{m-2} = Z_{m+1}\, Z_{m-1}\, Z_2^2 - Z_1\, Z_3\, Z_m^2$$

whenever $m-2, \ldots, m+2 \in I_r$.

Since $Z_2\, Z_3 \neq 0$, it follows from Theorem 4.4.1 on the existence and uniqueness of elliptic sequences with given initial values that (Z_n) satisfies (4.1). □

This leads to our main result in this section: every Somos 4 sequence containing a zero term has an equivalent sequence, a translate of which satisfies the elliptic sequence equation.

Theorem 6.4.4. *Let (h_n) be a Somos 4 sequence with non-zero coefficients and initial values, in which $h_r = 0$ for some index $r \notin \{0,1,2,3\}$, and let (Z_n) be an r–translate of an equivalent Somos 4 sequence given by*

$$Z_n = \frac{1}{h_{r+1}} \left(-\frac{h_{r+1}}{h_{r-1}}\right)^{\frac{1}{2}n(n-1)} h_{r+n} \quad \textit{for all } n \in I_r.$$

Then (Z_n) is a segment of an elliptic sequence.

Proof: Let (ℓ_n) be the equivalent Somos 4 sequence defined by

$$\ell_n = \frac{1}{h_{r+1}} \left(-\frac{h_{r+1}}{h_{r-1}}\right)^{\frac{1}{2}(n-r)(n-r-1)} h_n = Z_{n-r} \quad \text{for all } n \in I.$$

Then

$$\ell_r = 0, \;\; \ell_{r-1} = -1 \quad \text{and} \;\; \ell_{r+1} = 1,$$

and (Z_n) is the r–translate of (ℓ_n). Hence by Lemma 6.4.3, (Z_n) satisfies the elliptic sequence equation (4.1). □

If the first coefficient λ_1 is a square, then (h_n) has two additional equivalent sequences, r–translates of which also satisfy the elliptic sequence equation:

Theorem 6.4.5. *Let (h_n) be a Somos 4 sequence with non-zero coefficients and initial values, in which $h_r = 0$ for some index $r \notin \{0,1,2,3\}$, and the first coefficient λ_1 is a square. Then $\left(-\frac{h_{r-1}}{h_{r+1}}\right)$ is a square, and if (ℓ_n) is either of the equivalent sequences given by*

$$\ell_n = \beta^n \gamma\, h_n \quad \text{for all } n \in I,$$

where

$$\beta = \pm\sqrt{-\frac{h_{r-1}}{h_{r+1}}}, \quad \text{and} \quad \gamma = \frac{1}{h_{r+1}\,\beta^{r+1}},$$

then the r–translate of (ℓ_n) satisfies (4.1).

Proof: Since λ_1 is a square and $\lambda_1 = -\frac{h_{r+2}^2\, h_{r-1}}{h_{r+1}^3}$ by Theorem 6.4.2, it follows that $\left(-\frac{h_{r-1}}{h_{r+1}}\right)$ is a square, i.e., β is a rational number. So (ℓ_n) is a Somos 4 sequence with $\ell_{r-1} = -1$, $\ell_r = 0$ and $\ell_{r+1} = 1$. It follows by Lemma 6.4.3 that the r–translate of (ℓ_n) satisfies (4.1). □

(We will refer to Theorem 6.4.5 in chapter 7.) Theorem 6.4.4 has several consequences for (h_n). The first is that any Somos 4 sequence (h_n) with non-zero initial values is either infinite in at least one direction, or very short:

Theorem 6.4.6. *Let (h_n) be a Somos 4 sequence with non-zero coefficients and initial values, in which $h_r = 0$ for some index r with $|r|$ minimal. If $h_{r\pm N} = 0$ for some minimal positive integer N, then $4 \le N \le 12$.*

Proof: We know $N \ge 4$ from Theorem 6.4.2 and the fact that $h_{r+1}\, h_{r+2}\, h_{r+3} \neq 0$ if $r < 0$ or $h_{r-1}\, h_{r-2}\, h_{r-3} \neq 0$ if $r > 0$. Now let

$$Z_n = \frac{1}{h_{r+1}} \left(-\frac{h_{r+1}}{h_{r-1}}\right)^{\frac{1}{2}n(n-1)} h_{r+n} \quad \text{for all } n \in I_r.$$

By Theorem 6.4.4, (Z_n) is a segment of an elliptic sequence. Since $h_{r+n} = 0$ for $n \in I$ if and only if $Z_n = 0$, the result follows from Theorem 4.6.3 for elliptic sequences. □

Not surprisingly, the sequence (h_n) satisfies a global recursion "centred" at h_r, not merely the local recursion (6.2). (By a recursion being "global" we mean that it is a relationship between terms which are far apart in the sequence.)

Theorem 6.4.7. *Let (h_n) be a Somos 4 sequence with non-zero coefficients and initial values, in which $h_r = 0$ for some index $r \notin \{0,1,2,3\}$. Then (h_n) satisfies*

$$h_{r+m+n}\, h_{r+m-n}\, h_{r+t}^2 = \left(-\frac{h_{r+1}}{h_{r-1}}\right)^{-n+t} \left(h_{r+m+t}\, h_{r+m-t}\, h_{r+n}^2 - h_{r+n+t}\, h_{r+n-t}\, h_{r+m}^2\right)$$

whenever all these indices are in I.

Proof: Let (Z_n) be the r–translate of an equivalent Somos 4 sequence given by

$$Z_n = \frac{1}{h_{r+1}} \left(-\frac{h_{r+1}}{h_{r-1}}\right)^{\frac{1}{2}n(n-1)} h_{r+n} \quad \text{for all } n \in I_r.$$

Then by Theorems 6.4.4 and 4.1.2, (Z_n) satisfies the generalised EDS equation

$$Z_{m+n}\, Z_{m-n}\, Z_t^2 = Z_{m+t}\, Z_{m-t}\, Z_n^2 - Z_{n+t}\, Z_{n-t}\, Z_m^2$$

whenever all these indices are in I_r.

The result follows by substituting $Z_n = \frac{1}{h_{r+1}} \left(-\frac{h_{r+1}}{h_{r-1}}\right)^{\frac{1}{2}n(n-1)} h_{r+n}$ for all $n \in I_r$ and simplifying. □

Fixing n and t and replacing m by $m - r$, we get

Corollary 6.4.8. *Let (h_n) be a Somos 4 sequence with non-zero coefficients and initial values, in which $h_r = 0$ for some index r, and let $t, n \in \mathbb{N}$ with $t \leq n$ such that $r + n + t \in I$. Define constants $A_{r,n,t}$ and $B_{r,n,t}$ by*

$$A_{r,n,t} := \left(-\frac{h_{r+1}}{h_{r-1}}\right)^{-n+t} \left(\frac{h_{r+n}}{h_{r+t}}\right)^2$$

and

$$B_{r,n,t} := -\left(-\frac{h_{r+1}}{h_{r-1}}\right)^{-n+t} \left(\frac{h_{r+n+t}\, h_{r+n-t}}{h_{r+t}^2}\right).$$

Then (h_n) satisfies the degree $2n$ recursion

$$h_{m+n}\, h_{m-n} = A_{r,n,t}\, h_{m+t}\, h_{m-t} + B_{r,n,t}\, h_m^2,$$

whenever $m + n, m - n \in I$.

In other words, for every $n \geq 2$, the sequence (h_n) also satisfies $n-1$ Somos $2n$ recursions (for $t = 1, 2, \ldots, n-1$), in each of which at most two of the coefficients are non-zero.

Finally, if (h_n) is a Somos 4 sequence containing a zero term then taking every Mth term from any non-zero term h_t, we get (a segment of) another Somos 4 sequence. (We already knew this for elliptic sequences, by Theorem 4.6.1.)

Theorem 6.4.9. *Let (h_n) be a Somos 4 sequence with non-zero coefficients and initial values, in which $h_r = 0$ for some index $r \notin \{0, 1, 2, 3\}$. Let $t \in I$ and let M be a positive integer such that $h_{r+M} \neq 0$. Let (ℓ_n) be the subsequence obtained from (h_n) by taking every Mth term from h_t, i.e.,*

$$\ell_s = h_{t+sM} \qquad \text{whenever } t + sM \in I.$$

Then (ℓ_n) satisfies the Somos 4 recursion with coefficients

$$\left(-\frac{h_{r-1}}{h_{r+1}}\right)^M \left(\frac{h_{r+2M}}{h_{r+M}}\right)^2 \qquad \text{and} \qquad -\left(-\frac{h_{r-1}}{h_{r+1}}\right)^M \left(\frac{h_{r+3M}}{h_{r+M}}\right).$$

Proof: The result follows by setting $m = t - r + sM$, $n = 2M$ and $t = M$ in Theorem 6.4.7 and simplifying. □

6.5 Integrality properties

The surprising thing about Somos sequences is that, even though defined by a rational recursion (i.e., even though the computation of h_n involves dividing by h_{n-k} or h_{n+k}), the sequences Somos(4) to Somos(7) turn out to have only integer terms. This was first pointed out for Somos(6) by Michael Somos [27], though he did not give a proof. A simple proof for Somos(4) and Somos(5) was given by Janice Malouf, and a variant due to George Bergman is quoted by Gale in [13]. (The same method works for all Somos 4 and 5 sequences with integer coefficients, as long as the initial values are all 1.)

Dean Hickerson then showed that the original sequence Somos(6) is an integer sequence, and generalised the problem to include all Somos 6 sequences with

coefficients $\lambda_1 = \lambda_2 = 1$ (see [21]). He proved that if the initial values are not all 1 then, though the recursion does not necessarily give integers, it does give rational numbers whose denominators are products of powers of the numerators and denominators of the initial values. (Richard Stanley also solved this problem using similar methods.) Finally, Benjamin Lotto used Hickerson's method to prove the same thing for Somos(7).

Somos(8) and Somos(9) are not integer sequences (for example, the 17th term of Somos(8) is a fraction).

6.5.1 Writing each term as a rational function of the initial values and coefficients

Let (h_n) be a Somos k sequence with coefficients λ_i and initial values $h_0, \ldots, h_{k-1}$, defined over a set of indices I. Note that each term h_n is given by a rational function of the previous k terms and the coefficients λ_i. So if we replace the initial values $h_0, h_1, \ldots, h_{k-1}$ and coefficients $\lambda_1, \ldots, \lambda_{\lfloor \frac{k}{2} \rfloor}$ by variables $x_0, \ldots, x_{k-1}$, $x_k, \ldots, x_{k-1+\lfloor \frac{k}{2} \rfloor}$, then by applying the Somos k recursion repeatedly and cancelling common factors, we obtain a sequence of rational functions $\frac{f_n}{g_n}$, where f_n and g_n are coprime polynomials in $\mathbb{Z}[x_0, \ldots, x_{k-1+\lfloor \frac{k}{2} \rfloor}]$, which satisfies the recursion

$$\frac{f_n}{g_n} \cdot \frac{f_{n-k}}{g_{n-k}} = \sum_{i=1}^{\lfloor \frac{k}{2} \rfloor} x_{k-1+i} \cdot \frac{f_{n-i}}{g_{n-i}} \cdot \frac{f_{n-k+i}}{g_{n-k+i}} \quad \text{for all } n \in \mathbb{Z}. \tag{6.6}$$

It follows that, for each $k \geq 4$ and for all $n \in \mathbb{Z}$, there exist coprime polynomials f_n and g_n with integer coefficients such that, if (h_n) is *any* Somos k sequence (with coefficients λ_i and initial values $h_0, \ldots, h_{k-1}$, defined over a set of indices I), then

$$h_n = \frac{f_n(h_0, \ldots, h_{k-1}, \lambda_1, \ldots, \lambda_{\lfloor \frac{k}{2} \rfloor})}{g_n(h_0, \ldots, h_{k-1}, \lambda_1, \ldots, \lambda_{\lfloor \frac{k}{2} \rfloor})} \quad \text{for all } n \in I.$$

(Note that f_n and g_n depend only on n and k, and are the same for all Somos k sequences (h_n).)

Hickerson considered Somos sequences in which the coefficients λ_1 are all 1. For such sequences there exist coprime polynomials p_n and q_n in k variables with

integer coefficients such that

$$h_n = \frac{p_n(h_0, \ldots, h_{k-1})}{q_n(h_0, \ldots, h_{k-1})} \quad \text{for all } n \in I,$$

where

$$\frac{p_n(x_0, \ldots, x_{k-1})}{q_n(x_0, \ldots, x_{k-1})} = \frac{f_n(x_0, \ldots, x_{k-1}, 1, 1, \ldots, 1)}{g_n(h_0, \ldots, h_{k-1}, 1, 1, \ldots, 1)} \quad \text{for all } n \in \mathbb{Z}.$$

Techniques developed by Hickerson and expanded by Lotto (see [21]) have been used to prove

Theorem 6.5.1. *If $k \in \{4, 5, 6, 7\}$ then for all $n \in \mathbb{Z}$, q_n is a product of powers of the variables.*

(This is the "Condition H" referred to by Robinson in [21].) Since the initial values of Somos(k) are all 1, it follows from Theorem 6.5.1 that Somos(4) to Somos(7) are integer sequences.

It is easy to modify Hickerson's proof of Theorem 6.5.1 so that it works for Somos 4 sequences with arbitrary coefficients λ_1, λ_2, and we do this in the rest of this subsection. (We will need the polynomial dependence of f_n on the λ_i for our Theorem 6.6.8.)

We need the following lemma, adapted from [21].

Lemma 6.5.2. *If g_{2k} is a product of powers of the first k variables, and f_k is coprime to $f_{k+1}, \ldots, f_{2k}$ then for all $n \in \mathbb{Z}$, g_n is a product of powers of the first k variables.*

Proof: Note that the statement is trivially true for $0 \leq n \leq k-1$ (since then $f_n = x_n$ and $g_n = 1$). For $k \leq n \leq 2k-1$, noting that

$$\frac{f_n}{g_n} = \frac{g_{n-k}}{f_{n-k}} \cdot \sum_{i=1}^{\lfloor \frac{k}{2} \rfloor} x_{k-1+i} \cdot \frac{f_{n-i}}{g_{n-i}} \cdot \frac{f_{n-k+i}}{g_{n-k+i}} \quad \text{for all } n \in \mathbb{Z},$$

where $f_{n-k} = x_{n-k} \in \{x_0, x_1 \ldots, x_{k-1}\}$, and using the inductive assumption that $g_0, g_1, \ldots, g_{n-1}$ is a product of powers of $x_0, x_1 \ldots, x_{k-1}$, it is easy to show that g_n is a product of powers of $x_0, \ldots, x_{k-1}$ for $k \leq n \leq 2k-1$. This also holds

for $n = 2k$ by assumption. So suppose it holds for $n = 0, 1, \ldots, m-1$ for some $m > 2k$ (i.e., $g_0, \ldots, g_{m-1}$ are all products of powers of the variables $x_0, \ldots, x_{k-1}$). We prove it then holds for $n = m$.

Let (h_n) be *any* Somos k sequence. On the one hand, we have (using the Somos k recursion $m-k$ times to write h_m as a function of $h_1, \ldots, h_k$ and $\lambda_1, \ldots, \lambda_{\lfloor \frac{k}{2} \rfloor}$)

$$h_m = \frac{f_{m-1}(h_1, \ldots, h_k, \lambda_1, \ldots, \lambda_{\lfloor \frac{k}{2} \rfloor})}{g_{m-1}(h_1, \ldots, h_k, \lambda_1, \ldots, \lambda_{\lfloor \frac{k}{2} \rfloor})}, \tag{6.7}$$

and on the other hand (using the recursion $m-2k$ times to write h_m as a function of $h_{k+1}, \ldots, h_{2k}$ and $\lambda_1, \ldots, \lambda_{\lfloor \frac{k}{2} \rfloor}$)

$$h_m = \frac{f_{m-k-1}(h_{k+1}, \ldots, h_{2k}, \lambda_1, \ldots, \lambda_{\lfloor \frac{k}{2} \rfloor})}{g_{m-k-1}(h_{k+1}, \ldots, h_{2k}, \lambda_1, \ldots, \lambda_{\lfloor \frac{k}{2} \rfloor})}. \tag{6.8}$$

(We can do this because $m > 2k$.)

Now substituting for h_k in (6.7) and $h_{k+1}, \ldots, h_{2k}$ in (6.8) in terms of the initial values $h_0, \ldots, h_{k-1}$ and the coefficients $\lambda_1, \ldots, \lambda_{\lfloor \frac{k}{2} \rfloor}$, (i.e., substituting

$$h_{k+j} = \frac{f_{k+j}(h_0, \ldots, h_{k-1}, \lambda_1, \ldots, \lambda_{\lfloor \frac{k}{2} \rfloor})}{g_{k+j}(h_0, \ldots, h_{k-1}, \lambda_1, \ldots, \lambda_{\lfloor \frac{k}{2} \rfloor})}$$

for $j = 0, \ldots, k$), we obtain an expression for h_m whose denominator is a product of powers of the initial values and the polynomial f_k in the first case, and of the initial values and the polynomials $f_{k+1}, \ldots, f_{2k}$ in the second. Since f_k is coprime to $f_{k+1}, \ldots, f_{2k}$, the reduced denominator g_m must be a product of powers of the first k variables.

It follows by induction that the result holds for all $n \geq 0$. Since the reversed sequence $\ldots, h_3, h_2, h_1, h_0, \ldots$ is also a Somos k sequence with coefficients λ_i, the result then follows for $n < 0$ too. □

This leads to the following important result:

Theorem 6.5.3. *If $k = 4$ then, for all $n \in \mathbb{Z}$, g_n is a product of powers of the first four variables.*

Proof: By Lemma 6.5.2, we only need to check that the polynomial f_4 is coprime to the polynomials $f_5, \ldots, f_8$, and that g_8 is a product of powers of the first four variables.

By Theorem 6.2.2 (replacing the initial values $h_0, \ldots, h_3$ by the variables $x_0, \ldots, x_3$ and the coefficients λ_1, λ_2 by the variables x_4, x_5) we have

$$\begin{aligned}
f_4 &= x_4\, x_1\, x_3 + x_5\, x_2^2, \\
f_5 &= x_4\, x_2\, f_4 + x_5\, x_0\, x_3^2, \\
f_6 &= x_4^2\, x_0\, x_2\, x_3\, f_4 + x_4\, x_5\, x_0^2\, x_3^3 + x_5\, x_1\, f_4^2, \\
f_7 &= x_5^3\, x_0^3\, x_2\, x_3^4 + \left(x_4\, x_5^2\, x_0^2\, x_2^2\, x_3^2\right)\, f_4 \\
&\qquad + \left(x_4\, x_5\, x_0^2\, x_3^2\right)\, f_4^2 + x_4\, \left(x_4\, x_0\, x_2 + x_5\, x_1^2\right)\, f_4^3, \\
f_8 &= x_4\, x_5^3\, x_0^4\, x_3^6 \\
&\qquad + x_4\, x_5^2\, x_0^2\, x_3^4\, \left(3x_4\, x_0\, x_2 + 2x_5\, x_1^2\right)\, f_4 \\
&\qquad\qquad + 3x_4^2\, x_5\, x_0\, x_2\, x_3^2\, \left(x_4\, x_0\, x_2 + x_5\, x_1^2\right)\, f_4^2 \\
&\qquad\qquad\qquad + \left(x_4^3\, x_2^2 (x_4\, x_2\, x_0 + x_5\, x_1^2) + x_5^3\, x_1^3\, x_3\right) f_4^3,
\end{aligned}$$

while

$$\begin{aligned}
g_4 &= x_0, \\
g_5 &= x_0\, x_1, \\
g_6 &= x_0^2\, x_1\, x_2, \\
g_7 &= x_0^3\, x_1^2\, x_2\, x_3, \\
g_8 &= x_0^3\, x_1^3\, x_2^2\, x_3.
\end{aligned}$$

It is easy to see that the polynomials f_5, f_6, f_7 and f_8 are coprime to f_4. Since g_8 is a product of powers of x_0, x_1, x_2, x_3 the result follows by Lemma 6.5.2. □

6.5.2 Reasonable and unreasonable primes

Definition: Let (h_n) be a Somos 4 sequence and p a prime. If every term of (h_n) and the coefficients λ_i are p-integers, then for all $r \in \mathbb{N}$ the prime power p^r is said to be *reasonable* in (h_n); otherwise p^r is *unreasonable*.

Fortunately, for each Somos 4 sequence (h_n) only a few primes are unreasonable, and they are easy to find:

Theorem 6.5.4. *Let (h_n) be a Somos 4 sequence, and let p be an unreasonable prime in (h_n). Then either p divides the denominator of one of the coefficients λ_1, λ_2, or for every index t with $t, \ldots, t+3 \in I$, p divides either the numerator or the denominator of one of the four terms $h_t, \ldots, h_{t+3}$.*

Proof: If one of $h_t, \ldots, h_{t+3}$ is zero, then p divides it. Otherwise, the result follows from Theorem 6.5.3 applied to the t–translate of (h_n), which is also a Somos 4 sequence with the same coefficients and has $h_t, \ldots, h_{t+3}$ as initial values. □

So every set of 4 consecutive terms together with the denominators of the coefficients contains *all* the unreasonable primes. In particular, if the coefficients are integers and the initial values are all equal to 1, then all terms are integers.

Example: Looking at our earlier example (the Somos 4 sequence with initial values $2, 3, 5, 7$ and coefficients $\lambda_1 = 1$ and $\lambda_2 = 2$) again, we see that we can write it as

$$\ldots, \frac{1184}{3^2}, \frac{128}{3}, \frac{40}{3}, 4, 4, 2, 3, 5, 7, \frac{71}{2}, \frac{551}{2 \cdot 3}, \frac{1898}{3}, \frac{101125}{2 \cdot 3^2}, \ldots.$$

As expected (looking at the segment $4, 4, 2, 3$), the only unreasonable primes are 2 and 3.

Remark: If (h_n) is an elliptic sequence then it is easily proved using the doubling formula (4.5) that any prime appearing in the denominator of some term in (h_n) must also divide the numerator of h_2 or the denominator of h_2, h_3 or h_4. This is a stronger result than Theorem 6.5.4, which holds for general Somos 4 sequences.

6.6 Somos 4 sequences reduced modulo prime powers

Let (h_n) be a Somos 4 sequence, let p be a reasonable prime in (h_n), and let $r \in \mathbb{N}$. We are interested in the properties of the reduced sequence $(h_n \bmod p^r)$,

and particularly in how they change as r increases. This problem was posed by Raphael Robinson in [21].

6.6.1 Basic properties of $(h_n \bmod p^r)$

In this subsection we collect some basic properties of Somos 4 sequences reduced modulo prime powers, which we need in the next chapter.

Theorem 6.6.1. *Let (h_n) be a Somos 4 sequence with coefficients λ_1, λ_2, and let p^r be a reasonable prime power dividing some term h_k but coprime to $h_{k+1}\, h_{k-1}$. Then*

$$\lambda_1 \equiv \frac{h_{k+2}\, h_{k-2}}{h_{k+1}\, h_{k-1}} \bmod p^r \qquad \textit{and} \qquad \lambda_2 \equiv \frac{h_{k+3}\, h_{k-1}}{h_{k+1}^2} \bmod p^r.$$

Proof: This follows immediately from Theorem 6.2.1 on setting $h_k \equiv 0 \bmod p^r$ (since $p \nmid h_{k+1}\, h_{k-1}$). □

We will need the following result in the next chapter:

Theorem 6.6.2. *Let (h_n) be a Somos 4 sequence, and p a reasonable prime not dividing h_{-1}, h_0, h_1. Then (h_n) has an equivalent sequence each of whose terms is congruent to 1 modulo p if and only if (h_n) satisfies*

$$\frac{h_0\, h_2}{h_1^2} \equiv \frac{h_{-1}\, h_1}{h_0^2} \equiv \frac{h_{-2}\, h_0}{h_{-1}^2} \not\equiv 0 \bmod p.$$

Proof: The "only if" part of the theorem is obvious. For the "if" part, suppose (h_n) satisfies

$$\frac{h_0\, h_2}{h_1^2} \equiv \frac{h_{-1}\, h_1}{h_0^2} \equiv \frac{h_0\, h_{-2}}{h_{-1}^2} \equiv c \bmod p,$$

where $c \not\equiv 0 \bmod p$, and let (h'_n) be the equivalent sequence defined by

$$h'_n = c^{-\frac{1}{2}n(n+1)} \left(\frac{h_{-1}}{h_0}\right)^n \left(\frac{1}{h_0}\right) h_n \quad \text{for all } n \in I.$$

Then

$$h'_{-2} = c^{-1}\left(\frac{h_{-2}\,h_0}{h_{-1}^2}\right) \equiv 1 \bmod p,$$
$$h_{-1} = \left(\frac{h_0}{h_{-1}}\right)\left(\frac{1}{h_0}\right) h_{-1} \equiv 1 \bmod p,$$
$$h'_0 = \left(\frac{1}{h_0}\right) h_0 = 1,$$
$$h'_1 = c^{-1}\left(\frac{h_{-1}\,h_1}{h_0^2}\right) \equiv 1 \bmod p, \quad \text{and}$$
$$h'_2 = c^{-3}\left(\frac{h_{-1}}{h_0}\right)^2\left(\frac{1}{h_0}\right) h_2 = c^{-3}\left(\frac{h_{-1}\,h_1}{h_0^2}\right)\left(\frac{h_{-1}\,h_1}{h_0^2}\right)\left(\frac{h_0\,h_2}{h_1^2}\right) \equiv 1 \bmod p.$$

Substituting $h_n \equiv 1 \bmod p$ for $n = -2, \ldots, 2$ into the Somos 4 recursion gives

$$\lambda_1 + \lambda_2 \equiv 1 \bmod p,$$

and it is then easy to prove by induction that $h'_n \equiv 1 \bmod p$ for all $n \in \mathbb{Z}$. □

6.6.2 Primes dividing the coefficients

We now look at primes dividing one or both of the coefficients λ_i; it turns out that such primes occur very frequently in (h_n), or not at all.

Definition: Let (h_n) be a Somos 4 sequence. A reasonable prime power p^r which divides some term h_k is said to be *regular* in (h_n) if there exists a positive integer M such that for $n \in I$,

$$h_n \equiv 0 \bmod p^r \Leftrightarrow n \equiv k \bmod M.$$

Otherwise p^r is called *irregular*. M is called the *gap* of p^r.

For a fixed prime p, we will usually denote the gap of p^r by N_r for $r \in \mathbb{N}$.

Theorem 6.6.3. *Let (h_n) be a Somos 4 sequence with coefficients λ_1, λ_2, and let p be a reasonable prime which divides λ_1 but not λ_2. Then either*

1. *p divides no term of (h_n), or*

2. *p divides all terms of (h_n), or*

3. *p is regular in (h_n) with gap 2, and if p^w is the highest power of p dividing all multiples of p in (h_n) then $p^{2w} \mid \lambda_1$. If $p^{2w+1} \mid \lambda_1$ then no term of (h_n) is divisible by p^{w+1}; in other words, every multiple of p in (h_n) is divisible by exactly the same power of p.*

Proof: Suppose $p \mid h_k$, and let p^w be the highest power of p dividing all multiplies of p in (h_n). Then $h_k\, h_{k-4} = \lambda_1\, h_{k-1}\, h_{k-3} + \lambda_2\, h_{k-2}^2$ and $h_k\, h_{k+4} = \lambda_1\, h_{k+1}\, h_{k+3} + \lambda_2\, h_{k+2}^2$ imply that p^w divides both h_{k-2} and h_{k+2}. It follows that every second term in both directions from h_k is a multiple of p^w.

Now if p divides h_{k+1}, then $h_{k+1}\, h_{k-3} = \lambda_1\, h_k\, h_{k-2} + \lambda_2\, h_{k-1}^2$ implies that $p^w \mid h_{k-1}$. Similarly, if $p \mid h_{k-1}$ then p^w divides h_{k+1}. It follows that if p divides any two consecutive terms then all terms are multiples of p^w.

Now suppose that p has gap 2, i.e., $p \nmid h_{k+1}\, h_{k-1}$. Then since $h_{k+2}\, h_{k-2} = \lambda_1\, h_{k+1}\, h_{k-1} + \lambda_2\, h_k^2$ it follows that $p^{2w} \mid \lambda_1$. If some term of (h_n) is divisible by p^{w+1}, then for some choice of the index k, $p^w \parallel h_k$ and $p^{w+1} \mid h_{k+2}$. But then since $p^{2w+1} \mid h_{k+2}\, h_{k-2}$ and $p^{2w} \parallel \lambda_2\, h_k^2$, we have $p^{2w} \parallel \lambda_1$. The result follows. □

For example, in an EDS (h_n), if a prime p divides $\lambda_1 = h_2^2$ and not $\lambda_2 = -h_3$ then p is regular with gap 2.

Theorem 6.6.4. *Let (h_n) be a Somos 4 sequence with coefficients λ_1, λ_2, and let p be a reasonable prime which divides λ_2 but not λ_1. Then either*

1. *p divides no term of (h_n), or*

2. *p divides some two consecutive terms of (h_n), and no three consecutive terms are coprime to p, or*

3. *p is regular in (h_n) with gap 3, and if p^w is the highest power of p dividing all multiples of p in (h_n) then $p^w \mid \lambda_2$. If $p^{w+1} \mid \lambda_2$ then no term of (h_n) is divisible by p^{w+1}; in other words, every multiple of p in (h_n) is divisible by exactly the same power of p.*

Proof: Suppose $p \mid h_k$, and let p^w be the highest power of p dividing all multiples of p in (h_n). Then we have $h_{k+3}\, h_{k-1} = \lambda_1\, h_{k+2}\, h_k + \lambda_2\, h_{k+1}^2 \equiv 0 \bmod p$ and

$h_{k-3}\,h_{k+1} = \lambda_1\,h_{k-2}\,h_k + \lambda_2\,h_{k-1}^2 \equiv 0 \bmod p$. Hence if $p \nmid h_{k+1}$ then $p \mid h_{k-3}$, and if $p \nmid h_{k-1}$ then $p \mid h_{k+3}$. It follows that the maximum number of consecutive terms that can be coprime to p is two.

Now suppose that no two consecutive terms of (h_n) are divisible by p. Then $p \nmid h_{k-1}\,h_{k+1}$, and hence p divides both h_{k-3} and h_{k+3}. It follows that p divides every third term of (h_n) in both directions from h_k (and no other terms), i.e., p has gap 3.

Finally, suppose that p has gap 3. Then since $h_{k+3}\,h_{k-1} = \lambda_1\,h_{k+2}\,h_k + \lambda_2\,h_{k+1}^2$ it follows that $p^w \mid \lambda_2$. If some term of (h_n) is divisible by p^{w+1}, then for some choice of the index k, $p^w \parallel h_k$ and $p^{w+1} \mid h_{k+3}$. But then since $p^{w+1} \mid h_{k+3}\,h_{k-1}$, $p^w \parallel \lambda_1\,h_{k+2}\,h_k$ and $p \nmid h_{k+1}$, we have $p^w \parallel \lambda_2$. The result follows. □

For example, in an EDS (h_n), if a prime p divides $\lambda_2 = -h_3$ then p divides h_4 also, or p is regular with gap 3. (Of course in an EDS $p^w \parallel \lambda_2$, since $\lambda_2 = -h_3$.)

Theorem 6.6.5. *Let (h_n) be a Somos 4 sequence with coefficients λ_1, λ_2, and let p be a reasonable prime in (h_n), which divides both λ_1 and λ_2.*

Then p divides some two consecutive terms of (h_n). Also, p divides h_n or h_{n-4} whenever $n, n-4 \in I$, so the maximum number of consecutive terms that can be coprime to p is four.

Proof: It follows from $h_n\,h_{n-4} = \lambda_1\,h_{n-1}\,h_{n-3} + \lambda_2\,h_{n-2}^2$ that for all $n, n-4 \in I$ either h_n or h_{n-4} is divisible by p.

Let $p \mid h_k$. If $p \nmid h_{k-1}\,h_{k+1}$ then p divides both h_{k+3} and h_{k-3}. But p also divides $h_{k-2}\,h_{k+2}$, so p divides some two consecutive terms. □

For example, in an EDS (h_n), if a prime p divides $\lambda_1 = h_2^2$ and $\lambda_2 = -h_3$ then p divides all terms h_n with $|n| \neq 1$.

Theorem 6.6.6. *Let (h_n) be a Somos 4 sequence with coefficients λ_1, λ_2, and let p be a reasonable prime in (h_n), which divides neither λ_1 nor λ_2.*

Then either p divides all terms of (h_n), or there are at least three terms coprime to p between each two multiples of p in (h_n).

Proof: If p divides two terms of (h_n), let N be the minimal positive integer such that p divides h_k and h_{k+N} for some index k. Since $p \nmid \lambda_2$, if $N = 1$ (i.e., p divides h_k and h_{k+1}) then $h_{k+4}\,h_k = \lambda_1\,h_{k+3}\,h_{k+1} + \lambda_2\,h_{k+2}^2$ and $h_{k+1}\,h_{k-3} = \lambda_1\,h_k\,h_{k-2} + \lambda_2\,h_{k-1}^2$ imply that p divides both h_{k+2} and h_{k-1}. It follows that if p divides two consecutive terms then p divides all terms of (h_n).

Now note that if p divides h_k and either h_{k+2} or h_{k+3} then $h_{k+3}\,h_{k-1} = \lambda_1\,h_{k+2}\,h_k + \lambda_2\,h_{k+1}^2$ and $h_{k+2}\,h_{k-2} = \lambda_1\,h_{k+1}\,h_{k-1} + \lambda_2\,h_k^2$ imply that p divides one of h_{k+1} and h_{k-1}. So N cannot be 2 or 3, and hence $N \geq 4$. □

For example, in an EDS (h_n), if p is coprime to $\lambda_1 = h_2^2$ and $\lambda_2 = -h_3$ then p is regular with gap ≥ 4. Note that only if a reasonable prime p divides both λ_1 and λ_2 does p necessarily divide some term of a Somos 4 sequence (h_n).

The next theorem follows easily from the previous four.

Theorem 6.6.7. *Let (h_n) be a Somos 4 sequence with coefficients λ_1, λ_2, and let p be a reasonable prime which divides at least two terms of (h_n). Let N be the smallest positive integer such that p divides both h_k and h_{k+N} for some index k.*

1. *If $N = 1$ then either p divides all terms, or p divides λ_2.*
2. *If $N = 2$ then p divides λ_1 and not λ_2, and p is regular with gap 2.*
3. *If $N = 3$ then p divides λ_2 and not λ_1, and p is regular with gap 3.*
4. *If $N \geq 4$ then p does not divide λ_1 or λ_2.*

Remarks:

1. So if p is a reasonable prime dividing some but not all terms of (h_n), then p is coprime to $\lambda_1\,\lambda_2$ if and only if $N \geq 4$.
2. In an elliptic divisibility sequence (h_n), N can be 1 or 3 only if p divides $-h_1\,h_3 = \lambda_2$, and N can be 2 only if p divides $h_2^2 = \lambda_1$ and not $-h_1\,h_3 = \lambda_2$. This agrees with Theorem 6.6.7.
3. We prove in the next chapter that in fact a prime cannot divide just one term of a Somos 4 sequence (h_n).

4. We prove further (Theorem 7.6.4) that if $N \geq 4$ then p is regular. Hence all irregular primes divide λ_2 and some two consecutive terms of (h_n).

6.6.3 Periodicity in $(h_n \bmod p^r)$

The periodicity properties of Somos 4 sequences reduced modulo a prime power were considered by Robinson in [21]. His theoretical results are summarised in this section, and his experimental results in the next.

Let (h_n) be a Somos 4 sequence. Since the computation of h_n using the Somos 4 recursion involves division, it is not obvious that if four consecutive terms of $(h_n \bmod p^r)$ are repeated then the next term will also be repeated, even though it is true that each term depends only on the previous four terms. If $h_{s+j} \equiv h_{t+j} \bmod p^r$ then certainly $h_{s+4}\, h_s \equiv h_{t+4}\, h_t \bmod p^r$, but if $h_s \equiv h_t \equiv 0 \bmod p$ then we only know h_{s+4} and h_{t+4} are congruent modulo $\frac{p^r}{\gcd(p^r, h_s, h_t)}$.

However, Robinson showed that if some 4 consecutive terms of $(h_n \bmod p^r)$ which are coprime to p are repeated then $(h_n \bmod p^r)$ is periodic. This follows from the next theorem (which we will need in chapter 8).

Theorem 6.6.8. *Let (h_n) and (h'_n) be two Somos 4 sequences and let p^r be a reasonable prime power such that*

$$\lambda'_i \equiv \lambda_i \bmod p^r \text{ for } i = 1, 2,$$

and for some indices s and t

$$h'_{s+j} \equiv h_{t+j} \bmod p^r \quad \text{for} \quad j = 0, 1, 2, 3$$

and $p \nmid h_s\, h_{s+1}\, h_{s+2}\, h_{s+3}$. Then the t–translate of (h_n) is congruent modulo p^r to the s–translate of (h'_n), i.e.,

$$h'_{s+n} \equiv h_{t+n} \bmod p^r \qquad \text{whenever } n \in I_t \cap I'_s.$$

(Here (h_n) and (h'_n) could be the same or different sequences.)

Proof: Let $n \in I_t \cap I'_s$. By Theorem 6.5.3, we can use repeated application of the Somos 4 recursion (6.2) to write

$$h_{t+n} = \frac{f_n(h_t, h_{t+1}, h_{t+2}, h_{t+3}, \lambda_1, \lambda_2)}{g_n(h_t, h_{t+1}, h_{t+2}, h_{t+3}, \lambda_1, \lambda_2)}$$

and

$$h'_{s+n} = \frac{f_n(h'_s, h'_{s+1}, h'_{s+2}, h'_{s+3}, \lambda'_1, \lambda'_2)}{g_n(h'_s, h'_{s+1}, h'_{s+2}, h'_{s+3}, \lambda'_1, \lambda'_2)},$$

where g_n is a product of powers of the first four variables. Here the numerators are congruent modulo p^r, and the denominators are congruent modulo p^r and coprime to p. It follows that $h_{t+n} \equiv h'_{s+n} \bmod p^r$. □

Notice that we need 4 consecutive terms *coprime* to p to be the same in $(h_n \bmod p^r)$ and $(h'_n \bmod p^r)$ for Theorem 6.6.8 to work, as shown by the following example:

Example: Let $p = 5$, and let (h_n) and (h'_n) be the Somos 4 sequences with coefficients $\lambda_1 = \lambda'_1 = 2$, $\lambda_2 = \lambda'_2 = 8$ and initial values $5, 3, 1, 7$ and $25, 3, 1, 7$ respectively. Then $h_4 = \frac{2\cdot 7\cdot 3+8\cdot 1}{5} = 10$ and $h'_4 = \frac{2\cdot 7\cdot 3+8\cdot 1}{25} = 2$, so here $h_j \equiv h'_j \bmod p$ for $j = 0, 1, 2, 3$, but $h_4 \not\equiv h'_4 \bmod p$. The sequence $(h_n \bmod 5)$ is

$$\ldots, 0, 3, 1, 2, 0, 4, 1, 4, 0, 2, 1, 3, 0, 1, 1, 1, 0, 3, 1, 2, 0, 4, 1, 4, 0, 2, 1, 3, 0, 1, 1, 1, \ldots$$

(5 is regular with gap 4), and the sequence $(h'_n \bmod 5)$ is

$$\ldots, 0, 3, 1, 2, 2, 2, 0, 1, 2, 4, 4, 4, 0, 2, 4, 3, 3, 3, 0, 4, 3, 1, 1, 1, 0, 3, 1, 2, 2, 2, 0, 1, \ldots,$$

(5 is regular with gap 6).

If (h_n) and (h'_n) are the same sequence, then we have

Corollary 6.6.9. *[21]*

If (h_n) is a Somos 4 sequence, and if

$$h_{t+j} \equiv h_{s+j} \bmod p^r \text{ for } j = 0, 1, 2, 3$$

for some indices s and t and some reasonable prime p coprime to $h_t, \ldots, h_{t+3}$, then the sequence $(h_n \bmod p^r)$ is periodic, with period dividing $t - s$.

Robinson used Corollary 6.6.9 to prove that Somos(4) is periodic modulo p^r for every prime power p^r, by first proving that every five consecutive terms of Somos(4) are coprime (we will prove this in Corollary 7.6.3), and deducing that there are infinitely many blocks of four consecutive terms coprime to p, some two of which must therefore be congruent modulo p^r. This is not true in general for Somos 4 sequences, but we do have

Theorem 6.6.10. *Let (h_n) be a Somos 4 sequence containing at most one zero term, and p a reasonable prime such that there are at least four terms coprime to p between any two multiples of p. Then $(h_n \bmod p^r)$ is periodic for all r.*

Remarks:

1. The above proof will not work for Somos 4 sequences (h_n) in which p is irregular or has gap less than 5.

2. Notice that a Somos 4 sequence need not be periodic modulo every reasonable prime. For example, if p divides h_3 and h_4 in an EDS (h_n), then $(h_n \bmod p)$ is not periodic.

3. The above proof of mod p^r periodicity also works for Somos (5) (see [21]), but not for Somos(6) or (7), since for these sequences it is not true that every consecutive $k+1$ terms are pairwise coprime. The modulo p^r periodicity of Somos (6) and (7) remains open.

4. Robinson mentions that Michael Somos has a completely different proof of mod p^r periodicity in Somos(4) and Somos(5), but does not give a reference.

5. Theorem 6.6.10 gives an upper bound on the period mod p^r. Because there are $\left(p^{r-1}(p-1)\right)^4$ possibilities for a block of 4 terms coprime to p, we can find two such blocks $h_s, h_{s+1}, h_{s+2}, h_{s+3}$ and $h_t, h_{t+1}, h_{t+2}, h_{t+3}$ which are congruent mod p^r and such that

$$|s-t| \leq 5\left(p^{r-1}(p-1)\right)^4$$

(since the furthest apart these blocks can be is if we go through all possible blocks coprime to p before repeating one and each such block is followed by a term which is a multiple of p). By Corollary 6.6.9 it follows that the period is at most $5\left(p^{r-1}(p-1)\right)^4$. So modulo p^r periodicity in infinite Somos 4 sequences has been proved for regular primes with $N_1 \geq 5$ and for primes not dividing any term of the sequence, but with a very large upper bound on the period. In chapters 8 and 9 we will prove a much better upper bound for the period of any Somos 4 sequence reduced modulo any power of a regular prime with gap ≥ 4.

6.7 Robinson's conjectures

Robinson [21] made several conjectures on the properties of Somos(4) reduced modulo a prime power, the results of computer experiments. In this section we give his conjectures and outline the results we have obtained in trying to prove them for all Somos 4 sequences.

6.7.1 The pattern of zeroes

Note that, whereas every prime divides some non-zero term of every elliptic divisibility sequence, there are many primes that do not divide any term of a given Somos 4 sequence. For example, no term of Somos(4) is divisible by 5 or 29. However, for primes which do divide some term, we have

Conjecture 6.7.1. *If p is an odd prime dividing some term of Somos*(4)*, then for all $r \in \mathbb{N}$, p^r also divides some term.*

The restriction to odd primes is necessary; for example, if (h_n) is Somos(4) then 2 divides h_5 but no term of (h_n) is divisible by 4.

We prove in Theorem 7.6.6 that if (h_n) is *any* Somos 4 sequence and p is a reasonable prime coprime to $\lambda_1 \lambda_2$ and dividing some term but not all terms of (h_n), then either all multiples of p in (h_n) are divisible by exactly the same power of p, or for all $r \in \mathbb{N}$ some term of (h_n) is divisible by p^r. It remains open whether there are any primes other than 2 for which this happens in Somos(4), but we give an example of a Somos 4 sequence S and an odd prime p for which it does.

Conjecture 6.7.2. *Every prime power dividing some term of Somos*(4) *is regular in Somos*(4)*.*

(For example, in Somos(4) every 17th term is divisible by 11.) We prove in Theorem 7.6.4 that a reasonable prime p in a Somos 4 sequence (h_n) is irregular if and only if p divides some two consecutive terms but not all terms of (h_n), and that in this case $p \mid \lambda_2$. We also prove (Theorem 7.6.2) that if p is coprime to

$\lambda_1 \lambda_2$ then every power of p dividing some term of (h_n) is regular in (h_n). Since Somos(4) has $\lambda_1 = \lambda_2 = 1$, this proves Conjecture 6.7.2.

We have not yet been able to prove that if p is a regular prime dividing $\lambda_1 \lambda_2$ and p^r divides some term of (h_n) then p^r is regular.

Remark: Robinson mentions that Clifford S. Gardner has proved Conjecture 6.7.2, but does not give a reference.

Conjecture 6.7.3. *If p is a regular prime with gap N_1 in Somos(4) then some multiple of N_1 is "near" p.*

Robinson found that every prime $p < 2000$ that divides some term of Somos(4) has gap $N_1 < 1.1\,p + 6$. We prove in Theorem 7.6.5 that, if (h_n) is any Somos 4 sequence and p is a regular prime in (h_n), then some multiple of N_1 lies within the Hasse bound $2\sqrt{p}$ of $(p+1)$. Hence $N_1 \leq (p+1) + 2\sqrt{p}$, which is a sharper bound than $N_1 < 1.1p + 6$. This proves Conjecture 6.7.3.

Conjecture 6.7.4. *Let p be an odd prime with gap N_1 in Somos(4), and let p^w be the highest power of p such that all multiples of p in Somos(4) are divisible by p^w. If p^r appears as a factor in Somos(4) for $r > w$, then its gap is*

$$N_r = p^{r-w}\, N_1.$$

(Obviously $N_r = N_1$ for $r \leq w$.)

We prove in Theorem 7.6.7 that if (h_n) is any Somos 4 sequence and p is a regular odd prime with gap $N_1 \geq 4$ in (h_n), then for $r \geq w$ the gap of p^r in (h_n) is $N_r = p^{r-w}\, N_1$ (unless of course no term of (h_n) is divisible by p^{w+1}). The restriction $N_1 \geq 4$ is necessary, as shown by the example following Theorem 5.3.1 for EDSs. Since $N_1 \geq 5$ for every regular prime p in Somos(4), this proves Conjecture 6.7.4. We also prove a slightly weaker result for the $p = 2$ case, and the same example of an EDS used to show that Theorem 5.3.1 for EDSs cannot be improved is valid here.

6.7.2 Periodicity

Robinson found the period of Somos(4) modulo p to be unpredictable, but observed that once this is known, the period of Somos(4) modulo p^r for $r \in \mathbb{N}$ seems to follow:

Conjecture 6.7.5. *Let (h_n) be Somos(4). Then for all odd primes p and $r \in \mathbb{N}$, the period of $(h_n \bmod p^r)$ is equal to p^{r-1} times the period of $(h_n \bmod p)$.*

The $p = 2$ case is different. Robinson observed that the periods of Somos(4) modulo 2^r for $r \in \mathbb{N}$ are

$$\pi_1 = 5, \;\; \pi_2 = 10, \;\; \pi_3 = 10, \;\; \pi_4 = 20, \;\; \pi_5 = 40, \;\; \pi_6 = 80, \ldots,$$

i.e., Somos(4) has the same period modulo 8 as modulo 4.

We prove in Theorem 8.5.6 that if (h_n) is *any* Somos 4 sequence and p is an odd prime which is regular in (h_n) with gap $N_1 \geq 5$ then, as long as p^{w+1} divides some term of (h_n), there exists a positive integer u such that for $r \in \mathbb{N}$ the period of $(h_n \bmod p^r)$ is

$$\pi_r = \begin{cases} \pi_1 & \text{if} \quad r \leq u, \text{ and} \\ p^{r-u}\,\pi_1 & \text{if} \quad r \geq u. \end{cases}$$

Moreover, p^u divides every multiple of p in (h_n).

Since $N_1 \geq 5$ for every prime p dividing some term of Somos(4), this proves Conjecture 6.7.5 for such primes if u is always 1 for Somos(4) and if Conjecture 6.7.1 holds. It remains open whether u is always 1 for Somos(4) as Robinson's observations suggest, but we have found other Somos 4 sequences in which $u > 1$.

For odd primes p which do not divide λ_1 or any term of a Somos 4 sequence (h_n), we prove in Theorem 9.4.6 that there exist positive integers u and w with $u \leq w$ such that for $r \in \mathbb{N}$ the period of $(h_n \bmod p^r)$ is

$$\pi_r = \begin{cases} \pi_1 & \text{if } r \leq u, \\ p\,\pi_{r-1} & \text{if } u+1 \leq r \leq w \\ p^2\,\pi_w, \; p\,\pi_w \text{ or } \pi_w & \text{if } r = w+1 \\ p\,\pi_{r-1} & \text{if } r \geq w+2. \end{cases}$$

Moreover, $\pi_{w+1} = p\,\pi_w$ unless $\pi_w = \pi_1$ or $p\ \pi_1$. If $\pi_w = \pi_1$ then π_{w+1} is either $p\,\pi_1$ or $p^2\,\pi_1$, and if $\pi_w = p\ \pi_1$ then π_{w+1} is either $p\,\pi_w$ or π_w.

Although we have not yet been able to prove that $\pi_{w+1} = p\,\pi_w$ for all Somos 4 sequences (h_n), we suspect that this is true; in other words, that as r increases from 1 the period of $(h_n \bmod p^r)$ remains the same until r reaches some value u, and then increases by a factor of p each time.

Robinson's Conjecture 6.7.5 is that if (h_n) is Somos(4) and p is an odd prime then $\pi_{r+1} = p\,\pi_r$ for every $r \in \mathbb{N}$. We have proved that if $\pi_2 = p\,\pi_1$ then $\pi_{r+1} = p\,\pi_r$ for all but at most one value w of r, and if $\pi_{w+1} \neq p\,\pi_w$ then p does not divide any term of (h_n) and either

$$p \parallel \pi_w \text{ and } \pi_{w+1} = \pi_w, \quad \text{or} \quad p \nmid \pi_w \text{ and } \pi_{w+1} = p^2\,\pi_w.$$

Note that Conjecture 6.7.5 is the only one of Robinson's conjectures which applies to primes not dividing any term of the sequence.

Conjecture 6.7.6. *If (h_n) is Somos(4) and p is a regular prime with gap N_1 in (h_n), then the period of $(h_n \bmod p)$ is a multiple of N_1, and a divisor of $(p-1)\,N_1$.*

We prove in section 8.5 that if (h_n) is any Somos 4 sequence and p is a regular prime with gap $N_1 \geq 5$ in (h_n), then the period π_1 of $(h_n \bmod p)$ is a multiple of N_1 and a divisor of $2\,(p-1)\,N_1$. We find the conditions under which π_1 is a divisor of $(p-1)\,N_1$, and give an example of a Somos 4 sequence with $\tau_1 \nmid (p-1)$. We have not yet found a prime p with $\tau_1 \nmid (p-1)$ in Somos(4), so we have not proved or disproved Robinson's conjecture 6.7.6.

We also find an expression for the period π_r in terms of two constants related to the sequence, which can be easily found if we know the gap N_1 of p in (h_n). We have not been able to extend these results to the case where $N_1 \leq 4$.

Remark: Robinson found similar modular properties in Somos(5) (see [21]). In Somos(6) and (7) the multiples of a given prime are not regularly spaced, and it is not known whether Somos(6) and (7) are periodic modulo p^r.

The modulo p^r behaviour observed by Robinson in Somos(4) is so similar to that which we proved for elliptic divisibility sequences in chapter 5 that we are

tempted to try to rewrite our proofs to work for all Somos 4 sequences. We cannot do this, however, because Somos 4 sequences only satisfy the local recursion (6.2), while our proofs often use the elliptic sequence formula (4.1) with values of n other than 2.

In the next chapter we describe a recent result relating Somos 4 sequences (h_n) to the sequence of points $Q + [n]P$ on an elliptic curve, and we use this to prove that Robinson's first four conjectures on the pattern of zeroes hold for $(h_n \bmod p^r)$.

In chapter 8 we prove that for any Somos 4 sequence (h_n) and regular prime p with gap $N_1 \geq 5$ in (h_n) such that p^r divides some term h_k, there exists another Somos 4 sequence (ℓ_n) which is equivalent to (h_n) and congruent modulo p^r to (a segment of) the $(-k)$–translate of an EDS (Z_n). This will allow us to use our results on symmetry and periodicity in EDSs modulo p^r from chapter 5 to prove similar results for $(h_n \bmod p^r)$.

Unfortunately this method only works for primes p appearing in the sequence with gap ≥ 5. In chapter 9 we use the relationship between Somos 4 sequences and elliptic curves to prove a partial result on Robinson's Conjecture 6.7.5 for primes not dividing λ_1 or any term of (h_n).

Finally, we note that we have not proved any results about primes which divide λ_1 but not any term of (h_n), nor about prime powers p^r where $r \geq 2$ and p is irregular or has gap $N_1 \leq 4$. We have also not covered the periodicity of $(h_n \bmod p^r)$ in the $p = 2$ case, or in the case where some term of (h_n) is divisible by p but no term is divisible by p^r.

Chapter 7

The relationship between Somos 4 sequences and elliptic curves

In the same way that (almost) every elliptic sequence is associated with the sequence of points $[n]P$ where $P = (0, 0)$ is a non-singular point on an elliptic curve E, it has recently been proved that (almost) every Somos 4 sequence whose first coefficient λ_1 is a square is associated with the sequence of points $Q + [n]P$ for two non-singular points $P = (0, 0)$ and Q on an elliptic curve E.

In this chapter we prove this relationship and look at some of the immediate consequences for the theory of Somos 4 sequences. In chapter 9 we will use these results to prove that a weakened version of Robinson's Conjecture 6.7.5 holds for primes not dividing any term of a given Somos 4 sequence.

When describing a Somos 4 sequence S in this chapter we will find it convenient to treat s_{-1}, s_0, s_1, s_2 as initial values instead of s_0, s_1, s_2, s_3, since the associated point Q has a simpler formula in terms of these values. We will also assume that any Somos 4 sequence containing a zero term h_k terminates at h_k, instead of at h_{k+3} or h_{k-3}.

7.1 Going from curve to sequence

Let E be an elliptic curve over $\mathbb{Q}$ with equation

$$y^2 + a_1\,xy + a_3\,y = x^3 + a_2\,x^2 + a_4\,x + a_6.$$

For convenience, we restate the formulae (3.3) for the associated constants b_i:

$$\left.\begin{aligned} b_2 &= a_1^2 + 4a_2 \\ b_4 &= a_1\,a_3 + 2a_4 \\ b_6 &= a_3^2 + 4a_6 \\ b_8 &= a_1^2\,a_6 - a_1\,a_3\,a_4 + 4a_2\,a_6 + a_2\,a_3^2 - a_4^2, \end{aligned}\right\} \qquad (7.1)$$

and recall that

$$b_4^2 + 4b_8 = b_2\,b_6.$$

Let $P = (\overline{x}, \overline{y})$ and $Q = (x_0, y_0)$ be distinct non-singular points on $E(\mathbb{Q})$. For all $n \in \mathbb{Z}$ such that $Q + [n]P \neq \mathcal{O}$, denote the point $Q + [n]P$ by (x_n, y_n). Then the x-coordinates $\ldots, x_{-1}, x_0, x_1, \ldots$ form a sequence of rational numbers (x_n). If $Q + [k]P = \mathcal{O}$ for some minimum $k > 0$ then x_k is not defined, and we consider the sequence (x_n) to terminate to the right at x_{k-1}. Similarly, if $Q + [k]P = \mathcal{O}$ for some maximum $k < 0$ then the sequence (x_n) terminates to the left at x_{k+1}.

Definition: We define an associated family $S_{E,Q,P}$ of sequences (s_n) of rational numbers by

$$\begin{aligned} &s_{-1}, s_0 \quad \text{arbitrary non-zero rational numbers,} \\ &s_{n+1} = -\frac{(x_n - \overline{x})\,s_n^2}{s_{n-1}} \quad \text{for } n \geq 0, \quad \text{and} \\ &s_{n-1} = -\frac{(x_n - \overline{x})\,s_n^2}{s_{n+1}} \quad \text{for } n \leq -1. \end{aligned}$$

If $s_k = 0$ for some minimum $k > 0$ then s_{k+1} is not defined and we assume the sequence S terminates to the right at s_k. Similarly, if $s_k = 0$ for some maximum $k < 0$ then s_{k-1} is not defined, and we assume the sequence terminates to the left at s_k. Let I denote the set of indices n for which s_n is defined (i.e., before the sequence terminates).

A sequence $S \in S_{E,Q,P}$ is said to be *associated with* the sequence of points $Q + [n]P$ on E.

Remarks:

1. If $s_k = 0$ for some minimum $k > 0$ then we must have $x_{k-1} = \overline{x}$, i.e., $Q + [k-1]P = -P$ and $Q + [k]P = \mathcal{O}$. Similarly, if $s_k = 0$ for some maximum $k < 0$ then $x_{k+1} = \overline{x}$, i.e., $Q + [k+1]P = P$ and $Q + [k]P = \mathcal{O}$. So $S \in S_{E,Q,P}$ will contain a zero if and only if $Q \in \langle P \rangle$.

2. So S is in $S_{E,Q,P}$ if and only if

$$s_n = 0 \;\Leftrightarrow\; Q + [n]P = \mathcal{O}$$

and

$$x_n - \overline{x} = -\frac{s_{n-1}\, s_{n+1}}{s_n^2} \qquad \text{whenever } s_{n-1}, s_n, s_{n+1} \in I.$$

3. The condition $s_{-1}, s_0 \neq 0$ is equivalent to the condition $Q \neq P, \mathcal{O}$.

4. Note that all sequences $S \in S_{E,Q,P}$ have the same set of indices I. If $Q = [k]P$ for some $k > 0$ (so $s_{-k} = 0$) and P has infinite order, then the sequence is infinite to the right, and $I = \{-k, -k+1, \ldots\}$. If $Q = [k]P$ for $k > 0$ and P has finite order N, then the sequence terminates to both left and right, i.e., $I = \{-k, -k+1, \ldots, -k+N\}$. If $Q \notin \langle P \rangle$ then the sequence is infinite in both directions, i.e., $I = \mathbb{Z}$. Similar comments hold if $k < 0$.

5. The sequence (x_n) determines the sequence (s_n) given s_{-1} and s_0, and (s_n) determines (x_n).

   We can write each term s_n in terms of s_{-1}, s_0 and the x_i as follows:

Theorem 7.1.1. *Let $P = (\overline{x}, \overline{y})$ and $Q = (x_0, y_0)$ be distinct non-singular rational points on an elliptic curve E, and for all $n \in \mathbb{Z}$ such that $Q + [n]P \neq \mathcal{O}$, denote the point $Q + [n]P$ by (x_n, y_n). Let $S \in S_{E,Q,P}$. Then for all $n > 0$ in I,*

$$s_n = (-1)^{\frac{1}{2}n(n+1)} (x_{n-1} - \overline{x}) (x_{n-2} - \overline{x})^2 \ldots (x_1 - \overline{x})^{n-1} (x_0 - \overline{x})^n s_0 \left(\frac{s_0}{s_{-1}}\right)^n.$$

Similarly, for all $n > 1$ such that $-n$ is in I,

$$s_{-n} = (-1)^{\frac{1}{2}n(n-1)} (x_{-n+1} - \overline{x}) (x_{-n+2} - \overline{x})^2 \ldots (x_{-1} - \overline{x})^{n-1} s_0 \left(\frac{s_{-1}}{s_0}\right)^n.$$

Proof: The proof is by induction on n. It is easy to see that the result holds for $n = 1$ and $n = 2$, so assume it holds up to $n-1$ for some $n > 0$. Then by definition of s_n and the induction hypothesis (since $1 + 2 \cdot \frac{1}{2}(n-1)n - \frac{1}{2}(n-2)(n-1)$ has the same parity as $\frac{1}{2}n(n+1)$),

$$\begin{aligned} s_n &= -\frac{(x_{n-1} - \overline{x})\, s_{n-1}^2}{s_{n-2}} \\ &= -\frac{(x_{n-1} - \overline{x}) \left((-1)^{\frac{1}{2}(n-1)n} \left(\prod_{j=0}^{n-2} (x_j - \overline{x})^{n-1-j} \right) s_0 \left(\frac{s_0}{s_{-1}} \right)^{n-1} \right)^2}{(-1)^{\frac{1}{2}(n-2)(n-1)} \left(\prod_{j=0}^{n-3} (x_j - \overline{x})^{n-2-j} \right) s_0 \left(\frac{s_0}{s_{-1}} \right)^{n-2}} \\ &= (-1)^{\frac{1}{2}n(n+1)} (x_{n-1} - \overline{x}) (x_{n-2} - \overline{x})^2 \ldots (x_1 - \overline{x})^{n-1} (x_0 - \overline{x})^n s_0 \left(\frac{s_0}{s_{-1}} \right)^n. \end{aligned}$$

The proof for $n < 0$ is similar. □

We will usually assume that $P = (0, 0)$. Since the linear transformation

$$x \mapsto x - \overline{x}, \quad y \mapsto y - \overline{y}$$

moves $P = (\overline{x}, \overline{y})$ to the origin but does not affect the sequences $S \in S_{E,Q,P}$, the general case is not essentially different from the case $P = (0, 0)$.

It was proved recently by Nelson Stephens that every sequence in $S_{E,Q,P}$ is a Somos 4 sequence, whose coefficients and initial values can be calculated from the coefficients of E and coordinates of Q. (This also follows from independent work by Noam Elkies.)

Theorem 7.1.2. *[28]*

Let E be an elliptic curve over $\mathbb{Q}$ with equation

$$E : y^2 + a_1\, xy + a_3\, y = x^3 + a_2\, x^2 + a_4\, x,$$

and let $P = (0, 0)$ and $Q = (x_0, y_0)$ be non-singular rational points on E such that $Q \neq P, \mathcal{O}, -P$. Denote the point $Q + [n]P$ by (x_n, y_n) whenever $Q + [n]P \neq \mathcal{O}$.

Then the sequences S in $S_{E,Q,P}$ are precisely the Somos 4 sequences with coefficients

$$\lambda_1 = a_3^2 = b_6 \qquad \textit{and} \qquad \lambda_2 = a_4 (a_4 + a_1\, a_3) - a_3^2\, a_2 = -b_8, \tag{7.2}$$

and initial values

$$\left.\begin{array}{l} s_{-1},\, s_0 \quad \textit{arbitrary non-zero rational numbers} \\ s_1 = -\dfrac{x_0\, s_0^2}{s_{-1}} \\ s_2 = -\dfrac{(a_4\, x_0 - a_3 y_0)\, s_0^3}{s_{-1}^2}. \end{array}\right\} \qquad (7.3)$$

Proof: We first prove that every sequence in $S_{E,Q,P}$ is a Somos 4 sequence with coefficients and initial values given by (7.2) and (7.3). Let $S \in S_{E,Q,P}$ with initial values s_{-1}, s_0. By Theorem 3.5.1,

$$x_{n-1} \cdot x_n^2 \cdot x_{n+1} = -b_8 - b_6\, x_n$$

whenever $Q + [n]P \neq -P$, $\mathcal{O}$ or P (i.e., whenever $s_{n-1}\, s_n\, s_{n+1} \neq 0$). Substituting $x_t = -\frac{s_{t-1}\, s_{t+1}}{s_t^2}$ for $t = n-1, n, n+1$, we obtain

$$\frac{s_{n-2}\, s_n}{s_{n-1}^2} \cdot \left(\frac{s_{n-1}\, s_{n+1}}{s_n^2}\right)^2 \cdot \frac{s_n\, s_{n+2}}{s_{n+1}^2} = -b_8 - b_6 \left(-\frac{s_{n-1}\, s_{n+1}}{s_n^2}\right).$$

Simplifying and multiplying by s_n^2 gives

$$s_{n-2}\, s_{n+2} = b_6\, s_{n-1}\, s_{n+1} - b_8\, s_n^2,$$

so S is a Somos 4 sequence with coefficients

$$\lambda_1 = b_6 \qquad \text{and} \qquad \lambda_2 = -b_8.$$

Finally, since by Theorem 3.5.1 we have $x_0^2\, x_1 = a_4\, x_0 - a_3\, y_0$, we can use Theorem 7.1.1 to write s_1 and s_2 in terms of s_{-1}, s_0, the coordinates of Q and the coefficients of E as

$$\begin{aligned} s_1 &= -\frac{x_0\, s_0^2}{s_{-1}}, \qquad \text{and} \\ s_2 &= -\frac{(x_0^2\, x_1)\, s_0^3}{s_{-1}^2} = -\frac{(a_4\, x_0 - a_3\, y_0)\, s_0^3}{s_{-1}^2}. \end{aligned}$$

For the converse, let S be any Somos 4 sequence whose coefficients and initial values satisfy (7.2) and (7.3), and let S' be the sequence in $S_{E,Q,P}$ with initial values $s'_{-1} = s_{-1}$ and $s'_0 = s_0$. Then we have proved (above) that the

coefficients and initial values s'_1, s'_2 of S' satisfy (7.2) and (7.3), so they must be the same as those of S. Since a Somos 4 sequence is uniquely defined by its coefficients and four initial values, it follows that $S' = S$, and hence that S is in $S_{E,Q,P}$. □

We illustrate the above process with an example:

Example: Let E be the elliptic curve

$$E : y^2 + y = x^3 - x^2 - x$$

(so $a_1 = 0$, $a_3 = 1$, $a_2 = a_4 = -1$), let $P = (0, 0)$ and let $Q = (2, 1)$. Choose $s_{-1} = s_0 = 1$. Then Theorem 7.1.2 gives

$$\lambda_1 = a_3^2 = 1 \quad \text{and} \quad \lambda_2 = a_4^2 + a_1\, a_3\, a_4 - a_3^2\, a_2 = 2,$$

while

$$s_1 = -\frac{x_0\, s_0^2}{s_{-1}} = -2 \quad \text{and} \quad s_2 = -\frac{(a_4\, x_0 - a_3\, y_0)\, s_0^3}{s_{-1}^2} = 3.$$

So the Somos 4 sequence S with coefficients 1 and 2 and initial values $1, 1, -2, 3$ is in $S_{E,Q,P}$. This is

$$0, 1, 1, -2, 3, 11, -4, -115, -411, 2554, \ldots.$$

(Note that S terminates to the left at $s_{-2} = 0$, corresponding to $Q - [2]P = \mathcal{O}$.) So $x_{-1}, x_0, x_1, \ldots$ are

$$\frac{(0)(1)}{1^2}, \frac{(1)(-2)}{1^2}, \frac{(1)(3)}{(-2)^2}, \frac{(-2)(11)}{3^2}, \frac{(3)(-4)}{11^2}, \ldots$$

i.e.,

$$0, -2, \frac{3}{4}, -\frac{22}{9}, -\frac{12}{121}, \ldots$$

Remarks:

1. The reason we assumed $Q \neq -P$ (i.e., $s_1 \neq 0$) was so that the sequence would not terminate at s_1, i.e., so that s_2 would be defined.

2. Note that the coefficients λ_1, λ_2 of $S \in S_{E,Q,P}$ depend entirely on the coefficients a_i of E, not on the point Q. The initial values depend on both Q and the coefficients of E, as well as on the choice of s_{-1}, s_0.

3. If $S \in S_{E,Q,P}$ then the first coefficient λ_1 of S must be a square.

4. If the a_i are all integers then λ_1 and λ_2 are integers, coprime if and only if a_3 and a_4 are coprime.

5. Let t be an index with $t-1, t \in I$. If we replace Q by $Q' = Q + [t]P$ (so $x'_n = x_{n+t}$ whenever $Q' + [n]P \neq \mathcal{O}$) and choose s'_{-1}, s'_0 to be s_{t-1}, s_t, then $S' \in S_{E,Q',P}$ is the t–translate of $S \in S_{E,Q,P}$. (This can easily be proved using Theorem 7.1.1.) It follows that, for $t-1, t \in I$, each sequence in $S_{E,Q+[t]P,P}$ is the t–translate of a sequence in $S_{E,Q,P}$, and each sequence in $S_{E,Q,P}$ is the $(-t)$–translate of a sequence in $S_{E,Q+[t]P,P}$. (For this reason, $Q \in \{P, \mathcal{O}, -P\}$ is not an important special case.)

A different choice s'_{-1}, s'_0 for the initial values s_{-1}, s_0 gives an equivalent sequence S', also associated with the sequence of points $Q + [n]P$ on E:

Theorem 7.1.3. *Let S be any sequence in $S_{E,Q,P}$ for a given elliptic curve E and points Q and $P = (0,0)$. Then $S_{E,Q,P}$ is precisely the set of equivalent sequences S' given by*

$$s'_n = \beta^n \gamma\, s_n \quad \text{for all } n \in I,$$

where γ, β are non-zero rational numbers.

Proof: Let S' be any sequence in $S_{E,Q,P}$, and set

$$\gamma = \frac{s'_0}{s_0} \quad \text{and} \quad \beta = \frac{s'_0\, s_{-1}}{s'_{-1}\, s_0}.$$

Then it follows easily from Theorem 7.1.1 that

$$s'_n = \left(\frac{s'_0}{s_0}\right)\left(\frac{s'_0\, s_{-1}}{s'_{-1}\, s_0}\right)^n s_n \quad \text{for all } n \in I,$$

and hence that S' is equivalent to S under

$$s'_n = \gamma\, \beta^n\, s_n \quad \text{for all } n \in I.$$

Now note that as s'_{-1} and s'_0 run through all non-zero rational numbers, so do γ and β. It follows that all equivalent sequences S' of the above form are in $S_{E,Q,P}$. □

7.2 Going from sequence to curve

We now prove the converse of Theorem 7.1.2; that is, we prove that for (almost) every Somos 4 sequence S whose first coefficient λ_1 is a square, there exists an elliptic curve E and non-singular rational points $P = (0, 0)$ and Q on E such that S is in $S_{E,Q,P}$, and we find all such E, Q pairs for a given sequence S.

We consider the cases where λ_1 is a non-zero square, where λ_1 is not a square and where $\lambda_1 = 0$ separately, and we have to exclude the special case where S is equivalent to the constant sequence $\ldots, 1, 1, 1, \ldots$, for reasons which will become clear.

7.2.1 When λ_1 is a positive square

We first consider the case where $\lambda_1 \neq 0$ and λ_1 is a square.

Theorem 7.2.1. *[28]*

Let S be a Somos 4 sequence with non-zero initial values s_{-1}, s_0, s_1 and coefficients λ_1, λ_2, where λ_1 is a positive square and S is not equivalent to the constant sequence $\ldots, 1, 1, 1, \ldots$ Let I be the set of indices for which s_n is defined. Then the elliptic curves

$$E : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x$$

and non-singular rational points $P = (0, 0)$ and $Q = (x_0, y_0)$ on E for which $S \in S_{E,Q,P}$ are precisely those for which

$$\left.\begin{aligned} b_8 &= -\lambda_2 \\ b_6 &= \lambda_1 \\ b_4 &= \frac{s_{-1}\, s_2}{s_0\, s_1} + \frac{\lambda_1\, s_0^2}{s_{-1}\, s_1} + \frac{s_{-2}\, s_1}{s_{-1}\, s_0} \\ b_2 &= \frac{b_4^2 + 4b_8}{b_6} \end{aligned}\right\} \tag{7.4}$$

and

$$x_0 = -\frac{s_{-1}\, s_1}{s_0^2}, \qquad y_0 = \frac{-a_4 s_{-1}\, s_0\, s_1 + s_{-1}^2\, s_2}{a_3\, s_0^3}. \tag{7.5}$$

Proof: First let $P = (0,0)$ and $Q = (x_0, y_0)$ be non-singular rational points on an elliptic curve

$$E : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x$$

such that $S \in S_{E,Q,P}$; we need to show that E and Q satisfy (7.4) and (7.5).

Since $S \in S_{E,Q,P}$ we have

$$x_n = -\frac{s_{n-1}\, s_{n+1}}{s_n^2}, \quad \text{for } n = -1, 0, 1,$$

and by Theorem 7.1.2 we have

$$\lambda_1 = b_6 \ \text{ and } \ \lambda_2 = -b_8.$$

By Theorem 3.5.1 (since $a_3 \neq 0$ because $\lambda_1 = b_6 = a_3^2 \neq 0$) we have

$$y_0 = \frac{a_4\, x_0 - x_0^2\, x_1}{a_3} = \frac{-a_4\, s_{-1}\, s_0\, s_1 + s_{-1}^2\, s_2}{a_3\, s_0^3},$$

and (since x_0 is defined and non-zero because $s_{-1}\, s_0\, s_1 \neq 0$)

$$\begin{aligned} b_4 &= x_0\, x_1 - \frac{b_6}{x_0} + x_0\, x_{-1} \\ &= \frac{s_{-1}\, s_2}{s_0\, s_1} + \frac{\lambda_1\, s_0^2}{s_{-1}\, s_1} + \frac{s_{-2}\, s_1}{s_{-1}\, s_0}. \end{aligned}$$

The expression for b_2 follows from $b_4^2 + 4b_8 = b_2\, b_6$. So if $S \in S_{E,Q,P}$, then E and Q must satisfy (7.4) and (7.5).

For the converse, let E be an elliptic curve

$$E : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x$$

where the b_i are given by (7.4), and let $Q = (x_0, y_0)$ be a rational point on E given by (7.5). (It is easy to prove that Q does indeed lie on E.) We need to show that $P = (0,0)$ and Q are non-singular, and that S is in $S_{E,Q,P}$.

The point $P = (0, 0)$ is singular if and only if $a_3 = a_4 = 0$, so P is non-singular since $a_3^2 = b_6 = \lambda_1 \neq 0$. By Theorem 3.4.2 (since $x_0 \neq 0$) Q is singular if and only if x_0 and y_0 satisfy

$$2a_3\, y_0 + a_1\, a_3\, x_0 + b_6 = 0$$

and

$$2x_0^3 - b_4\, x_0 - b_6 = 0.$$

Substituting for λ_1, x_0 and $a_3\, y_0$ from (7.4) and (7.5) and using the fact that $b_4 = 2a_4 + a_1\, a_3$, we can rewrite these equations as

$$2\left(\frac{s_{-1}^2\, s_2}{s_0^3}\right) - b_4\left(\frac{s_{-1}\, s_1}{s_0^2}\right) + \lambda_1 = 0 \tag{7.6}$$

and

$$2\left(-\frac{s_{-1}\, s_1}{s_0^2}\right)^3 + b_4\left(\frac{s_{-1}\, s_1}{s_0^2}\right) - \lambda_1 = 0. \tag{7.7}$$

Using (7.4) to replace $b_4\, \frac{s_{-1}\, s_1}{s_0^2} - \lambda_1$ by $\frac{s_{-1}^2\, s_2}{s_0^3} + \frac{s_{-2}\, s_1^2}{s_0^3}$ and rearranging shows that (7.6) and (7.7) are equivalent to the condition

$$\frac{s_0\, s_2}{s_1^2} = \frac{s_{-1}\, s_1}{s_0^2} = \frac{s_{-2}\, s_0}{s_{-1}^2}.$$

But by Theorem 6.3.4 this is true if and only if S is equivalent to the constant sequence $\ldots, 1, 1, 1, \ldots$, contrary to our assumption. Hence Q must be a non-singular point. Since by (7.4) and (7.5) the sequence S has coefficients $\lambda_1 = b_6$ and $\lambda_2 = -b_8$ and initial values

$$s_1 = -\frac{x_0\, s_0^2}{s_{-1}}, \quad \text{and } s_2 = -\frac{(a_4\, x_0 - a_3\, y_0)\, s_0^3}{s_{-1}^2},$$

it follows from Theorem 7.1.2 that $S \in S_{E,Q,P}$. □

Equivalently, we have

Theorem 7.2.2. *[28]*

Let S be a Somos 4 sequence with non-zero initial values s_{-1}, s_0, s_1 and coefficients λ_1, λ_2, where λ_1 is a positive square and S is not equivalent to the

constant sequence $\ldots, 1, 1, 1, \ldots$ *Let* I *be the set of indices for which* s_n *is defined. Then the elliptic curves*

$$E : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x$$

and non-singular rational points $P = (0, 0)$ *and* $Q = (x_0, y_0)$ *on* E *for which* $S \in S_{E,Q,P}$ *are precisely those for which*

$$\left.\begin{aligned} a_4 &\quad \textit{arbitrary} \\ a_3 &= \pm\sqrt{\lambda_1} \\ a_1 &= \frac{1}{a_3}\left(-2a_4 + \frac{s_{-1}\,s_2}{s_0\,s_1} + \frac{\lambda_1\,s_0^2}{s_{-1}\,s_1} + \frac{s_{-2}\,s_1}{s_{-1}\,s_0}\right) \\ a_2 &= -\frac{a_4^2 + \lambda_2}{\lambda_1} + \frac{a_4\,\left(s_{-1}^2\,s_2 + s_{-2}\,s_1^2 + \lambda_1\,s_0^3\right)}{\lambda_1\,s_{-1}\,s_0\,s_1} \end{aligned}\right\} \tag{7.8}$$

and

$$x_0 = -\frac{s_{-1}\,s_1}{s_0^2}, \qquad y_0 = \frac{-a_4 s_{-1}\,s_0\,s_1 + s_{-1}^2\,s_2}{a_3\,s_0^3}. \tag{7.9}$$

Proof: We can write the a_i in terms of a_4, a_3 and the b_i as follows:

$$\begin{aligned} a_3 &= \pm\sqrt{b_6}, \\ a_1 &= \frac{1}{a_3}\left(-2a_4 + b_4\right), \quad \text{and} \\ a_2 &= \frac{a_4^2 + a_1\,a_3\,a_4 + b_8}{a_3^2} = \frac{a_4\,b_4 + b_8 - a_4^2}{b_6}. \end{aligned}$$

As a_4 varies over all rational numbers and a_3 over $\sqrt{\lambda_1}$ and $-\sqrt{\lambda_1}$, E varies over all elliptic curves for which the b_i satisfy (7.4). The result therefore follows from Theorem 7.2.1. □

We illustrate the above process with an example:

Example: Let S be the Somos 4 sequence with coefficients $\lambda_1 = 1$ and $\lambda_2 = 2$ and initial values $s_{-2} = 0, s_{-1} = 1, s_0 = 1, s_1 = -2, s_2 = -3$. Choose $a_4 = -1$ and $a_3 = 1$. Then Theorem 7.2.2 gives

$$a_1 = \frac{1}{a_3}\left(-2a_4 + \frac{s_{-1}\,s_2}{s_0\,s_1} + \frac{\lambda_1\,s_0^2}{s_{-1}\,s_1} + \frac{s_{-2}\,s_1}{s_{-1}\,s_0}\right) = 0,$$

$$a_2 = -\frac{(a_4^2 + \lambda_2)}{\lambda_1} + \frac{a_4\,\left(s_{-1}^2\,s_2 + s_{-2}\,s_1^2 + \lambda_1\,s_0^3\right)}{\lambda_1\,s_{-1}\,s_0\,s_1} = -1,$$

and

$$x_0 = -\frac{s_{-1}\, s_1}{s_0^2} = 2, \qquad y_0 = \frac{-a_4 s_{-1}\, s_0\, s_1 + s_{-1}^2\, s_2}{a_3\, s_0^3} = 1.$$

So if E is the elliptic curve

$$E : y^2 + y = x^3 - x^2 - x,$$

$P = (0, 0)$ and $Q = (2, 1)$, then S is in $S_{E,Q,P}$.

A different choice of a_4 and the sign of a_3 gives a different E, Q pair for which S is in $S_{E,Q,P}$, but all such curves E are birationally equivalent with $|u| = 1$:

Theorem 7.2.3. *Let S be a Somos 4 sequence with non-zero initial values s_{-1}, s_0, s_1 and coefficients λ_1, λ_2, such that $\lambda_1 \neq 0$ and $S \in S_{E,Q,P}$ for some non-singular rational point Q on an elliptic curve E.*

Then S is in $S_{E',Q',P'}$ if and only if E' is birationally equivalent to E, with $P = (0, 0)$ and Q mapped to $P' = (0, 0)$ and Q' respectively, under the admissible change of variables

$$x = x' \quad \text{and} \quad y = u\, y' + s\, x'$$

for some $s \in \mathbb{Q}$ and $u = \pm 1$.

Proof: Let E' be an elliptic curve and $P' = (0, 0)$ and $Q' = (x'_0, y'_0)$ non-singular rational points on E' such that $S \in S_{E',Q',P'}$. By Theorem 7.2.1, $b'_i = b_i$ for $i = 2, 4, 6, 8$, and hence by Theorem 3.3.5 E and E' are birationally equivalent under an admissible change of variables

$$x = x' \quad \text{and} \quad y = u\, y' + s\, x' \tag{7.10}$$

with $u = \pm 1$, $r = t = 0$ and $s \in \mathbb{Q}$. By Theorem 3.3.3,

$$u = \frac{a_3}{a'_3} = \pm\sqrt{\frac{b_3}{b'_3}} = \pm 1 \quad \text{and} \quad s = \frac{a_4 - a'_4}{a_3}.$$

Clearly (7.10) maps $P = (0, 0)$ to $P' = (0, 0)$. By Theorem 7.2.1, the coordinates of Q' are

$$x'_0 = -\frac{s_{-1}\, s_1}{s_0^2} = u^2\, x_0,$$

and

$$\begin{aligned} y_0' &= \frac{-a_4'\, s_{-1}\, s_0\, s_1 + s_{-1}^2\, s_2}{a_3'\, s_0^3} \\ &= \frac{-\left(a_4 - s\, a_3\right)\, s_{-1}\, s_0\, s_1 + s_{-1}^2\, s_2}{u\, a_3\, s_0^3} \\ &= u\, \left(\frac{-a_4\, s_{-1}\, s_0\, s_1 + s_{-1}^2\, s_2}{a_3\, s_0^3}\right) - s\, \left(-\frac{s_{-1}\, s_1}{s_0^2}\right) \\ &= u\, y_0 - s\, x_0 \end{aligned}$$

(where we have used the fact that $u = \pm 1$). Hence (7.10) maps Q to Q'.

Finally, we note that as E' and Q' vary through all elliptic curves and rational points for which $S \in S_{E',Q',P'}$, a_4' varies through all rational numbers and a_3' through $\pm\sqrt{\lambda_1}$. Hence s varies through all rational numbers, and u through ± 1, and so E' varies through all curves birationally equivalent to E with $u = \pm 1$ and $r = t = 0$. It follows that S is in $S_{E',Q',P'}$ for all of these curve – point pairs, and no others. □

Remarks:

1. Note that demanding $s_{-1}\, s_0\, s_1\, s_2 \neq 0$ is not a big restriction — if any of the initial values is zero then S is a short sequence anyway, and hence not very interesting.

2. It is easy to prove that b_4 is invariant under translation of the sequence S (and hence that all the b_i are). Let $t \in \mathbb{Z}$ such that $t-2, \ldots, t+3 \in I$. From the Somos 4 recursion, we have

$$\lambda_2\, s_t^2 = s_{t-2}\, s_{t+2} - \lambda_1\, s_{t-1}\, s_{t+1}$$

and also

$$\lambda_2, s_{t+1}^2 = s_{t+3}\, s_{t-1} - \lambda_1\, s_t\, s_{t+2}.$$

It follows that

$$s_{t+1}^2\, \left(s_{t-2}\, s_{t+2} - \lambda_1\, s_{t-1}\, s_{t+1}\right) = s_t^2\, \left(s_{t+3}\, s_{t-1} - \lambda_1\, s_t\, s_{t+2}\right),$$

from which

$$\lambda_1\, s_t^3\, s_{t+2} + s_{t-2}\, s_{t+1}^2\, s_{t+2} = \lambda_1\, s_{t-1}\, s_{t+1}^3 + s_{t-1}\, s_t^2\, s_{t+3}.$$

Dividing by $s_{t-1}\, s_t\, s_{t+1}\, s_{t+2}$ and adding $\frac{s_{t-1}\, s_{t+2}}{s_t\, s_{t+1}}$ to both sides, we get

$$\frac{s_{t-1}\, s_{t+2}}{s_t\, s_{t+1}} + \frac{\lambda_1\, s_t^2}{s_{t-1}\, s_{t+1}} + \frac{s_{t-2}\, s_{t+1}}{s_{t-1}\, s_t} = \frac{s_t\, s_{t+3}}{s_{t+1}\, s_{t+2}} + \frac{\lambda_1\, s_{t+1}^2}{s_t\, s_{t+2}} + \frac{s_{t-1}\, s_{t+2}}{s_t\, s_{t+1}},$$

and it follows that b_4 is invariant under translation.

3. Not surprisingly, replacing S by a t–translate S' replaces Q by $Q + [t]P$, but does not change P or E (if we still choose the same a_4 and a_3). This is easily seen from the fact that the b_i are invariant under translation of the sequence and the fact that by definition of S and Theorem 3.5.1

$$Q = \left(-\frac{s_{-1}\, s_1}{s_0^2}, \frac{-a_4\, x_0 - x_0^2\, x_1}{a_3} \right),$$

while

$$Q + [t]P = \left(-\frac{s_{t-1}\, s_{t+1}}{s_t^2}, \frac{-a_4\, x_t - x_t^2\, x_{t+1}}{a_3} \right).$$

4. Note that the formulae (7.8) and (7.5) produce a singular point Q if and only if S is equivalent to the constant sequence (h_n) given by $h_n = 1$ for all $n \in \mathbb{Z}$. This is not unexpected, since if for some sequence S the formulae give a singular point Q on a curve E, then all translates of S must give the *same* point Q (i.e., the condition for Q to be singular must be invariant under translation of S). Otherwise, if say a t–translate S' gives a different point Q' on the same curve, then Q' must be non-singular (since an elliptic curve has at most one singular point), and so by the previous remark we would expect the original sequence S to give the point $Q' - [t]P$, which is non-singular. It follows that $\frac{s_{n-1}\, s_{n+1}}{s_n^2}$ is constant for all $n \in \mathbb{Z}$, and hence by Theorem 6.3.4 that S is equivalent to (h_n).

   If the constant sequence (h_n) is in $S_{E,Q,P}$ for some elliptic curve E and non-singular point Q, then the point $Q + [n]P$ has x-coordinate $x_n = -\frac{h_{n-1}\, h_{n+1}}{h_n^2} = -1$ for all $n \in \mathbb{Z}$, so x_n is constant for all $n \in \mathbb{Z}$. This is only possible if $P = -P$, and hence $a_3 = 0$. We cover this case in the next section.

5. If a t–translate (h_n) of S satisfies the elliptic sequence equation (4.1), then $\lambda_1 = h_2^2$ and $\lambda_2 = -h_3$, so λ_1 is always a square. In this case $s_{t+n} = h_n$ for all $n \in I$, so in particular $s_t = 0$ and $s_{t+1} = 1$. Since our expressions (7.8) for the a_i are invariant under translation by $t+2$, we can rewrite them as

$$\begin{aligned}
a_4 &\text{ arbitrary} \\
a_3 &= \pm h_2 \\
a_1 &= \frac{1}{a_3}\left(-2a_4 + \frac{s_{t+1}\,s_{t+4}}{s_{t+2}\,s_{t+3}} + \frac{\lambda_1\,s_{t+2}^2}{s_{t+1}\,s_{t+3}} + \frac{s_t\,s_{t+3}}{s_{t+1}\,s_{t+2}}\right) \\
&= \frac{1}{a_3}\left(-2a_4 + \frac{h_4}{h_2\,h_3} + \frac{h_2^4}{h_3}\right) \\
&= \frac{1}{a_3}\left(\frac{h_4 + h_2^5 - 2h_2\,h_3\,a_4}{h_2^2\,h_3}\right)
\end{aligned}$$

and

$$\begin{aligned}
a_2 &= -\frac{(a_4^2 + \lambda_2)}{\lambda_1} + \frac{a_4\,\left(s_{t+1}^2\,s_{t+4} + s_t\,s_{t+3}^3 + \lambda_1\,s_{t+2}^3\right)}{\lambda_1\,s_{t+1}\,s_{t+2}\,s_{t+3}} \\
&= -\frac{(a_4^2 - h_3)}{h_2^2} + a_4\left(\frac{h_4 + h_2^5}{h_2^3\,h_3}\right) \\
&= \frac{h_2\,h_3^2 + (h_4 + h_2^5)a_4 - h_2\,h_3 a_4^2}{h_2^3 h_3}.
\end{aligned}$$

If we choose the sign of a_3 so that $a_3 = h_2$, then these are the same as our expressions for the a_i in Theorem 4.5.3. So with this choice Theorems 7.2.2 and 4.5.3 give the same elliptic curve E.

7.2.2 When λ_1 is not a square

If S is a Somos 4 sequence whose first coefficient λ_1 is not a square, then S is not in $S_{E,Q,P}$ for any non-singular point Q on an elliptic curve E. However, we now prove that S has an equivalent Somos 4 sequence which is.

Theorem 7.2.4. *Let S be a Somos 4 sequence with non-zero initial values s_{-1}, s_0, s_1 and coefficients λ_1, λ_2, where $\lambda_1 \neq 0$ and S is not equivalent to the*

constant sequence $\ldots, 1, 1, 1, 1, \ldots$ *Let* θ *be the squarefree part of* λ_1*, and let* S' *be the equivalent sequence given by*

$$s'_n = \theta^{\frac{1}{2}n(n+1)}\, s_n \qquad \text{for all } n \in I.$$

Then $S' \in S_{E,Q,P}$ *for some elliptic curve* E *and non-singular rational points* $P = (0, 0)$ *and* Q *on* E*.*

If we denote $Q + [n]P$ *by* (x_n, y_n) *whenever* $Q + [n]P \neq \mathcal{O}$*, then we have*

$$x_n = -\theta \cdot \frac{s_{n-1}\, s_{n+1}}{s_n^2}$$

whenever $n - 1, n, n + 1 \in I$*.*

In particular, E *has parameters*

$$\left.\begin{aligned} a_4 &\quad \text{arbitrary} \\ a_3 &= \pm\theta^2 \sqrt{\tfrac{\lambda_1}{\theta}} \\ a_1 &= \frac{1}{a_3}\left(-2a_4 + \theta^2 \cdot \left(\frac{s_{-1}\, s_2}{s_0\, s_1} + \frac{\lambda_1\, s_0^2}{s_{-1}\, s_1} + \frac{s_{-2}\, s_1}{s_{-1}\, s_0}\right)\right) \\ a_2 &= -\frac{a_4^2 + \theta^4\, \lambda_2}{\theta^3\, \lambda_1} + \frac{a_4}{\theta\, \lambda_1}\left(\frac{s_{-1}^2\, s_2 + s_{-2}\, s_1^2 + \lambda_1\, s_0^3}{s_{-1}\, s_0\, s_1}\right) \end{aligned}\right\} \tag{7.11}$$

and Q *is the point* (x_0, y_0) *with*

$$x_0 = -\theta \cdot \frac{s_{-1}\, s_1}{s_0^2} \qquad \text{and} \qquad y_0 = \frac{-\theta\, a_4\, s_{-1}\, s_0\, s_1 + \theta^3\, s_{-1}^2\, s_2}{a_3\, s_0^3}. \tag{7.12}$$

Proof: By Theorem 6.3.1, the first coefficient of S' is

$$\lambda'_1 = \theta^{6\cdot\frac{1}{2}}\, \lambda_1 = \theta^4 \cdot \frac{\lambda_1}{\theta},$$

which is clearly a square. It therefore follows from Theorem 7.2.2 that $S' \in S_{E,Q,P}$ for some elliptic curve E and non-singular rational point Q on E. Equations (7.11) and (7.12) are obtained by substituting $s'_n = \theta^{\frac{1}{2}n(n+1)}\, s_n$ for all n in the corresponding equations in Theorem 7.2.2 and simplifying.

Finally, we note that, whenever $n-1, n, n+1 \in I$,

$$\begin{aligned} x_n &= -\frac{s'_{n-1}\, s'_{n+1}}{{s'_n}^2} \\ &= -\frac{\theta^{\frac{1}{2}(n-1)(n)}\, s_{n-1} \cdot \theta^{\frac{1}{2}(n+1)(n+2)}\, s_{n+1}}{\left(\theta^{\frac{1}{2}n(n+1)}\, s_n\right)^2} \\ &= -\theta \cdot \frac{s_{n-1}\, s_{n+1}}{s_n^2}. \end{aligned}$$

□

Note that $s'_{-1} = s_{-1}$ and $s'_0 = s_0$, and if λ_1 is a square then $\theta = 1$ and $S' = S$.

Example: Let S be the Somos 4 sequence with initial values

$$s_{-1} = 1, \quad s_0 = 1, \quad s_1 = 2, \quad s_2 = -1,$$

and coefficients

$$\lambda_1 = 12 \quad \text{and} \quad \lambda_2 = 1.$$

Note that λ_1 is not a square; the squarefree part is $\theta = 3$.

Let S' be the equivalent Somos 4 sequence given by

$$s'_n = \theta^{\frac{1}{2}n(n+1)}\, s_n \quad \text{for all } n \in I.$$

So S' has initial values

$$s_{-2} = -75, \quad s'_{-1} = 1, \quad s'_0 = 1, \quad s'_1 = (3)(2) = 6, \quad s'_2 = (3^3)(-1) = -27,$$

and coefficients

$$\lambda'_1 = \theta^3\, \lambda_1 = 3^3 \cdot 12 = 3^4 \cdot 4 \quad \text{and} \quad \lambda_2 = \theta^4\, \lambda_2 = 3^4.$$

Let E be the elliptic curve with coefficients

$$\begin{aligned} a_4 &= 0 \\ a_3 &= \sqrt{\lambda_1} = 3^2 \cdot 2 = 18 \\ a_1 &= \frac{1}{a_3}\left(\frac{s_{-1}\, s_2}{s_0\, s_1} + \frac{\lambda_1\, s_0^2}{s_{-1}\, s_1} + \frac{s_{-2}\, s_1}{s_{-1}\, s_0}\right) = \frac{107}{2} \\ a_2 &= -\frac{\lambda_2}{\lambda_1} = -324, \end{aligned}$$

let $P = (0, 0)$ and let $Q = (x_0, y_0)$ where

$$x_0 = -\frac{s_{-1}\, s_1}{s_0^2} = -6, \qquad y_0 = \frac{-a_4 s_{-1}\, s_0\, s_1 + s_{-1}^2\, s_2}{a_3\, s_0^3} = -\frac{3}{2},$$

That is,

$$E : y^2 + \frac{107}{2}\, x\, y + 18y = x^3 - 324\, x^2,$$

$P = (0, 0)$, and $Q = (-6, -\frac{3}{2})$. Then by Theorem 7.2.2, $S' \in S_{E,Q,P}$.

7.2.3 When λ_1 is zero

We now consider the case where $\lambda_1 = 0$ and $\lambda_2 \neq 0$. It is easy to show that if $h_k = 0$ for some index k, then $h_{k-2} = 0$ if $k \geq 0$, or $h_{k+2} = 0$ if $k < 0$, so the sequence terminates after two terms. Hence the only important case here is the one where the sequence contains no zero terms.

Theorem 7.2.5. *Let S be a Somos 4 sequence with non-zero initial values s_{-1}, s_0, s_1 and coefficients $\lambda_1 = 0$ and $\lambda_2 \neq 0$. If*

$$s_{-1}^2\, s_2 = s_1^2\, s_{-2}, \qquad \textit{or, equivalently,} \qquad \lambda_2 = \left(\frac{s_{-1}\, s_2}{s_0\, s_1}\right)^2,$$

then the elliptic curves

$$E : y^2 + a_1 xy + a_3 y = x^3 + a_2 x^2 + a_4 x$$

and rational points $P = (0, 0)$ and $Q = (x_0, y_0)$ on E for which $S \in S_{E,Q,P}$ are precisely those which satisfy

$$x_0 = -\frac{s_{-1}\, s_1}{s_0^2}, \quad y_0 \textit{ arbitrary},$$

and

$$\begin{aligned} a_3 &= 0 \\ a_4 &= \frac{s_{-1}\, s_2}{s_0\, s_1} \\ a_1 &\textit{ arbitrary} \\ a_2 &= \frac{y_0^2 + a_1\, x_0\, y_0 - x_0^3 - a_4\, x_0}{x_0^2}, \end{aligned}$$

with the additional restriction that if S is equivalent to the constant sequence $\ldots, 1, 1, 1, \ldots$, then $2y_0 - a_1 \left(\frac{s_{-1}\, s_1}{s_0^2}\right) \neq 0$.

Otherwise, if S does not satisfy $s_{-1}^2\, s_2 = s_1^2\, s_{-2}$, then there is no E, Q pair with $S \in S_{E,Q,P}$.

Proof: First note that, since $s_{-1}, s_0, s_1, s_2 \neq 0$, S contains no zero terms and $I = \mathbb{Z}$.

Suppose $S \in S_{E,Q,P}$ for some elliptic curve $E : y^2 + a_1\, x\, y + a_3\, y = x^3 + a_2\, x^2 + a_4\, x$ and non-singular rational points $P = (0, 0)$ and $Q = (x_0, y_0)$. Then by Theorem 7.1.2 $\lambda_1 = a_3^2$, so

$$a_3 = 0.$$

Note that $a_3 = 0$ implies $-P = P$, i.e., that P has order 2 in $E(\mathbb{Q})$. So we must have $x_{n-1} = x_{n+1}$ for all $n \in \mathbb{Z}$, and hence S must satisfy

$$-\frac{s_{n+2}\, s_n}{s_{n+1}^2} = -\frac{s_n\, s_{n-2}}{s_{n-1}^2},$$

i.e.,

$$s_{n+2}\, s_{n-1}^2 = s_{n-2}\, s_{n+1}^2$$

for all $n \in \mathbb{Z}$. So in particular,

$$s_2\, s_{-1}^2 = s_{-2}\, s_1^2.$$

Equivalently, since $\lambda_1 = 0$ implies $s_2\, s_{-2} = \lambda_2\, s_0^2$, we have

$$\lambda_2 = \frac{s_2\, s_{-2}}{s_0^2} = \left(\frac{s_{-1}\, s_2}{s_0\, s_1}\right)^2.$$

By Theorem 3.5.1 with $a_3 = 0$, $x_n\, x_{n+1} = a_4$ for all $n \in \mathbb{Z}$, so S must satisfy

$$\frac{s_{n-1}\, s_{n+2}}{s_{n+1}\, s_n} = a_4 \quad \text{for all } n \in \mathbb{Z},$$

and in particular

$$a_4 = \frac{s_{-1}\, s_2}{s_0\, s_1}.$$

Finally, since Q lies on E, y_0, a_1 and a_2 must satisfy $y_0^2 + a_1\, x_0\, y_0 = x_0^3 + a_2\, x_0^2 + a_4\, x_0$. So

$$a_2 = \frac{y_0^2 + a_1\, x_0\, y_0 - x_0^3 - a_4\, x_0}{x_0^2}.$$

For the converse, let S satisfy $s_{-1}^2\, s_2 = s_1^2\, s_{-2}$ or, equivalently, $\lambda_2 = \left(\frac{s_{-1}\, s_2}{s_0\, s_1}\right)^2$. Let E be any elliptic curve $E : y^2 + a_1xy = x^3 + a_2x^2 + a_4x$ with $a_4 = \frac{s_{-1}\, s_2}{s_0\, s_1}$, and let $Q = (x_0, y_0)$ be any rational point on E with $x_0 = \frac{-s_{-1}\, s_1}{s_0^2}$. Clearly $P = (0, 0)$ is on E and is non-singular, since $a_4 \neq 0$.

By Theorem 3.4.2, Q is singular if and only if

$$2y_0 + a_1\, x_0 = 0 \quad \text{and} \quad 2x_0^3 - b_4\, x_0 = 0.$$

Since $x_0 = \frac{-s_{-1}\, s_1}{s_0^2} \neq 0$ and $b_4 = 2a_4 = 2\left(\frac{s_{-1}\, s_2}{s_0\, s_1}\right)$, the condition that $2x_0^3 - b_4\, x_0 = 0$ is equivalent to the condition that

$$2\left(\frac{s_{-1}^2\, s_1^2}{s_0^4}\right) = 2\left(\frac{s_{-1}\, s_2}{s_0\, s_1}\right),$$

i.e.,

$$\frac{s_{-1}\, s_1}{s_0^2} = \frac{s_0\, s_2}{s_1^2}.$$

Since S also satisfies $\frac{s_0\, s_2}{s_1^2} = \frac{s_{-2}\, s_0}{s_{-1}^2}$, it follows from Theorem 6.3.4 that Q is singular if and only if $2y_0 + a_1\, x_0 = 0$ and S is equivalent to the constant sequence $\ldots, 1, 1, 1, \ldots$ If this is not the case, then since S satisfies

$$\lambda_1 = 0 = b_6, \qquad \lambda_2 = a_4^2 = -b_8$$

and

$$s_1 = -\frac{x_0\, s_0^2}{s_{-1}}, \qquad s_2 = \frac{a_1\, s_0\, s_1}{s_{-1}} = -\frac{a_4\, x_0\, s_0^3}{s_{-1}^2},$$

it follows from Theorem 7.1.2 that $S \in S_{E,Q,P}$. □

Remarks:

1. It is in fact easy to prove by induction using the Somos 4 recursion that if S is a Somos 4 sequence with $\lambda_1 = 0$ and $S \in S_{E,Q,P}$ for some E, Q (which implies $\lambda_2 = \left(\frac{s_{-1}\, s_2}{s_0\, s_1}\right)^2$) then each term s_n is given in terms of the initial values s_{-1}, s_0, s_1, s_2 by

$$s_n = \begin{cases} \lambda_2^{k^2} \left(\dfrac{s_1}{s_{-1}}\right)^k s_0 & \text{if } n = 2k, \quad \text{or} \\ \lambda_2^{k(k+1)} \left(\dfrac{s_1}{s_{-1}}\right)^k s_1 & \text{if } n = 2k+1. \end{cases}$$

2. If S is a Somos 4 sequence with $\lambda_1 = \lambda_2 = 0$ then S cannot be in $S_{E,Q,P}$ for any elliptic curve E, because E would need to have $a_3 = 0$ and $a_4 = 0$, and $P = (0, 0)$ is singular in any such curve.

7.3 Equivalent curves and sequences

It turns out, not surprisingly, that equivalent sequences are associated with equivalent curves:

Theorem 7.3.1. *Let $S \in S_{E,Q,(0,0)}$ and $S' \in S_{E',Q',(0,0)}$. Then S' is equivalent to S under*

$$s'_n = \alpha^{n^2} \beta^n \gamma \, s_n \quad \text{for all } n \in I$$

for some $\alpha, \beta, \gamma \in \mathbb{Q}$ if and only if E' is birationally equivalent to E, with Q mapped to Q', under the admissible change of variables

$$x = u^2 \, x' \quad \text{and} \quad y = u^3 \, y' + u^2 \, s \, x'$$

for

$$u = \pm \left(\frac{1}{\alpha} \right)$$

and some $s \in \mathbb{Q}$.

Proof: First let E and E' be birationally equivalent, with $P = (0, 0)$ and Q mapped to $P' = (0, 0)$ and Q' respectively, under the admissible change of variables

$$x = u^2 x' \quad \text{and} \quad y = u^3 y' + u^2 s x'$$

for some $u, s \in \mathbb{Q}$ with $u \neq 0$.

Let $\overline{S}$ be the particular sequence in $S_{E',Q',P'}$ with initial values

$$\overline{s}_{-1} = u^{-1} \, s_{-1} \quad \text{and} \quad \overline{s}_0 = s_0.$$

By Theorem 7.1.1, for all $n > 0$ in I,

$$\begin{aligned} s_n &= (-1)^{\frac{1}{2}n(n+1)}\, x_{n-1}\, x_{n-2}^2 \cdots x_1^{n-1}\, x_0^n\, s_0 \left(\frac{s_0}{s_{-1}}\right)^n \\ &= (-1)^{\frac{1}{2}n(n+1)} \left(u^2\, x'_{n-1}\right) \left(u^2\, x'_{n-2}\right)^2 \cdots \left(u^2\, x'_1\right)^{n-1} \left(u^2\, x'_0\right)^n \overline{s}_0 \left(\frac{\overline{s}_0}{u\,\overline{s}_{-1}}\right)^n \\ &= \left(u^2\right)^{\frac{1}{2}n(n+1)} \cdot \left(\frac{1}{u}\right)^n \cdot \left((-1)^{\frac{1}{2}n(n+1)}\, x'_{n-1}\, {x'_{n-2}}^2 \ldots {x'_1}^{n-1}\, {x'_0}^n\, \overline{s}_0 \left(\frac{\overline{s}_0}{\overline{s}_{-1}}\right)^n\right) \\ &= u^{n^2}\, \overline{s}_n. \end{aligned}$$

The proof for $n < -1$ is similar. It follows that $\overline{S}$ is equivalent to S under

$$\overline{s}_n = u^{-n^2} s_n \quad \text{for all } n \in I,$$

and hence by Theorem 7.1.3 that $S_{E',Q',P'}$ is precisely the set of equivalent sequences S' given by

$$s'_n = \left(\frac{1}{u}\right)^{n^2} \beta^n\, \gamma\, s_n \;=\; \left(-\frac{1}{u}\right)^{n^2} (-\beta)^n\, \gamma\, s_n \quad \text{for all } n \in I,$$

for non-zero rational numbers γ, β.

For the converse, let S and S' be equivalent under

$$s'_n = \alpha^{n^2}\, \beta^n\, \gamma\, s_n \quad \text{for all } n \in I,$$

where α, β and γ are rational numbers, and let $\overline{E}$ be the particular curve birationally equivalent to E, with Q mapped to $\overline{Q}$, under the admissible change of variables

$$x = u^2\, x' \quad \text{and} \quad y = u^3\, y',$$

where

$$u = \pm\frac{1}{\alpha}.$$

Then by the above argument, S' is in $S_{\overline{E},\overline{Q},(0,0)}$, and it follows by Theorem 7.2.3 that the elliptic curves E' and non-singular rational points $P' = (0,0)$ and Q' on E' for which $S' \in S_{E',Q',P'}$ are precisely those E' which are birationally equivalent to E, with $P = (0,0)$ and Q mapped to $P' = (0,0)$ and Q' respectively, under the admissible change of variables

$$x = u^2 x' \quad \text{and} \quad y = u^3 y' + u^2 s x'$$

for some $s \in \mathbb{Q}$. □

Remarks:

1. Notice that if $S \in S_{E,Q,P}$ is equivalent to $S' \in S_{E',Q',P'}$ under $s'_n = u^{-n^2}\,\beta^n\,\gamma\,s_n$ for all $n \in I$, then the coefficients of S' are $u^{-6}\,b_6 = b'_6$ and $u^{-8}\,(-b_8) = -b'_8$, as expected.

2. If $|u| = 1$ then $S_{E',Q',P'} = S_{E,Q,P}$, as in Theorem 7.1.3.

3. Varying u and s over all rational numbers with $u \neq 0$ gives all elliptic curves E' birationally equivalent to E over $\mathbb{Q}$ with $P = (0,0)$ mapped to $P' = (0,0)$. Notice however that varying u, γ and β over all non-zero rational numbers does not give all sequences S' equivalent to S (see Theorem 6.3.2).

7.4 Singular curves and sequences

We now characterise the Somos 4 sequences S which are associated with singular elliptic curves; in other words, we find a condition on the coefficients and initial values of S under which all curves E with $S \in S_{E,Q,P}$ are singular. (Remember that P and Q are still non-singular points.)

Definition: Let S be a Somos 4 sequence in which $\lambda_1 \neq 0$ and $s_{k-1}, s_k, s_{k+1} \neq 0$ for some index k. Let

$$\chi = \frac{s_{k-1}\,s_{k+2}}{s_k\,s_{k+1}} + \frac{\lambda_1\,s_k^2}{s_{k-1}\,s_{k+1}} + \frac{s_{k-2}\,s_{k+1}}{s_{k-1}\,s_k}.$$

(It is easy to see that χ is invariant under translation of S.) Then the *discriminant* of S is

$$\Delta(S) = \left(\frac{\lambda_2}{\lambda_1^2}\right)\,\chi^4 + \chi^3 - 8\left(\frac{\lambda_2^2}{\lambda_1^2}\right)\,\chi^2 - 36\,\lambda_2\,\chi + 16\left(\frac{\lambda_2^3}{\lambda_1^2}\right) - 27\lambda_1^2.$$

A Somos 4 sequence S is said to be *singular* if $\Delta(S) = 0$, and *singular modulo* p for a prime p if $\Delta(S) \equiv 0 \bmod p$. Otherwise S is said to be *non-singular*, or *non-singular modulo* p.

Clearly $\Delta(S)$ is invariant under translation of S, since χ is. It is also easy to prove that if S and S' are equivalent Somos 4 sequences with $s_n = \alpha^{n^2}\,\beta^n\,\gamma\,s_n$ for all n, then

$$\Delta(S) = \alpha^{12}\,\Delta(S').$$

If S satisfies the elliptic sequence equation (4.1) then (setting $k = 2$), $\chi = \frac{s_4+s_2^5}{s_2\,s_3}$, and so $\Delta(S)$ matches our definition of the discriminant of an elliptic sequence from section 4.5.2.

Since in Theorem 7.2.1 $b_4 = \chi$, $b_6 = \lambda_1$ and $b_8 = -\lambda_2$ and the discriminant of an elliptic curve is

$$\Delta = -b_2^2\,b_8 - 8b_4^3 - 27b_6^2 + 9b_2\,b_4\,b_6,$$

we have

Theorem 7.4.1. *If $S \in S_{E,Q,P}$, then the discriminant of S is equal to the discriminant of E. So singular Somos 4 sequences are associated with singular elliptic curves.*

7.5 Sequences containing a zero term

Recall from section 3.10 that for every elliptic curve E there exists a sequence of "division polynomials" $(\psi_n(x,y))$ such that for every rational point $P = (x,y)$ on E, the x-coordinate of $[n]P$ is given by

$$x - \frac{\psi_{n-1}(x,y)\cdot\psi_{n+1}(x,y)}{\psi_n(x,y)^2} \quad \text{whenever } [n]P \neq \mathcal{O},$$

and $(\psi_n(x,y))$ is an elliptic sequence.

We now show that if $P = (0,0)$ and $Q = [k]P$ for some $k \in \mathbb{Z}$, then $S_{E,Q,P}$ contains (a segment of) the k–translate of the sequence $(\psi_n(0,0))$ obtained by substituting the coordinates of P into the sequence of division polynomials.

Theorem 7.5.1. *Let E be an elliptic curve over $\mathbb{Q}$ containing the point $P = (0,0)$, let $Q = [k]P$ for some $k \neq 0, 1$, and let $S \in S_{E,Q,P}$ be the associated Somos 4 sequence with initial values*

$$s_{-1} = \psi_{k-1}(0,0) \quad \textit{and} \quad s_0 = \psi_k(0,0),$$

where ψ_n is the nth division polynomial of E.

Then

$$s_n = \psi_{k+n}(0,0) \quad \textit{for all } n \in I.$$

Proof: Denote $Q + [n]P = (x_n, y_n)$ for all $n \in \mathbb{Z}$, and note that

$$Q + [n]P = [k+n]P \quad \text{whenever } [n+k]P \neq \mathcal{O}.$$

It follows by Theorem 3.10.2 and the definition of $S \in S_{E,Q,P}$ that

$$-\frac{s_{n-1}\, s_{n+1}}{s_n^2} = x_n = -\frac{\psi_{k+n-1}(0,0) \cdot \psi_{k+n+1}(0,0)}{\psi_{k+n}(0,0)^2}$$

whenever $n-1, n, n+1 \in I$. Since $s_{-1} = \psi_{k-1}(0,0)$ and $s_0 = \psi_k(0,0)$, it is now easy to prove by induction that $s_n = \psi_{n+k}(0,0)$ for all $n \in I$. □

Remarks:

1. By Theorem 3.10.9 ($\psi_n(0,0)$) is an elliptic sequence, and by Theorem 3.10.3 it has initial values $\psi_1(0,0) = 1$, $\psi_2(0,0) = a_3$, and $\psi_3(0,0) = b_8$. Hence it is also a Somos 4 sequence with coefficients

   $$\psi_2(0,0)^2 = a_3^2 = b_6 \quad \text{and} \quad -\psi_1(0,0) \cdot \psi_3(0,0) = -b_8,$$

   as expected from Theorem 7.1.2.

2. Note that at most four sequences in $S_{E,Q,P}$ can have a translate satisfying the elliptic sequence equation (4.1). (This follows from the fact that (s_n) and $(\gamma\, \beta^n\, s_n)$ cannot both satisfy (4.1) unless $\beta = \pm 1$ and $\gamma = \pm 1$.)

3. Let S be any Somos 4 sequence whose first coefficient λ_1 is a positive square and which contains a zero term s_{-k} (not in the initial values s_{-1}, s_0, s_1, s_2). By Theorem 7.2.1 there exists an elliptic curve E containing non-singular rational points $P = (0,0)$ and Q such that S is in $S_{E,Q,P}$. Since $s_k = 0$, Q must be $[k]P$. By Theorem 7.1.3 the sequences in $S_{E,Q,P}$ are precisely the Somos 4 sequences equivalent to S given by

   $$s'_n = \beta^n\, \gamma\, s_n$$

for all $n \in I$, and by Theorem 7.5.1 one of these is (a segment of) the k–translate of the sequence $(\psi_n(0,0))$. This recalls Theorem 6.4.5, in which we showed that for any Somos 4 sequence S whose first coefficient λ_1 is a square and which contains a zero term s_r, there exists an equivalent sequence (ℓ_n) given by

$$\ell_n = \beta^n \gamma \, s_n \quad \text{for all } n \in I$$

for some $\beta, \gamma \in \mathbb{Q}$, an r–translate of which satisfies the elliptic sequence equation (4.1).

Finally, let E be an elliptic curve containing the non-singular rational points $P = (0,0)$ and $Q = [t]P$, let M be any positive integer such that $[M]P \neq \mathcal{O}$, and let $Z_n = \psi_n(0,0)$ for all $n \in \mathbb{Z}$, where ψ_n is the nth division polynomial of E. We recall that since the sequence (Z_n) is an elliptic sequence, by Theorem 4.6.2 the subsequence (ℓ_n) obtained by taking every Mth term from Z_t is a Somos 4 sequence with coefficients $\left(\frac{Z_{2M}}{Z_M}\right)^2$ and $-\frac{Z_{3M}}{Z_M}$. Not surprisingly, (ℓ_n) turns out to be equivalent to the Somos 4 sequences associated with the sequence of points $[t]P + [n][M]P$, $n \in \mathbb{Z}$, on E.

Theorem 7.5.2. *Let E be an elliptic curve containing the non-singular rational points $P = (0,0)$ and $Q = [t]P$, let M be any positive integer such that $[M]P \neq \mathcal{O}$, and let $[M]P = (\overline{x}, \overline{y})$. Let $Z_n = \psi_n(0,0)$ for all $n \in \mathbb{Z}$, where ψ_n is the nth division polynomial of E, and let (ℓ_n) be the sequence obtained from (Z_n) by taking every Mth term from Z_t, i.e.,*

$$\ell_n = Z_{t+nM} \quad \textit{for all } n \in \mathbb{Z}.$$

Let (s_n) be the particular Somos 4 sequence in $S_{E,[t]P,[M]P}$ whose initial values are $s_{-1} = \frac{1}{Z_M} \cdot \ell_{-1}$ and $s_0 = \ell_0$. Then (s_n) is equivalent to (ℓ_n) (or a segment of ℓ_n if (s_n) terminates) under

$$s_n = \left(\frac{1}{Z_M}\right)^{n^2} \ell_n \quad \textit{for all } n \in I.$$

Proof: We have for all $n-1, n, n+1 \in I$, by definition of $S \in S_{E,[t]P,[M]P}$,

$$-\frac{s_{n-1}\, s_{n+1}}{s_n^2} \;=\; x([t]P + [n][M]P) - x([M]P) \;=\; x([t+nM]P) - x([M]P).$$

But by Theorem 3.10.2 for the division polynomials, whenever $[t+nM]P \neq \mathcal{O}$,

$$\begin{aligned} & x([t+nM]P) - x([M]P) \\ & \quad = -\frac{Z_{t+nM+1}\, Z_{t+nM-1}}{Z_{t+nM}^2} - \frac{Z_{M+1}\, Z_{M-1}}{Z_M^2} \\ & \quad = -\frac{1}{Z_M^2} \cdot \frac{1}{Z_{t+nM}^2} \cdot \left(Z_{t+nM+1}\, Z_{t+nM-1}\, Z_M^2 - Z_{M+1}\, Z_{M-1}\, Z_{t+nM}^2 \right) \end{aligned}$$

Since by Theorem 3.10.9 (Z_n) satisfies the elliptic sequence equation (4.1), this is

$$\begin{aligned} -\frac{s_{n-1}\, s_{n+1}}{s_n^2} &= -\frac{1}{Z_M^2} \cdot \left(\frac{Z_{t+(n+1)M}\, Z_{t+(n-1)M}}{Z_{t+nM}^2} \right) \\ &= -\frac{1}{Z_M^2} \cdot \left(\frac{\ell_{n+1}\, \ell_{n-1}}{\ell_n^2} \right) \end{aligned}$$

for all $n-1, n, n+1 \in I$.

Since $s_{-1} = \frac{1}{Z_M} \cdot \ell_{-1}$ and $s_0 = \ell_0$, it is easy to prove by induction that

$$s_n = \left(\frac{1}{Z_M} \right)^{n^2} \ell_n \quad \text{for all } n \in I.$$

So S is equivalent to (ℓ_n) (or a segment of (ℓ_n) if $[t+nM]P = \mathcal{O}$ for some $n \in \mathbb{Z}$, since then $s_n = \ell_n = 0$ but S terminates while (ℓ_n) doesn't). □

7.6 Consequences for Somos 4 sequences modulo prime powers

The relationship between Somos 4 sequences and elliptic curves has some important consequences for the pattern of zeroes in a Somos 4 sequence considered modulo a reasonable prime power. In particular, it allows us to prove almost all Robinson's conjectures on the pattern of zeroes in $(h_n \bmod p^r)$.

Theorem 7.6.1. *Let S be a Somos 4 sequence with non-zero initial values and coefficients λ_1, λ_2, and let θ be the squarefree part of λ_1. Let p be a reasonable prime in S, not dividing λ_1 and not dividing all terms of S, such that either*

1. *p divides some term of S and p does not divide λ_2, or*
2. *p does not divide any term of S, and S is not equivalent to a sequence which is constant modulo p.*

Then there exists an elliptic curve $E/\mathbb{Q}$ containing rational points $P = (0, 0)$ and Q such that the coefficients a_i of E are p-integers, P and Q are non-singular modulo p, and for $r \in \mathbb{N}$,

$$s_n \equiv 0 \bmod p^r \;\Leftrightarrow\; Q + [n]P \equiv \mathcal{O}_{p^r} \bmod p^r.$$

The x-coordinate of $Q + [n]P$ is

$$x_n = -\theta \cdot \frac{s_{n-1}\, s_{n+1}}{s_n^2} \qquad \text{whenever } n-1, n, n+1 \in I.$$

Proof: We first note that by Theorem 6.6.7 there are at least three terms coprime to p between every two multiples of p in S. So we may replace S by a translate in which $p \nmid s_{-1}\, s_0\, s_1$ and $s_2 \neq 0$.

Now let S' be the equivalent Somos 4 sequence given by

$$s'_n = \theta^{\frac{1}{2}n(n+1)}\, s_n \qquad \text{for all } n \in I.$$

(So since $p \nmid \theta$, s'_n and s_n are divisible by the same power of p for all $n \in I$.) By Theorem 7.2.4 there exists an elliptic curve E over $\mathbb{Q}$ containing non-singular rational points $P = (0, 0)$ and $Q = (x_0, y_0)$ such that $S' \in S_{E,Q,P}$. In particular the coefficients of E are given by

$$\begin{aligned}
a_4 &\text{ arbitrary} \\
a_3 &= \pm\sqrt{\lambda_1} \\
a_1 &= \frac{1}{a_3}\left(-2a_4 + \frac{s'_{-1}\, s'_2}{s'_0\, s'_1} + \frac{\lambda_1\, {s'_0}^2}{s'_{-1}\, s'_1} + \frac{s'_{-2}\, s'_1}{s'_{-1}\, s'_0}\right) \\
a_2 &= -\frac{({a_4}^2 + \lambda_2)}{\lambda_1} + \frac{a_4\left({s'_{-1}}^2\, s'_2 + s'_{-2}\, {s'_1}^2 + \lambda_1\, {s'_0}^3\right)}{\lambda_1\, s'_{-1}\, s'_0\, s'_1},
\end{aligned}$$

and Q is given by

$$x_0 = -\frac{s'_{-1}\, s'_1}{{s'_0}^2}, \qquad y_0 = \frac{-a_4\, {s'_{-1}}^2\, s'_0\, s'_1 + {s'_{-1}}^2\, s'_2}{a_3\, {s'_0}^3}.$$

Choose a_4 to be a p-integer. Then since the s'_n and λ_1 are p-integers and $p \nmid \lambda_1\, s'_{-1}\, s'_0\, s'_1$, the a_i are p-integers, and for any $r \in \mathbb{N}$ we may consider the reduced curve $E \bmod p^r$. Furthermore, since p does not divide $a_3 = \pm\sqrt{\lambda_1}$, $P \bmod p = (0,0)$ is a non-singular point. We now prove that $Q \bmod p$ is non-singular. By Theorem 3.4.2, if $Q \bmod p$ is a singular point then x_0 and y_0 satisfy

$$2a_3\, y_0 + a_1\, a_3\, x_0 + a_3^2 \equiv 0 \bmod p$$

and

$$2{x_0}^3 - b_4\, x_0 - a_3^2 \equiv 0 \bmod p.$$

Substituting $a_3^2 = \lambda_1$, $x_0 = -\frac{s'_{-1}\, s'_1}{{s'_0}^2}$ and $a_3\, y_0 = \frac{-a_4\, {s'_{-1}}^2\, s'_0\, s'_1 + {s'_{-1}}^2\, s'_2}{{s'_0}^3}$, multiplying both equations by ${s'_0}^3$ and using the fact that $b_4 = 2a_4 + a_1\, a_3$, we get two conditions on S':

$$2\, {s'_{-1}}^2\, s'_2 - b_4\, s'_{-1}\, s'_0\, s'_1 + \lambda_1\, {s'_0}^3 \equiv 0 \bmod p$$

and

$$2\left(-\frac{s'_{-1}\, s'_1}{s'_0}\right)^3 + b_4\, s'_{-1}\, s'_0\, s'_1 - \lambda_1\, {s'_0}^3 \equiv 0 \bmod p.$$

Substituting $b_4 = \frac{s'_{-1}\, s'_2}{s'_0\, s'_1} + \frac{\lambda_1\, {s'_0}^2}{s'_{-1}\, s'_1} + \frac{s'_{-2}\, s'_1}{s'_{-1}\, s'_0}$ and cancelling terms, we get

$${s'_{-1}}^2\, s'_2 - s'_{-2}\, {s'_1}^2 \equiv 0 \bmod p$$

and

$$-2\left(\frac{{s'_{-1}}^3\, {s'_1}^3}{{s'_0}^3}\right) + {s'_{-1}}^2\, s'_2 + s'_{-2}\, {s'_1}^2 \equiv 0 \bmod p,$$

from which

$$\frac{s'_0\, s'_2}{{s'_1}^2} \equiv \frac{s'_{-1}\, s'_1}{{s'_0}^2} \equiv \frac{s'_0\, s'_{-2}}{{s'_{-1}}^2} \bmod p.$$

But by Theorem 6.6.2, this is true if and only if S' and hence S is equivalent to a sequence all of whose terms are congruent to 1 modulo p, contrary to our assumption. Hence $Q \bmod p$ must be non-singular.

By Theorem 7.2.4, if we denote $Q+[n]P$ by (x_n, y_n) whenever $Q+[n]P \neq \mathcal{O}$ in $E(\mathbb{Q})$, we have

$$x_n = -\theta \cdot \frac{s_{n-1}\, s_{n+1}}{s_n^2} \quad \text{whenever } n-1, n, n+1 \in I,$$

and

$$s_n = 0 \;\Leftrightarrow\; Q + [n]P = \mathcal{O} \text{ in } E(\mathbb{Q}).$$

Since p does not divide θ or any two consecutive terms of S, it follows from Theorem 3.7.3 that $s_n \equiv 0 \bmod p^r$ if and only if the weighted Z-coordinate of $Q + [n]P$ is divisible by p^r, i.e., if and only if $Q + [n]P \equiv \mathcal{O}_{p^r} \bmod p^r$. □

Note that if S is equivalent to a sequence S' which is constant (and non-zero) modulo p (say $s_n = \prod_{i=1}^{k} \theta^{a_i n^2 + b_i n + c_i}\, s'_n$ for all $n \in I$), then either p does not appear in S (if $p \nmid \prod_{i=1}^{k} \theta_i$) or at most two terms of S are coprime to p (if p divides some θ_i).

7.6.1 The pattern of zeroes in $(s_n \bmod p^r)$

We can now show that almost all prime powers are regular in S.

Theorem 7.6.2. *Let S be a Somos 4 sequence with non-zero initial values, and let p be a reasonable prime in S, not dividing λ_1 or λ_2. Then every power of p which divides some term of S is regular in S.*

Proof: First suppose p divides some term but not all terms of S. Then by Theorem 7.6.1, for every $r \in \mathbb{N}$ there exists an elliptic curve E over $\mathbb{Z}_{p^r}$ containing points $P = (0, 0)$ and Q such that $P \bmod p = (0, 0)$ and $Q \bmod p$ are non-singular in $E \bmod p$ and

$$s_n \equiv 0 \bmod p^r \;\Leftrightarrow\; Q + [n]P = \mathcal{O}_{p^r} \text{ in } E(\mathbb{Z}_{p^r}).$$

It follows that p^r is regular in S with gap N_r equal to the order of the point $(0, 0)$ in $E(\mathbb{Z}_{p^r})$.

If p divides all terms of S, then let S' be the sequence obtained from S by dividing each term by p repeatedly until some term is coprime to p. Then since p^r is regular in S', it must be regular in S. □

Since Somos(4) is an integer sequence with $\lambda_1 = \lambda_2 = 1$, it follows that every prime power dividing some term of Somos(4) is regular. This proves Robinson's conjecture 6.7.2. Furthermore, since Somos(4) has $s_0 = s_1 = s_2 = s_3 = 1$, every regular prime has gap ≥ 5 in Somos(4), and so we have

Corollary 7.6.3. *Every five consecutive terms of Somos*(4) *are pairwise coprime.*

Theorem 7.6.2 together with Theorem 6.6.7 shows that there are few irregular primes in a Somos 4 sequence, and they are easy to find.

Theorem 7.6.4. *Let S be a Somos* 4 *sequence with non-zero initial values, and let p be a reasonable prime in S, dividing at least two terms of S. Then p is irregular if and only if p divides some two consecutive terms but not all terms of S, and in this case p also divides λ_2.*

We have not yet been able to prove that if p is a regular prime dividing $\lambda_1\,\lambda_2$ (i.e., with gap ≤ 3) and p^r divides some term of S then p^r is regular, though it seems likely that this is the case.

Hasse's Theorem for elliptic curves gives us a bound on the gap of p:

Theorem 7.6.5. *Let S be a Somos* 4 *sequence and let p be a regular prime with gap N_1 in S. Then*

$$N_1 \leq (p+1) + 2\sqrt{p}.$$

Proof: If p is a regular prime with gap $N_1 \geq 4$ in a Somos 4 sequence S, then $p \nmid \lambda_1\,\lambda_2$, and so by Theorem 7.6.1 there exists an elliptic curve E over $\mathbb{Z}_p$ containing points $P = (0,0)$ and Q such that $P \bmod p = (0,0)$ and $Q \bmod p$ are non-singular in $E \bmod p$ and

$$s_n \equiv 0 \bmod p \;\Leftrightarrow\; Q + [n]P = \mathcal{O}_p \text{ in } E(\mathbb{Z}_p).$$

So the gap N_1 of p in S is equal to the order of $P \bmod p$ in $E(\mathbb{Z}_p)$. Hasse's Theorem 3.6.1 applied to the elliptic curve $E/\mathbb{Z}_p$ shows that N_1 satisfies the Hasse bound $N_1 \leq (p+1) + 2\sqrt{p}$. Since every prime with gap $N_1 \leq 4$ in S also satisfies this bound, the result follows. □

This proves Robinson's Conjecture 6.7.3 for Somos(4). It also follows that no prime p can divide just one term of a Somos 4 sequence S.

Theorem 7.6.6. *Let S be a Somos 4 sequence, and let p be a reasonable odd prime coprime to $\lambda_1 \lambda_2$ and dividing some term but not all terms of (h_n).*

Then either all multiples of p in (h_n) are divisible by exactly the same power of p, or for all $r \in \mathbb{N}$ some term of (h_n) is divisible by p^r.

Proof: By Theorem 7.6.1, for every $r \in \mathbb{N}$ there exists an elliptic curve E over $\mathbb{Z}_{p^r}$ containing points $P = (0,0)$ and Q such that $P \bmod p = (0,0)$ and $Q \bmod p$ are non-singular in $E \bmod p$ and

$$s_n \equiv 0 \bmod p^r \;\Leftrightarrow\; Q + [n]P = \mathcal{O}_{p^r} \text{ in } E(\mathbb{Z}_{p^r}).$$

For $r \in \mathbb{N}$ let the order of the point $P = (0,0)$ in $E(\mathbb{Z}_{p^r})$ be N_r; *if* p^r appears as a factor in S then it has gap N_r.

Now let p^w be the highest power of p which divides every multiple of p in S, and let $p^{\bar{w}}$ be the highest power of p dividing $\psi_{N_1}(0,0)$. Then $Q \bmod p^w \in \langle P \bmod p^w \rangle$ in $E(\mathbb{Z}_{p^w})$.

If some term of S is divisible by p^{w+1} then $Q \bmod p^{w+1} \in \langle P \bmod p^{w+1} \rangle$ in $E(\mathbb{Z}_{p^{w+1}})$. Since $N_w = N_1$ and $N_{w+1} > N_1$ we must have $w = \bar{w}$. By Theorem 3.9.6 $Q \bmod p^r \in \langle P \bmod p^r \rangle$ in $E(\mathbb{Z}_{p^r})$ for all $r > \bar{w}$, so p^r appears in S for all $r > w$ and hence for all $r \in \mathbb{N}$.

Hence either no term of S is divisible by p^{w+1} (in which case $w < \bar{w}$), or for every $r \in \mathbb{N}$, some term of S is divisible by p^r. □

In Somos(4), every fifth term is divisible by 2 but no term is divisible by 4. It remains open whether $p = 2$ is the only prime for which this happens in Somos(4) (as conjectured by Robinson in Conjecture 6.7.1), but the following example gives another Somos 4 sequence S in which it does happen for odd p.

Example: Let S be the Somos 4 sequence with coefficients $\lambda_1 = 1$ and $\lambda_2 = -26$ and initial values $1, 1, -2, 7$, and let $p = 5$.

Then (s_n mod 25) is

$$\ldots, 1, 1, -2, 7, 3, 20, 22, 13, 17, 19, 14, 4, 15, 1, 11, 16, 8, 7, 18, 20, \ldots$$

So every seventh term of S is divisible by 5, but no term is divisible by 25.

An expression for the gap N_r of p^r in S in terms of N_1 now follows from Theorem 3.9.5 on the order of P mod p^r in $E(\mathbb{Z}_{p^r})$:

Theorem 7.6.7. *Let S be a Somos 4 sequence containing at most one zero term, and p a regular prime in S with gap $N_1 \geq 4$. Let w be the highest power of p such that all multiples of p in S are divisible by p^w, and suppose some term of S is divisible by p^{w+1}.*

For all $r \in \mathbb{N}$, p^r is regular in S; let the gap of p^r be N_r. Then if p is odd, or $p = 2$ and $w \geq 2$,

$$N_r = \begin{cases} N_1 & \text{if } r \leq w \\ p\, N_{r-1} & \text{if } r > w. \end{cases}$$

Otherwise, if $p = 2$ and $w = 1$ then for some $v \geq 2$,

$$N_r = \begin{cases} N_1 & \text{if } r = 1, \\ 2\, N_1 & \text{if } 2 \leq r \leq v, \\ 2\, N_{r-1} & \text{if } r > v. \end{cases}$$

Theorem 7.6.7 cannot be improved for $p = 2$, as shown by the example following Theorem 5.3.1 for EDSs. If $w = 1$ then v can be arbitrarily large.

Since every regular prime in Somos(4) has gap ≥ 5, this proves Robinson's Conjecture 6.7.4.

Remark: Note that if $r \leq w$ and $p^r \mid s_k$ then s_{k+N_1} may be divisible by a power of p greater than p^w, so we cannot define w by $p^w \parallel s_{k+N_1}$ as we did for EDSs in chapter 4.

Chapter 8

Primes appearing in a Somos 4 sequence

In this chapter let (h_n) be a Somos 4 sequence with coefficients λ_1, λ_2 and let p be a regular prime with gap $N_1 \geq 5$ in (h_n). Let $r \in \mathbb{N}$ such that p^r divides some term h_k, where if h_k is zero we assume it is not in the initial values of (h_n).

We prove that there exists a Somos 4 sequence (ℓ_n) which is equivalent to (h_n) and congruent modulo p^r to the $(-k)$–translate of an elliptic divisibility sequence (Z_n) (or to a segment of (Z_n) if (h_n) terminates). This will allow us to convert our results on the symmetry and periodicity of EDSs from chapter 5 to results on Somos 4 sequences. We also find a more general equation satisfied by the sequence $(h_n \bmod p^r)$, of which the Somos 4 recursion is a special case.

8.1 Finding an equivalent Somos 4 sequence congruent to an EDS modulo p^r

We define two rational constants

$$\xi = \frac{1}{h_{k+1}} \quad \text{and} \quad \mu = -\frac{h_{k+1}}{h_{k-1}},$$

and let (ℓ_n) be the equivalent Somos 4 sequence given by

$$\ell_n = \xi\, \mu^{\frac{1}{2}(n-k)(n-k-1)}\, h_n \quad \text{for all } n \in I.$$

The coefficients of (ℓ_n) are

$$\lambda'_1 = \mu^3 \lambda_1 \quad \text{and} \quad \lambda'_2 = \mu^4 \lambda_2.$$

(These are coprime to p by Theorem 6.6.7.) Note that (since $\xi \mu \neq 0$) ℓ_n is non-zero if and only if h_n is, so the ℓ_n are defined for indices n in the same set I as the h_n. Also note that μ and ξ depend on k, i.e., on which of the terms divisible by p^r we choose for h_k.

We have

$$\begin{aligned}
\ell_{k-1} &= \xi \, \mu^{\frac{1}{2}(-1)(-2)} \, h_{k-1} = \left(\frac{1}{h_{k+1}}\right) \left(-\frac{h_{k+1}}{h_{k-1}}\right) h_{k-1} = -1, \\
\ell_k &= \xi \, \mu^{\frac{1}{2}(0)(-1)} \, h_k = \frac{h_k}{h_{k+1}}, \qquad \text{and} \\
\ell_{k+1} &= \xi \, \mu^{\frac{1}{2}(1)(0)} \, h_{k+1} = \left(\frac{1}{h_{k+1}}\right) h_{k+1} = 1.
\end{aligned}$$

Furthermore, since ξ and μ are coprime to p, each term ℓ_n is divisible by the same power of p as the corresponding term h_n. So in particular, $\ell_k \equiv 0 \bmod p^r$ and $\ell_{k+j} \not\equiv 0 \bmod p$ for $1 \leq |j| \leq 4$.

We also have

Lemma 8.1.1. *For $j = 1, 2, 3, 4$,*

$$\ell_{k-j} \equiv -\ell_{k+j} \bmod p^r$$

(if ℓ_{k-j} and ℓ_{k+j} both exist i.e., unless $h_k = 0$ and $j = 4$).

Proof: If $k - 4 \in I$, let (b_n) be the $(k-4)$–translate of (ℓ_n), i.e., let

$$b_n = \ell_{k-4+n} \quad \text{whenever } k - 4 + n \in I.$$

Then $b_0, b_1, \ldots, b_7$ are defined, and b_8 is defined unless $k + 4 \notin I$. Of these terms, only b_4 is divisible by p. We have $b_4 = \ell_k \equiv 0 \bmod p^r$, $b_3 = -1$ and $b_5 = 1$. It follows from Theorem 6.2.2 that

$$\begin{aligned}
b_6 &\equiv \frac{1}{b_1 \, b_2} \left(\lambda'_1 \, \lambda'_2 \, b_3^3\right) \equiv -\frac{\lambda'_1 \, \lambda'_2}{b_1 \, b_2} \bmod p^r, \\
b_7 &\equiv \frac{1}{b_1^2 \, b_2 \, b_3} \left({\lambda'_2}^3 \, b_2 \, b_3^4\right) \equiv -\frac{{\lambda'_2}^3}{b_1^2} \bmod p^r,
\end{aligned}$$

and if b_8 is defined,

$$b_8 \equiv \frac{1}{b_1^3\, b_2^2\, b_3}\left(\lambda_1'\, {\lambda_2'}^3\, b_0\, b_3^6\right) \equiv -\frac{\lambda_1'\, {\lambda_2'}^3\, b_0}{b_1^3\, b_2^2} \bmod p^r.$$

But by Theorem 6.2.1 with $k = 2$ we have

$$\lambda_1' \equiv -\frac{b_2^2\, b_5\, b_1}{b_3^3\, b_1} \equiv b_2^2 \bmod p^r \quad \text{and} \quad \lambda_2' \equiv \frac{b_3\, b_1^2\, b_5}{b_3^3\, b_1} \equiv b_1 \bmod p^r.$$

Hence

$$b_6 \equiv -b_2 \bmod p^r, \quad b_7 \equiv -b_1 \bmod p^r, \quad \text{and} \quad b_8 \equiv -b_0 \bmod p^r$$

if b_8 is defined, and the result follows. If $k - 4 \notin I$ (i.e., $h_k = 0$ and $k < 0$), then we obtain a similar proof by letting (b_n) be the $(k+4)$–translate of the reversed sequence (ℓ_{-n}), i.e., letting

$$b_n = \ell_{k+4-n} \quad \text{whenever} \quad k + 4 - n \in I.$$

□

Now let (Z_n) be any elliptic divisibility sequence such that

$$Z_j \equiv \ell_{k+j} \bmod p^r \quad \text{for} \quad |j| \le 4$$

(as long as ℓ_{k+j} exists), Z_3 and Z_4 are coprime, and $Z_5 \neq 0$. (So by Theorem 4.6.4, (Z_n) has no zero terms other than Z_0.) We know such a sequence (Z_n) exists, because we can construct one in the following way (among other ways):

1. Set $Z_{-1} = -1,\ Z_0 = 0,\ Z_1 = 1$. So $Z_j \equiv \ell_{k+j} \bmod p^r$ for $|j| \le 1$.

2. Set $Z_2 = \ell_{k+2} \bmod p^r$, $\chi = \ell_{k+4}\, \ell_{k+2}^{-1} \bmod p^r$ and $Z_4 = Z_2\, \chi$. So $Z_j \equiv \ell_{k+j} \bmod p^r$ for $j = 2, 4$ and Z_2 divides Z_4.

3. Use the Chinese Remainder Theorem to find an integer Z_3 that is congruent to ℓ_{k+3} modulo p^r and coprime to Z_4. (We can do this because $p \nmid \ell_{k+4}$ and hence $p \nmid Z_4$. Any of the $\phi(Z_4)$ possibilities for $Z_3 \bmod Z_4$ will do.)

4. Finally, check that $Z_5 \neq 0$. If $Z_5 = 0$ then, since $Z_5 Z_1 = Z_4 Z_2^3 - Z_1 Z_3^3$ and Z_3 is coprime to Z_2 and Z_4, the terms Z_2, Z_3 and Z_4 must all be ± 1. In this case, multiply Z_3 by any integer (except 1) that is congruent to 1 modulo p^r and coprime to Z_4 to make $Z_5 \neq 0$.

It follows from Lemma 8.1.1 and the fact that $Z_{-n} = -Z_n$ for all $n \in \mathbb{Z}$ that $Z_j \equiv \ell_{k+j}$ for $j = -2, -3, -4$ too, as long as ℓ_{k+j} exists.

Theorem 8.1.2. *The Somos 4 sequence (ℓ_n) is congruent to a $(-k)$–translate of the EDS (Z_n) modulo p^r. That is,*

$$\ell_{k+n} \equiv Z_n \bmod p^r \qquad \textit{whenever} \quad k+n \in I.$$

Proof: By Theorem 6.6.1 and Lemma 8.1.1, since $\ell_k \equiv 0 \bmod p^r$, (ℓ_n) has coefficients

$$\lambda_1' \equiv \frac{\ell_{k+2}\,\ell_{k-2}}{\ell_{k+1}\,\ell_{k-1}} \equiv \ell_{k+2}^2 \bmod p^r \qquad \text{and} \qquad \lambda_2' \equiv \frac{\ell_{k+3}\,\ell_{k-1}}{\ell_{k+1}^2} \equiv -\ell_{k+3} \bmod p^r.$$

The sequence $Z_{-3}, Z_{-2}, Z_{-1}, Z_0, Z_1, \ldots$ is also a Somos 4 sequence, with coefficients

$$Z_2^2 \equiv \ell_{k+2}^2 \equiv \lambda_1' \bmod p^r \qquad \text{and} \qquad -Z_1\,Z_3 \equiv -\ell_{k+3} \equiv \lambda_2' \bmod p^r$$

and initial values

$$Z_j \equiv \ell_{k+j} \bmod p^r \qquad \text{for } j = 1, \ldots, 4.$$

The sequence (Z_n) has no zero terms other than Z_0, so it never terminates to the right of Z_0. It follows by Theorem 6.6.8 that

$$\ell_{k+n} \equiv Z_n \bmod p^r \qquad \text{for} \quad n = -3, -2, \ldots$$

till the sequence (ℓ_n) terminates to the right (if it does), i.e., for all $n \in I \cap \{-3, -2, \ldots\}$.

This establishes the result for $n \geq -3$. The proof for $n \leq -4$ (i.e., for $n \in I \cap \{\ldots, -5, -4\}$) is similar, using the fact that $Z_j \equiv \ell_{k+j} \bmod p^r$ for $j = -1, -2, -3, -4$. □

We illustrate the above process with an example:

Example: Let (h_n) be the Somos 4 sequence with initial values $h_0 = 2,\ h_1 = 3,\ h_2 = 5,\ h_3 = 7$ and coefficients $\lambda_1 = 1$ and $\lambda_2 = 2$, i.e.,

$$\ldots,\ \frac{128}{3},\ \frac{40}{3},\ 4,\ 4,\ 2,\ 3,\ 5,\ 7,\ \frac{71}{2},\ \frac{551}{6},\ \frac{1898}{3},\ \ldots$$

and let $p^r = 5$ and $k = -3$. So $\xi = \frac{1}{h_3} = \frac{1}{7} \equiv 3 \bmod 5$ and $\mu = -\frac{h_3}{h_1} = -\frac{7}{3} \equiv 1 \bmod 5$. Then (ℓ_n) is the equivalent Somos 4 sequence given by $\ell_n = \xi\, \mu^{\frac{1}{2}(n-k)(n-k-1)}\, h_n$ for all $n \in I$, i.e.,

$$\begin{aligned}
\ell_{k-2} &= \xi\, \mu^3\, h_{k-2} = \frac{1}{7} \cdot \left(-\frac{7}{3}\right)^3 \cdot 2 \equiv 1 \bmod 5 \\
\ell_{k-1} &= \xi\, \mu\, h_{k-1} = \frac{1}{7} \cdot \left(-\frac{7}{3}\right) \cdot 3 = -1 \\
\ell_k &= \xi\, h_k = \frac{5}{7} \equiv 0 \bmod 5 \\
\ell_{k+1} &= \xi\, h_{k+1} = 1 \\
\ell_{k+2} &= \xi\, \mu\, h_{k+2} = \frac{1}{7} \cdot \left(-\frac{7}{3}\right) \cdot \frac{71}{2} \equiv -1 \bmod 5 \\
\ell_{k+3} &= \xi\, \mu^3\, h_{k-1} = \frac{1}{7} \cdot \left(-\frac{7}{3}\right)^3 \cdot \frac{551}{6} \equiv 3 \bmod 5 \\
\ell_{k+4} &= \xi\, \mu^6\, h_{k-1} = \frac{1}{7} \cdot \left(-\frac{7}{3}\right)^6 \cdot \frac{551}{6} \equiv 3 \bmod 5.
\end{aligned}$$

The coefficients of (ℓ_n) are $\lambda_1' = \mu^3\, \lambda_1 = \left(-\frac{7}{3}\right)^3$ and $\lambda_2' = \mu^4\, \lambda_2 = \left(-\frac{7}{3}\right)^4 \cdot 2$.

A possibility for (Z_n) is the EDS with initial values $Z_2 = -1$, $Z_3 = 3$ and $Z_4 = -2$, i.e.,

$$\ldots, 0, 1, -1, 3, -2, -25, 87, -683 \ldots$$

Remarks:

1. This technique does not work if p^r does not divide some term of (h_n), because every EDS (Z_n) has $Z_0 = 0$. So if (h_n) has no term divisible by p^r then we cannot find an equivalent sequence congruent to a translate of an EDS.

2. If p were a regular prime with gap $N_1 \leq 4$ then we could still have found an EDS (Z_n) whose initial values were congruent to ℓ_{k+1}, ℓ_{k+2}, ℓ_{k+3}, ℓ_{k+4} modulo p^r. However, we could not then conclude that $Z_n \equiv \ell_{k+n} \bmod p^r$ for *all* $n \in I$, since to use Theorem 6.6.8 we need four consecutive terms of the two sequences to be congruent modulo p^r *and* coprime to p.

   Theorem 8.1.2 means that

$$Z_n \equiv \ell_{k+n} = \xi\, \mu^{\frac{1}{2}n(n-1)}\, h_{k+n} \bmod p^r \quad \text{whenever} \quad n+k \in I. \tag{8.1}$$

In the remainder of this chapter we use this relationship to prove results on the reduced Somos 4 sequence $(h_n \bmod p^r)$ from known results about the elliptic divisibility sequence (Z_n).

Theorem 8.1.3. *Let (h_n) be a Somos 4 sequence defined over a set of indices I, and p a regular prime with gap $N_1 \geq 5$ in (h_n). Let $r \in \mathbb{N}$, and let h_k be a term divisible by p^r. Then for all n with $k-n, k+n \in I$,*

$$h_{k+n} \equiv -\mu^n\, h_{k-n} \bmod p^r.$$

Proof: Since (Z_n) is an EDS, $Z_{-n} = -Z_n$ for all $n \in \mathbb{Z}$. If we replace Z_t by $\xi\, \mu^{\frac{1}{2}t(t-1)}\, h_{k+t}$ modulo p^r for all t with $k+t \in I$, this becomes

$$\xi\, \mu^{\frac{1}{2}(-n)(-n-1)}\, h_{k-n} \equiv -\xi\, \mu^{\frac{1}{2}n(n-1)}\, h_{k+n} \bmod p^r,$$

or

$$h_{k-n} \equiv -\mu^{\frac{1}{2}(n^2-n)-\frac{1}{2}(n^2+n)}\, h_{k+n} \equiv -\mu^{-n}\, h_{k+n} \bmod p^r,$$

as required. □

8.2 The pattern of zeroes in $(h_n \bmod p^r)$

It follows from (8.1) (since $p \nmid \xi\, \mu$) that for every $1 \leq j \leq r$,

$$h_n \equiv 0 \bmod p^j \;\Leftrightarrow\; Z_{n-k} \equiv 0 \bmod p^j,$$

i.e., the pattern of zeroes in $(h_n \bmod p^j)$ is exactly the same as in $(Z_n \bmod p^j)$, shifted by k places.

We could have used this to prove Theorems 7.6.2, 7.6.4, 7.6.5, and 7.6.7 in the case where $N_1 \geq 5$, instead of using the connection between Somos 4 sequences and elliptic curves directly as we did in chapter 7. However, if we did this we would still be using elliptic curves indirectly, in the existence proof for the EDS (Z_n) (Theorem 4.4.8).

In fact, an obvious alternative choice for the elliptic sequence (Z_n) is the sequence $(\psi_n(0,0))$, where ψ_n is the nth division polynomial of some elliptic curve E with $(h_n) \in S_{E,Q,P}$ for points $P = (0,0)$ and $Q = [-k]P$.

8.3 A global recursion satisfied by $(h_n \bmod p^r)$

Theorem 8.1.2 implies that, when considered modulo p^r, (h_n) actually satisfies a global recursion "centred" at h_k, not merely the local recursion (6.2). (By a recursion being "global" we mean that it is a relationship between terms which are far apart in the sequence.)

Theorem 8.3.1. *Let (h_n) be a Somos 4 sequence and let p be a regular prime with gap $N_1 \geq 5$ such that p^r divides some term h_k. Then*

$$\mu^n\, h_{k+m+n}\, h_{k+m-n}\, h_{k+t}^2 + \mu^t\, h_{k+n+t}\, h_{k+n-t}\, h_{k+m}^2 + \mu^m\, h_{k+t+m}\, h_{k+t-m}\, h_{k+n}^2 \equiv 0 \bmod p^r$$

whenever all these indices are in I.

Proof: By the symmetric form of the generalised elliptic sequence equation (4.2b),

$$Z_{m+n}\, Z_{m-n}\, Z_t^2 + Z_{n+t}\, Z_{n-t}\, Z_m^2 + Z_{t+m}\, Z_{t-m}\, Z_n^2 = 0$$

for all $m, n, t \in \mathbb{Z}$. The result follows on substituting $Z_j \equiv \xi\, \mu^{\frac{1}{2}j(j-1)}\, h_{k+j} \bmod p^r$ whenever $k + j \in I$ and simplifying. □

Fixing n and t and replacing m by $m - k$, we get

Theorem 8.3.2. *Let (h_n) be a Somos 4 sequence and let p be a regular prime with gap $N_1 \geq 5$ such that p^r divides some term h_k. Let $n, t \in \mathbb{N}$ with $t \leq n$ such that $k+n+t, \ldots, k-n \in I$, and define constants $A_{k,n,t}$ and $B_{k,n,t}$ by*

$$A_{k,n,t} = \frac{h_{k+n}\, h_{k-n}}{h_{k+t}\, h_{k-t}} \bmod p^r$$

and

$$B_{k,n,t} = \frac{h_{k+t-n}\, h_{k+t+n}}{h_{k+t}^2} \bmod p^r.$$

Then the sequence $(h_n \bmod p^r)$ satisfies the degree $2n$ recursion

$$h_{m+n}\, h_{m-n} \equiv A_{k,n,t}\, h_{m+t}\, h_{m-t} + B_{k,n,t}\, h_m^2 \bmod p^r \tag{8.2}$$

whenever $m+n, \ldots, m-n \in I$. In other words, for every $n \geq 2$, $(h_n \bmod p^r)$ satisfies $n-1$ Somos $2n$ recursions (for $t = 1, 2, \ldots, n-1$), in each of which at most two of the coefficients are non-zero.

Proof: Replacing m by $m-k$ in Theorem 8.3.1 we get

$$\mu^n\, h_{m+n}\, h_{m-n}\, h_{k+t}^2 + \mu^t\, h_{k+n+t}\, h_{k+n-t}\, h_m^2 + \mu^{m-k}\, h_{t+m}\, h_{k+t-(m-k)}\, h_{k+n}^2 \equiv 0 \bmod p^r.$$

But by Theorem 8.1.3, we have $h_{k+n-t} \equiv -\mu^{n-t}\, h_{k+t-n}$ and $h_{k+t-(m-k)} \equiv -\mu^{t-(m-k)}\, h_{k-t+(m-k)} \equiv -\mu^{t-(m-k)}\, h_{m-t} \bmod p^r$, so

$$\mu^n\, h_{m+n}\, h_{m-n}\, h_{k+t}^2 - \mu^n\, h_{k+t+n}\, h_{k+t-n}\, h_m^2 - \mu^t\, h_{m+t}\, h_{m-t}\, h_{k+n}^2 \equiv 0 \bmod p^r.$$

This becomes, on dividing by $\mu^n\, h_{k+t}^2$,

$$h_{m+n}\, h_{m-n} \equiv \left(\frac{h_{k+t+n}\, h_{k+t-n}}{h_{k+t}^2}\right) h_m^2 + \left(\frac{h_{k+n} \cdot \mu^{-n}\, h_{k+n}}{h_{k+t} \cdot \mu^{-t}\, h_{k+t}}\right) h_{m+t}\, h_{m-t} \bmod p^r.$$

Again using Theorem 8.1.3, we get

$$\begin{aligned} h_{m+n}\, h_{m-n} &\equiv \left(\frac{h_{k+t-n}\, h_{k+t+n}}{h_{k+t}^2}\right) h_m^2 + \left(\frac{h_{k+n}\, h_{k-n}}{h_{k+t}\, h_{k-t}}\right) h_{m+t}\, h_{m-t} \bmod p^r \\ &\equiv B_{k,n,t}\, h_m^2 + A_{k,n,t}\, h_{m+t}\, h_{m-t} \bmod p^r. \end{aligned}$$

□

Remarks:

1. It remains an open problem whether for given $n \geq 2$ there is a Somos $2n$ recursion satisfied by (h_n). We proved this for Somos 4 sequences containing a zero term in Theorem 6.4.7, and some other special cases are known to hold. For example, it was observed by Michael Somos (and referred to by Jim Propp on his Somos sequence newsgroup [19]) that Somos(4) satisfies

$$h_{m+3}\, h_{m-3} = h_{m-1}\, h_{m+1} + 5\, h_m^2 \quad \text{for all } m \in \mathbb{Z}$$

   i.e., that Somos(4) is also a Somos 6 sequence. We can in fact prove using elliptic curves that all Somos 4 sequences with non-zero initial values and $\lambda_1 \neq 0$ satisfy

$$h_{m+3}\, h_{m-3} = \lambda_2^2\, h_{m-1}\, h_{m+1} + B_3\, h_m^2 \quad \text{for all } m \in \mathbb{Z}$$

   for some coefficient B_3 which we can find in terms of the initial values. The coefficients λ_2^2 and B_3 reduce modulo p^r to the above values $A_{k,3,1}$ and $B_{k,3,1}$.

2. It is easy to prove using the symmetry formula for $(h_n \bmod p^r)$ (which we will prove in the next section) that the coefficients $A_{k,n,t}$ and $B_{k,n,t}$ are independent of which k we choose for $p^r \mid h_k$ — i.e., adding a multiple of N_r to k does not change them.

3. Setting $n = 2$ and $t = 1$ in Theorem 8.3.2 and using Theorem 6.6.1 gives

$$A_{k,2,1} \equiv \frac{h_{k+2}\, h_{k-2}}{h_{k+1}\, h_{k-1}} \equiv \lambda_1 \bmod p^r$$

   and

$$B_{k,2,1} \equiv \frac{h_{k+3}\, h_{k-1}}{h_{k+1}^2} \equiv \lambda_2 \bmod p^r,$$

   as expected.

4. Theorem 8.3.2 also allows us to find a given term $h_k \bmod p^r$ in $\mathrm{O}(\log k)$ operations in $\mathbb{Z}_{p^r}$ if we know $h_0, \ldots, h_4$, using doubling formulae as we do for elliptic sequences.

The next result shows that taking every Mth term from a given term h_j gives a sequence (ℓ_s) which looks like a Somos 4 sequence when considered modulo p^r.

Theorem 8.3.3. *Let (h_n) be a Somos 4 sequence and let p be a regular prime with gap $N_1 \geq 5$ such that p^r divides some term h_k. Let M and j be integers with $M \not\equiv 0 \bmod N_1$, and let (ℓ_s) be the subsequence of (h_s) consisting of the Mth terms from h_j, i.e., let*

$$\ell_s = h_{j+sM} \qquad \text{whenever } \; j + sM \in I.$$

Then (ℓ_s) satisfies

$$\ell_{m+n}\,\ell_{m-n} \equiv A_{k,nM,tM}\;\ell_{m+t}\,\ell_{m-t} + B_{k,nM,tM}\;\ell_m^2 \bmod p^r$$

whenever all these terms are defined. In particular,

$$\ell_{m+2}\,\ell_{m-2} \equiv A\,\ell_{m+1}\,\ell_{m-1} + B\,\ell_m^2 \bmod p^r,$$

where $A = A_{k,2M,M} = \frac{h_{k+2M}\,h_{k-2M}}{h_{k+M}\,h_{k-M}} \bmod p^r$ and $B = B_{k,2M,M} = \frac{h_{k-M}\,h_{k+3M}}{h_{k+M}^2} \bmod p^r$.

Proof: The result follows on replacing m by $mM + j$ and n and t by nM and tM in Theorem 8.3.2, and then setting $n = 2$, $t = 1$. □

In Theorem 4.6.2 we proved that if (h_n) is an EDS and $\ell_s = h_{j+sM}$ for all s and some j and $M \not\equiv 0 \bmod N_1$, then (ℓ_s) is in fact a Somos 4 sequence. It remains an open problem whether this is true for every Somos 4 sequence (h_n), but it was true for the (few) examples I looked at. For example, taking every 4th term of Somos(4) gives a Somos 4 sequence with coefficients $\lambda_1' = 169$ and $\lambda_2' = 476557$ and initial values $1, 2, 59, 83313$.

8.4 Symmetry in $(h_n \bmod p^r)$

Next we prove some new results on the symmetry of the reduced Somos 4 sequence $(h_n \bmod p^r)$. The results in this section are obtained from the EDS results in chapter 5 simply by substituting

$$Z_t \equiv \xi\,\mu^{\frac{1}{2}t(t-1)}\,h_{k+t} \bmod p^r \qquad \text{for all } k + t \in I,$$

but it should be noted that we could also have used Theorem 8.3.1 to prove them from scratch, as we did for EDSs in chapter 5. (The proofs would then be generalisations of our proofs of the corresponding EDS results in chapter 5.)

Recall that (h_n) is a Somos 4 sequence, p is a regular prime with gap $N_1 \geq 5$ in (h_n) and $r \in \mathbb{N}$ such that p^r divides some term h_k. For $1 \leq j \leq r$, let the gap of p^j in (h_n) be N_j, and let p^w be the highest power of p dividing every multiple of p in (h_n). In this section and the next we assume that k has been chosen so that either $h_k \neq 0$ or $k < 0$, so that h_{k+N_r-1} and h_{k+N_r-2} are defined. (We can make this assumption without loss of generality because if necessary we can consider the reverse sequence defined by $h'_n = -h_n$ for all $n \in I$.)

Notice that, because all powers of p which divide some term are regular in (h_n), if $p^j \parallel h_t$ for some $j < r$ then $p^j \parallel h_{t+N_r}$ too, i.e., h_t and h_{t+N_r} are divisible by exactly the same power of p. If $p^r \mid h_t$, then $p^r \mid h_{t+sN_r}$ too, but one of h_t or h_{t+N_r} might be divisible by a higher power of p than p^r (although if $r \geq w$ they can't both be).

Definition: For $1 \leq j \leq r$, we define constants β_j and γ_j by

$$\gamma_j = \frac{h_{k+N_j-1}}{h_{k-1}} \cdot \frac{h_{k-2}}{h_{k+N_j-2}} \bmod p^j \qquad \text{and} \qquad \beta_j = -\gamma_j^{-k+1} \cdot \frac{h_{k+N_j-1}}{h_{k-1}} \bmod p^j.$$

Notice that h_{k+N_j-1}, h_{k+N_j-2}, h_{k-1} and h_{k-2} are all coprime to p (since p has gap $N_1 > 2$), so β_j and γ_j are defined and coprime to p.

The corresponding constants b_j and c_j for the EDS (Z_n) are

$$c_j = \frac{Z_{N_j-1}}{Z_{-1}} \cdot \frac{Z_{-2}}{Z_{N_j-2}} \bmod p^j \qquad \text{and} \qquad b_j = -c_j \cdot \frac{Z_{N_j-1}}{Z_{-1}} \bmod p^j.$$

Replacing Z_n by $\xi\, \mu^{\frac{1}{2}n(n-1)}\, h_{k+n}$ modulo p^r for all n and simplifying, we get

Theorem 8.4.1. *The constants β_j and γ_j for the Somos 4 sequence (h_n) are related to the constants b_j and c_j for the EDS (Z_n) by*

$$b_j \equiv \mu^{\frac{1}{2}N_j(N_j-1)}\, \gamma_j^k\, \beta_j \bmod p^j \qquad \textit{and} \qquad c_j \equiv \mu^{N_j}\, \gamma_j \bmod p^j.$$

Notice that if $k = 0$ and $h_{-1} \equiv -h_1 \bmod p^r$ then $\mu \equiv 1 \bmod p^r$ and we have $\gamma_r \equiv c_r \bmod p^r$ and $\beta_r \equiv b_r \bmod p^r$.

We can now obtain a symmetry formula for $(h_n \bmod p^r)$ from the symmetry formula for the reduced EDS $(Z_n \bmod p^r)$:

Theorem 8.4.2. *Let (h_n) be a Somos 4 sequence and p a regular prime with gap $N_1 \geq 5$ in (h_n). For $r \in \mathbb{N}$, let p^r divide some term h_k, and let p^r have gap N_r in (h_n). Then*

$$h_{t+sN_r} \equiv \left(\gamma_r^{N_r}\right)^{\frac{1}{2}s(s-1)} \gamma_r^{st} \left(-\beta_r\right)^s h_t \bmod p^r \tag{8.3}$$

for all integers s and t with $t, t + sN_r \in I$.

Proof: By Corollary 5.1.5 (replacing t by $t - k$)

$$Z_{t-k+sN_r} \equiv \left(c_r^{N_r}\right)^{\frac{1}{2}s(s-1)} c_r^{s(t-k)} \left(-b_r\right)^s Z_{t-k} \bmod p^r.$$

Replacing Z_n by $\xi\, \mu^{\frac{1}{2}n(n-1)}\, h_{k+n}$ modulo p^r for all n, and substituting $b_r \equiv \mu^{\frac{1}{2}N_r(N_r-1)}\, \gamma_r^k\, \beta_r \bmod p^r$ and $c_r \equiv \mu^{N_r}\gamma_r \bmod p^r$ from Theorem 8.4.1, we get

$$\begin{aligned}
&\xi\, \mu^{\frac{1}{2}(sN_r+t-k)(sN_r+t-k-1)}\, h_{sN_r+t} \\
&\quad \equiv \left(\mu^{N_r}\gamma_r\right)^{\frac{1}{2}N_r s(s-1)+s(t-k)} \left(-\mu^{\frac{1}{2}N_r(N_r-1)}\gamma_r^k\beta_r\right)^s \left(\xi\, \mu^{\frac{1}{2}(t-k)(t-k-1)}\, h_t\right) \bmod p^r,
\end{aligned}$$

which simplifies to

$$h_{t+sN_r} \equiv \left(\gamma_r^{N_r}\right)^{\frac{1}{2}s(s-1)} \gamma_r^{st} \left(-\beta_r\right)^s h_t \bmod p^r.$$

□

Setting $s = 1$ and $t = 0$ in the above theorem we get

Corollary 8.4.3. *Let (h_n) be a Somos 4 sequence and p a regular prime with gap $N_1 \geq 5$ in (h_n). For $r \in \mathbb{N}$, let p^r divide some term h_k, and let p^r have gap N_r in (h_n). If $p \nmid h_0$, then*

$$\beta_r \equiv -\frac{h_{N_r}}{h_0} \bmod p^r.$$

This would perhaps have made a neater definition for β_r, but of course we do not necessarily have $p \nmid h_0$. In the next chapter we do, and we will be defining the corresponding quantity β_r in this way.

Setting first $s = 1$ and then $s = -1$ in Theorem 8.4.2 gives

Corollary 8.4.4. *The constant γ_r satisfies*

$$\frac{h_{t+N_r}\, h_{t-N_r}}{h_t^2} \equiv \gamma_r^{N_r} \bmod p^r \qquad \textit{for all } t \not\equiv k \bmod N_1.$$

Theorem 8.4.2 says that, if $h_t \not\equiv 0 \bmod p$ and we consider the sequence (ℓ_n) obtained from (h_n) by taking every N_rth term from h_t and dividing by h_t, then modulo p^r the sth term $\ell_s = \frac{h_{t+sN_r}}{h_t}$ has the simple form

$$\ell_s \equiv \left(\gamma_r^{N_r}\right)^{\frac{1}{2}s(s-1)}\, \gamma_r^{st}\, (-\beta_r)^s \bmod p^r$$

for all s with $t + sN_r \in I$. So for every choice of $t \not\equiv 0 \bmod N_1$,

$$\ell_1 \equiv \gamma_r^t\, (-\beta_r) \bmod p^r \qquad \text{and} \qquad \ell_{-1}\, \ell_1 \equiv \gamma_r^{N_r} \bmod p^r,$$

and we can write ℓ_s in the even more simple form

$$\ell_s \equiv \ell_1^s\, (\ell_{-1}\, \ell_1)^{\frac{1}{2}s(s-1)} \equiv \ell_1^{\frac{1}{2}s(s+1)}\, \ell_{-1}^{\frac{1}{2}s(s-1)} \bmod p^r$$

for all s with $t + sN_r \in I$. So modulo p^r, (ℓ_n) looks like (a segment of) a sequence equivalent to the constant sequence $\ldots, 1, 1, 1, 1, \ldots$

Substituting for b_r and c_r in terms of β_r and γ_r in Lemma 5.1.2 gives

Theorem 8.4.5. *The constants μ, β_r and γ_r are related by*

$$\mu^{N_r} \cdot \gamma_r^{N_r - 2k} \cdot \beta_r^{-2} \equiv 1 \bmod p^r.$$

In fact, if N_r is even and $p > 2$ then

$$\mu^{\frac{N_r}{2}} \cdot \gamma_r^{\frac{N_r}{2} - k} \cdot \beta_r^{-1} \equiv 1 \bmod p^r.$$

Of course k is not uniquely determined by p^r; since p^r is regular we could have chosen any $k' = k + aN_r$ in I instead. It turns out that this would have affected μ, but not β_r or γ_r:

Theorem 8.4.6. *If we replace k by $k' = k + aN_r$ for some integer a, then μ is replaced by $\mu' \equiv \gamma_r^{2a}\,\mu \bmod p^r$. The constants β_r and γ_r are independent of our choice of k.*

Proof: Using Theorem 8.4.2 we have

$$\begin{aligned}\mu' &= -\frac{h_{(k+aN_r)+1}}{h_{(k+aN_r)-1}} = -\frac{h_{(k+1)+aN_r}}{h_{(k-1)+aN_r}} \\ &\equiv \frac{\left(\gamma_r^{N_r}\right)^{\frac{1}{2}a(a-1)}\,(\gamma_r)^{a(k+1)}\,(-\beta_r)^a}{\left(\gamma_r^{N_r}\right)^{\frac{1}{2}a(a-1)}\,(\gamma_r)^{a(k-1)}\,(-\beta_r)^a} \cdot \left(-\frac{h_{k+1}}{h_{k-1}}\right) \bmod p^r \\ &\equiv \gamma_r^{2a}\,\mu \bmod p^r.\end{aligned}$$

So $\mu \bmod p^r$ is multiplied by a factor of γ_r^2 every time k increases by N_r.

Similarly,

$$\begin{aligned}\gamma_r' &\equiv \frac{h_{N_r+k'-1}}{h_{k'-1}} \cdot \frac{h_{k'-2}}{h_{N_r+k'-2}} \bmod p^r \\ &\equiv \frac{\gamma_r^{k'-1}\,(-\beta_r)}{\gamma_r^{k'-2}\,(-\beta_r)} \bmod p^r \\ &\equiv \gamma_r \bmod p^r,\end{aligned}$$

and

$$\begin{aligned}\beta_r' &\equiv -\gamma_r^{-k'+1} \cdot \frac{h_{N_r+(k'-1)}}{h_{k'-1}} \bmod p^r \\ &\equiv -\gamma_r^{-k'+1} \cdot \left(\gamma_r^{k'-1}\,(-\beta_r)\right) \bmod p^r \\ &\equiv \beta_r \bmod p^r.\end{aligned}$$

So the constants γ_r and β_r are independent of our choice of k. □

8.4.1 Extending the symmetry formula to higher powers of p

As in the EDS case, if p^{3j} divides some term of (h_n) then we can find a formula for $\frac{h_{t+sN_j}}{h_t}$ that holds modulo p^{3j}, not just modulo p^j.

Theorem 8.4.7. *Let (h_n) be a Somos 4 sequence and p a regular prime with gap $N_1 \geq 5$ in (h_n). Let $j \in \mathbb{N}$ such that p^{3j} divides some term of (h_n), and let p^j have gap N_j in (h_n). Then*

$$\frac{h_{t+sN_j}}{h_t} \equiv \left(\frac{h_{t+N_j}}{h_t}\right)^{\frac{1}{2}s(s+1)} \left(\frac{h_{t-N_j}}{h_t}\right)^{\frac{1}{2}s(s-1)} \bmod p^{3j} \tag{8.4}$$

for all integers s and t such that $h_t \not\equiv 0 \bmod p$ and all these indices are in I.

Proof: Set $r = 3j$. Since (Z_n) is an EDS, by Theorem 5.2.1 (replacing r by j and t by $t-k$) we have

$$\frac{Z_{t-k+sN_j}}{Z_{t-k}} \equiv \left(\frac{Z_{t-k+N_j}}{Z_{t-k}}\right)^{\frac{1}{2}s(s+1)} \left(\frac{Z_{t-k-N_j}}{Z_{t-k}}\right)^{\frac{1}{2}s(s-1)} \bmod p^{3j}.$$

Substituting $Z_n \equiv \xi\, \mu^{\frac{1}{2}n(n-1)}\, h_{k+n} \bmod p^{3j}$ for all $n \in I$, we get

$$\begin{aligned}\frac{Z_{t-k+sN_j}}{Z_{t-k}} &\equiv \frac{\xi\, \mu^{\frac{1}{2}(st-k+N_j)(st-k+N_j-1)}\, h_{t+sN_j}}{\xi\, \mu^{\frac{1}{2}(t-k)(t-k-1)}\, h_t} \bmod p^{3j} \\ &\equiv \mu^{\frac{1}{2}(s^2N_j^2+sN_j(2t-2k-1))}\, \frac{h_{t+sN_j}}{h_t} \bmod p^{3j}.\end{aligned}$$

So

$$\frac{Z_{t-k+N_j}}{Z_{t-k}} \equiv \mu^{\frac{1}{2}(N_j^2+N_j(2t-2k-1))}\, \frac{h_{t+N_j}}{h_t} \bmod p^{3j},$$

and

$$\frac{Z_{t-k-N_j}}{Z_{t-k}} \equiv \mu^{\frac{1}{2}(N_j^2-N_j(2t-2k-1))}\, \frac{h_{t-N_j}}{h_t} \bmod p^{3j}.$$

It follows that

$$\mu^{\frac{1}{2}\left(s^2N_j^2+sN_j(2t-2k-1)\right)} \cdot \frac{h_{t+sN_j}}{h_t} \equiv$$
$$\left(\mu^{\frac{1}{2}\left(N_j^2+N_j(2t-2k-1)\right)}\, \frac{h_{t+N_j}}{h_t}\right)^{\frac{1}{2}s(s+1)} \left(\mu^{\frac{1}{2}(N_j^2-N_j(2t-2k-1))}\, \frac{h_{t-N_j}}{h_t}\right)^{\frac{1}{2}s(s-1)} \bmod p^{3j},$$

or simplifying,

$$\frac{h_{t+sN_j}}{h_t} \equiv \left(\frac{h_{t+N_j}}{h_t}\right)^{\frac{1}{2}s(s+1)} \left(\frac{h_{t-N_j}}{h_t}\right)^{\frac{1}{2}s(s-1)} \bmod p^{3j}.$$

□

Theorem 8.4.7 says that, if $h_t \not\equiv 0 \bmod p$ and we consider the sequence (ℓ_n) obtained from (h_n) by taking every N_jth term from h_t and dividing by h_t, then the simple form for the sth term $\ell_s = \frac{h_{t+sN_j}}{h_t}$ we found after Theorem 8.4.2 actually holds modulo p^{3j}, not just modulo p^j, i.e.,

$$\ell_s \equiv \ell_1^{\frac{1}{2}s(s+1)}\, \ell_{-1}^{\frac{1}{2}s(s-1)} \bmod p^{3j} \quad \text{for all } t + sN_j \in I.$$

Remarks:

1. We cannot phrase Theorem 8.4.7 as a formula involving N_r and holding modulo p^{3r} (as we phrased Theorem 5.2.1), because in this chapter r is fixed, and we only have $Z_n \equiv \xi\, \mu^{\frac{1}{2}n(n-1)}\, h_{n+k} \bmod p^r$, not modulo p^{3r}.

2. Notice that, unlike Theorem 8.4.2, Theorem 8.4.7 gives no information for $s = 1$.

8.4.2 Writing β_r and γ_r in terms of β_w and γ_w

In this section we find a simple expression for the constants β_r and γ_r in terms of β_w and γ_w (and hence in terms of h_{k-1}, h_{k-2}, h_{k+N_1-1} and h_{k+N_1-2} modulo p^w), in the case where p is an odd prime. We will need these expressions in the next section to prove our Theorem 8.5.6, which describes how the period of $(h_n \bmod p^r)$ increases as r increases.

If $r \le w$ (i.e., if p^r divides all multiples of p in (h_n)), then of course $N_r = N_w = N_1$ and so $\gamma_r \equiv \gamma_w \bmod p^r$ and $\beta_r \equiv \beta_w \bmod p^r$ by definition of γ_r and β_r.

Theorem 8.4.8. *If p is an odd prime then for $r > w$,*

$$\gamma_r \equiv \gamma_{r-1}^p \bmod p^r$$

and

$$\beta_r \equiv \left(\gamma_{r-1}^{\frac{1}{2}N_{r-1}(p-1)}\, \beta_{r-1}\right)^p \bmod p^r.$$

Proof: Since $r > w$, by Theorem 7.6.7 $N_r = p\, N_{r-1}$. Since (Z_n) is an EDS, by Corollary 5.4.2 we have

$$c_r \equiv c_{r-1}^p \bmod p^r \quad \text{and} \quad b_r \equiv \left(c_{r-1}^{\frac{1}{2}(p-1)\,N_{r-1}}\, b_{r-1}\right)^p \bmod p^r.$$

The result follows on substituting $\gamma_r \equiv \mu^{-N_r}\, c_r \bmod p^r$ and $\beta_r \equiv \mu^{-\frac{1}{2}N_r(N_r-1)}\, \gamma_r^{-k}\, b_r \bmod p^r$ from Theorem 8.4.1 and simplifying. □

It now follows by an easy induction that

Theorem 8.4.9. *If p is an odd prime then for $r \geq w$,*

$$\gamma_r \equiv \gamma_w^{p^{r-w}} \bmod p^r$$

and

$$\beta_r \equiv \left(\gamma_w^{\frac{1}{2}N_1\,(p^{r-w}-1)}\, \beta_w\right)^{p^{r-w}} \bmod p^r.$$

8.5 The period of $(h_n \bmod p^r)$

We now prove that if p is odd then for every $r \geq w$ the sequence $(h_n \bmod p^r)$ is periodic with period $p^{r-w}\, \tau_w\, N_w$, where τ_w is a constant that can be found from the sequence. This confirms Robinson's conjectures 6.7.5 and 6.7.6 for the case where some term of the sequence is divisible by p but not by p^2 (i.e., $w = 1$). The proofs in this section are in most cases generalisations of the corresponding elliptic divisibility sequence proofs in chapter 5.

8.5.1 The constant τ_r

We first prove $(h_n \bmod p^r)$ is periodic.

Definition: For $1 \leq j \leq r$ we define a constant τ_j to be the least positive integer for which

$$\gamma_j^{\tau_j} \equiv 1 \bmod p^j$$

and

$$(-\beta_j)^{\tau_j} \equiv \gamma_j^{-\frac{1}{2}\tau_j(\tau_j-1)\,N_j} \bmod p^j.$$

Theorem 8.5.1. *Let (h_n) be a Somos 4 sequence and let p be a regular prime p with gap $N_1 \geq 5$. For $r \in \mathbb{N}$ let p^r divide some term h_k and let the gap of p^r in (h_n) be N_r. Then $(h_n \bmod p^r)$ is periodic with period $\pi_r = \tau_r N_r$.*

Proof: By Theorem 8.4.2,

$$h_{t+\tau_r N_r} \equiv \left(\gamma_r^{N_r}\right)^{\frac{1}{2}\tau_r(\tau_r-1)} \, (\gamma_r^{\tau_r})^t \, (-\beta_r)^{\tau_r} \, h_t \equiv h_t \bmod p^r$$

whenever these indices are in I. It follows that $(h_n \bmod p^r)$ is periodic with period π_r dividing $\tau_r N_r$. Furthermore, since $h_n \equiv 0 \bmod p^r$ if and only if $n \equiv k \bmod N_r$, the period π_r must be a multiple of N_r — say $\pi_r = vN_r$.

Now since (by definition of π_r)

$$h_{t+vN_r} \equiv \left(\gamma_r^{N_r}\right)^{\frac{1}{2}v(v-1)} \, (\gamma_r^{v})^t \, (-\beta_r)^{v} \, h_t \equiv h_t \bmod p^r,$$

we have

$$\left(\gamma_r^{N_r}\right)^{\frac{1}{2}v(v-1)} \, (\gamma_r^{v})^t \, (-\beta_r)^{v} \equiv 1 \bmod p^r$$

for all t with $t, t+vN_r \in I$. It follows (setting $t = k+2$ and $t = k+1$ and dividing) that $\gamma_r^v \equiv 1 \bmod p^r$, and hence that $\left(\gamma_r^{N_r}\right)^{\frac{1}{2}v(v-1)} \, (-\beta_r)^v \equiv 1 \bmod p^r$. Hence (by definition of τ_r) $v \geq \tau_r$, and so $v = \tau_r$. □

So τ_r is the number of zeroes before $(h_n \bmod p^r)$ starts repeating.

Remarks:

1. Theorem 8.5.1 matches the alternative definition of τ_r for EDSs given in Theorem 5.6.2. It is also easy to prove that τ_r is the smallest positive integer such that

   $$\gamma_r^{\tau_r} \equiv 1 \bmod p^r \quad \text{and} \quad (-\beta_r)^{\tau_r^2} \left(\mu^{-N_r}\right)^{\frac{1}{2}\tau_r(\tau_r-1)} \equiv 1 \bmod p^r,$$

   which matches the first definition of τ_r for EDSs in chapter 5. We use the fact that $\mu^{-N_r} \equiv \gamma_r^{N_r-2k} \cdot \beta_r^{-2} \bmod p^r$, so

   $$\begin{aligned}(-\beta_r)^{\tau_r^2} \left(\mu^{-N_r}\right)^{\frac{1}{2}\tau_r(\tau_r-1)} &\equiv (-\beta_r)^{\tau_r^2} \left(\gamma_r^{N_r-2k} \cdot \beta_r^{-2}\right)^{\frac{1}{2}\tau_r(\tau_r-1)} \bmod p^r \\ &\equiv (-\beta_r)^{\tau_r} \cdot \gamma_r^{\frac{1}{2}\tau_r(\tau_r-1)N_r} \cdot (\gamma_r^{\tau_r})^{-k(\tau_r-1)} \bmod p^r \\ &\equiv (-\beta_r)^{\tau_r} \cdot \gamma_r^{\frac{1}{2}\tau_r(\tau_r-1)N_r} \bmod p^r,\end{aligned}$$

   since $\gamma_r^{\tau_r} \equiv 1 \bmod p^r$.

2. Since $\mathbb{Z}_{p^r}^*$ has $p^{r-1}(p-1)$ elements, it follows easily from the definition that τ_r divides $2\,p^{r-1}(p-1)$. So the period π_r of $(h_n \bmod p^r)$ is a multiple of N_r and a divisor of $2(p-1)N_r$ (see Corollary 8.5.4).

3. Clearly τ_r does not change if k increases by a multiple of N_r (since τ_r depends only on β_r, γ_r and N_r, which are independent of k). This makes sense, as the period of $(h_n \bmod p^r)$ should not depend on our choice of k.

4. Note that, although the sequences $(h_n \bmod p^r)$ and $(\ell_n \bmod p^r)$ have zeroes in the same places, they do not necessarily have the same period.

Lemma 8.5.2. *[30]*

The integers y which satisfy

$$\gamma_r^y \equiv 1 \bmod p^r \tag{8.5a}$$

and

$$(-\beta_r)^y \equiv \gamma_r^{-\frac{1}{2}y(y-1)N_r} \bmod p^r \tag{8.5b}$$

are precisely the multiples of τ_r.

Proof: First let $a\,\tau_r$ be any multiple of τ_r. Then

$$\gamma_r^{a\,\tau_r} \equiv (\gamma_r^{\tau_r})^a \equiv 1 \bmod p^r,$$

and

$$\begin{aligned}(-\beta_r)^{a\,\tau_r} &\equiv \gamma_r^{-\frac{1}{2}a\,\tau_r(\tau_r-1)N_r} \bmod p^r \\ &\equiv \gamma_r^{-\frac{1}{2}(a\,\tau_r)(a\,\tau_r-1)N_r} \cdot \gamma_r^{\frac{1}{2}(a\,\tau_r)(a\,\tau_r-\tau_r)N_r} \bmod p^r \\ &\equiv \gamma_r^{-\frac{1}{2}(a\,\tau_r)(a\,\tau_r-1)N_r} \cdot (\gamma_r^{\tau_r})^{\frac{1}{2}a(a-1)\tau_r N_r} \bmod p^r \\ &\equiv \gamma_r^{-\frac{1}{2}(a\,\tau_r)(a\,\tau_r-1)N_r} \bmod p^r.\end{aligned}$$

So all multiples of τ_r satisfy (8.5a) and (8.5b).

For the converse, note that if y is a solution of (8.5a) and (8.5b) then by Theorem 8.4.2, for all t with $t, t+yN_r \in I$,

$$\begin{aligned}h_{t+yN_r} &\equiv (\gamma_r^y)^t\ \gamma_r^{\frac{1}{2}y(y-1)N_r}\,(-\beta_r)^y\,h_t \bmod p^r \\ &\equiv h_t \bmod p^r.\end{aligned}$$

It follows that yN_r is a multiple of the period $\pi_r = \tau_r N_r$, and hence that y is a multiple of τ_r. □

If p is odd then we can find τ_r from the order of γ_r and $-\beta_r$ in $\mathbb{Z}_{p^r}^*$:

Theorem 8.5.3. *Let p be an odd prime, and let n_r and ν_r be the orders of γ_r and $(-\beta_r)$ respectively modulo p^r. Then*

$$\tau_r = \begin{cases} \frac{1}{2}\operatorname{lcm}(n_r, \nu_r) & \text{if } N_1 \text{ is odd and } \frac{\nu_r}{2} \text{ and } n_r \text{ are even} \\ & \text{and divisible by exactly the same power of } 2, \\ 2\operatorname{lcm}(n_r, \nu_r) & \text{if } N_1 \text{ is odd and } n_r \text{ is even and divisible by} \\ & \text{at least as high a power of } 2 \text{ as } \nu_r \text{ is, or} \\ \operatorname{lcm}(n_r, \nu_r) & \text{otherwise.} \end{cases}$$

Proof: Notice that since p is odd, N_r and N_1 have the same parity. Since $\gamma_r^{\tau_r} \equiv (-\beta_r)^{2\tau_r} \equiv 1 \bmod p^r$, τ_r is a multiple of n_r and of ν_r if ν_r is odd, or $\frac{\nu_r}{2}$ if ν_r is even.

If n_r or ν_r is even, let $y = \frac{1}{2}\operatorname{lcm}(n_r, \nu_r)$. So γ_r^y and β_r^y are $\equiv \pm 1 \bmod p^r$, and τ_r is a multiple of y. Note that $\gamma_r^y \equiv 1 \bmod p^r$ if and only if y is a multiple of n_r, i.e., if and only if ν_r is divisible by a higher power of 2 than n_r. If this is the case, then ν_r is even and y is an odd multiple of $\frac{\nu_r}{2}$, so $(-\beta_r)^y \equiv -1 \bmod p^r$.

But $\gamma_r^{-\frac{1}{2}y(y-1)N_r} \equiv -1 \bmod p^r$ if and only if n_r is even and $\frac{1}{2}\,y\,(y-1)\,N_r$ is an odd multiple of $\frac{n_r}{2}$, i.e., $y(y-1)N_r$ is an odd multiple of n_r. If $n_r \mid y$ then this is true if and only if n_r is even, N_1 is odd and y is an odd multiple of n_r. It follows from Lemma 8.5.2 that $\tau_r \mid y$ (and hence $\tau_r = y$) if and only if N_1 is odd, n_r and ν_r are both even, and n_r and $\frac{\nu_r}{2}$ are divisible by the same power of 2.

Now let $z = \operatorname{lcm}(n_r, \nu_r)$. So $\gamma_r^z \equiv \beta_r^z \equiv 1 \bmod p^r$. Note that $\gamma_r^{-\frac{1}{2}z(z-1)N_r} \equiv 1 \bmod p^r$ if and only if $\frac{1}{2}z(z-1)N_r$ is a multiple of n_r, i.e., if and only if N_1 is even or z is odd or $\frac{z}{2}$ is a multiple of n_r. So by Lemma 8.5.2 $\tau_r \mid z$ if and only if N_1 is even or n_r and ν_r are both odd or ν_r is divisible by a higher power of 2 than n_r. It follows that $\tau_r = z$ if and only if N_1 is even or n_r is odd, or $\frac{\nu_r}{2}$ is divisible by a higher power of 2 than n_r.

Finally, we note that since $\gamma_r^z \equiv (-\beta_r)^z \equiv 1 \bmod p^r$, it follows from Lemma 8.5.2 that τ_r divides $2\,z$. The result follows. □

If (h_n) is an EDS then we have $\gamma_r^{N_r} \equiv \beta_r^2 \bmod p^r$, so if N_1 is odd then n_r is divisible by a higher power of 2 than ν_r, and so the case $\tau_r = 2\,\mathrm{lcm}(n_r, \nu_r)$ does not occur. This agrees with Theorem 5.6.3.

Corollary 8.5.4. *If p is an odd prime, then τ_1 divides $p-1$ if and only if N_1 is even or μ is a quadratic residue modulo p. Otherwise τ_1 divides $2\,(p-1)$.*

Proof: Since $\mathrm{lcm}(n_1, \nu_1) \mid (p-1)$, it follows from Theorem 8.5.3 that $\tau_1 \nmid (p-1)$ if and only if N_1 is odd and n_1 is divisible by the same power of 2 as $p-1$, i.e., if and only if N_1 is odd and γ_1 is a quadratic non-residue modulo p. Since $(\gamma_1\,\mu)^{N_1} \equiv \left(\beta_1\,\gamma_1^k\right)^2 \bmod p$, this is true if and only if N_1 is odd and μ is a quadratic non-residue modulo p. □

Note that τ_1 now divides $2\,(p-1)$, whereas in the EDS case (Theorem 5.6.1) we had $\tau_1 \mid (p-1)$. The difference is that if (h_n) is an EDS then $\mu = 1$ is a quadratic residue modulo p.

8.5.2 How the period of $(h_n \bmod p^r)$ increases with r

The next theorem describes the behaviour of τ_r as r increases for odd primes p.

Theorem 8.5.5. *Let (h_n) be a Somos 4 sequence and p an odd regular prime with gap $N_1 \geq 5$ in (h_n). Let $r \in \mathbb{N}$ such that p^r divides some term of (h_n), and let p^w be the highest power of p which divides every multiple of p in (h_n). Then there exists a positive integer $u \leq w$ such that*

$$\tau_r = \begin{cases} \tau_1 & \text{if} \quad r \leq u \\ p^{r-u}\,\tau_1 & \text{if} \quad u \leq r \leq w, \text{ and} \\ \tau_w & \text{if} \quad r \geq w. \end{cases}$$

Proof: We first consider the case $2 \leq r \leq w$. Then $N_r = N_{r-1} = N_1$, so by definition of β_r and γ_r,

$$\gamma_r \equiv \gamma_{r-1} \bmod p^{r-1} \qquad \text{and} \qquad \beta_r \equiv \beta_{r-1} \bmod p^{r-1}.$$

So by Theorem 2.2.5, for any integer x we have

$$\gamma_r^{xp} \equiv 1 \bmod p^r \;\Leftrightarrow\; \gamma_r^x \equiv 1 \bmod p^{r-1} \;\Leftrightarrow\; \gamma_{r-1}^x \equiv 1 \bmod p^{r-1} \tag{8.6a}$$

and

$$(-\beta_r)^{xp} \equiv \gamma_r^{-\frac{1}{2}xp\,(xp-1)N_r} \bmod p^r \;\Leftrightarrow\; (-\beta_r)^x \equiv \gamma_r^{-\frac{1}{2}x\,(xp-1)N_r} \bmod p^{r-1}$$
$$\Leftrightarrow\; (-\beta_{r-1})^x \equiv \gamma_{r-1}^{-\frac{1}{2}x\,(xp-1)N_{r-1}} \bmod p^{r-1}.$$

But $-\frac{1}{2}x\,(xp-1) = -\frac{1}{2}x(x-1) - \frac{1}{2}(p-1)x^2$, so we can write

$$\gamma_{r-1}^{-\frac{1}{2}x\,(xp-1)N_{r-1}} \equiv \gamma_{r-1}^{-\frac{1}{2}x(x-1)N_{r-1}} \cdot \left(\gamma_{r-1}^x\right)^{-\frac{1}{2}(p-1)xN_{r-1}} \bmod p^{r-1}.$$

It follows that *if* $\gamma_{r-1}^x \equiv 1 \bmod p^{r-1}$, then

$$(-\beta_r)^{x\,p} \equiv \gamma_r^{-\frac{1}{2}x\,p(x\,p-1)N_r} \bmod p^r$$
$$\Leftrightarrow (-\beta_{r-1})^x \equiv \gamma_{r-1}^{-\frac{1}{2}x(x-1)N_{r-1}} \bmod p^{r-1}. \tag{8.6b}$$

Clearly π_r is a multiple of π_{r-1} (since if the sequence repeats modulo p^r then it must repeat modulo p^{r-1}), so by Theorem 8.5.1 τ_r is a multiple of τ_{r-1}. Setting $x = \tau_{r-1}$ in (8.6a) and (8.6b) shows that $\tau_r \mid p\,\tau_{r-1}$ by Lemma 8.5.2. It follows that $\tau_r = \tau_{r-1}$ or $p\,\tau_{r-1}$.

If τ_r is divisible by p, then setting $x = \frac{\tau_r}{p}$ in (8.6a) and (8.6b) shows that $\frac{\tau_r}{p}$ is a multiple of τ_{r-1} by Lemma 8.5.2. Hence $\tau_r = p\,\tau_{r-1}$ if τ_r is divisible by p, and otherwise $\tau_r = \tau_{r-1}$. It follows that $\tau_r = \tau_1$ for $r = 1, 2, \ldots, u$ for some $u \le w$, and $\tau_r = p\,\tau_{r-1} = p^{r-u}\,\tau_1$ for $u+1 \le r \le w$.

We now consider the case $r \ge w+1$. Then $N_r = pN_{r-1}$, and by Theorem 8.4.8 (since p is odd),

$$\gamma_r \equiv \gamma_{r-1}^p \bmod p^r$$

and

$$-\beta_r \equiv \left(-\gamma_{r-1}^{\frac{1}{2}N_{r-1}\,(p-1)}\,\beta_{r-1}\right)^p \bmod p^r.$$

Hence by Theorem 2.2.5, for any integer x we have

$$\gamma_r^x \equiv 1 \bmod p^r \;\Leftrightarrow\; (\gamma_{r-1}^p)^x \equiv 1 \bmod p^r$$
$$\Leftrightarrow\; \gamma_{r-1}^x \equiv 1 \bmod p^{r-1}.$$

Similarly,

$$\begin{aligned}
(-\beta_r)^x &\equiv \gamma_r^{-\frac{1}{2}x(x-1)N_r} \bmod p^r \\
&\Leftrightarrow \left(-\gamma_{r-1}^{\frac{1}{2}N_{r-1}(p-1)}\,\beta_{r-1}\right)^{px} \equiv \left(\gamma_{r-1}^{p}\right)^{-\frac{1}{2}x(x-1)pN_{r-1}} \bmod p^r \\
&\Leftrightarrow (-\beta_{r-1})^{px} \equiv \left(\gamma_{r-1}^{-\frac{1}{2}x(x-1)}\right)^{p^2 N_{r-1}} \cdot \left(\left(\gamma_{r-1}^{x}\right)^{-\frac{1}{2}(p-1)N_{r-1}}\right)^{p} \bmod p^r \\
&\Leftrightarrow (-\beta_{r-1})^{x} \equiv \left(\gamma_{r-1}^{-\frac{1}{2}x(x-1)}\right)^{pN_{r-1}} \cdot \left(\gamma_{r-1}^{x}\right)^{-\frac{1}{2}(p-1)N_{r-1}} \bmod p^{r-1}.
\end{aligned}$$

If $\gamma_{r-1}^x \equiv 1 \bmod p^{r-1}$, then (since p is odd) $\gamma_{r-1}^{\frac{1}{2}x(x-1)} \equiv \pm 1 \bmod p^{r-1}$, and we have

$$(-\beta_r)^x \equiv \gamma_r^{-\frac{1}{2}x(x-1)N_r} \bmod p^r \Leftrightarrow (-\beta_{r-1})^x \equiv \gamma_{r-1}^{-\frac{1}{2}x(x-1)N_{r-1}} \bmod p^{r-1}.$$

It follows by the definition of τ_{r-1} and τ_r that $\tau_r = \tau_{r-1}$ for all $r \geq w+1$. □

In other words, for $r \geq w$ the reduced sequence $(h_n \bmod p^r)$ repeats after the same number of zeroes no matter what r is.

Since $N_r = N_1$ for $r \leq w$, $N_r = p^{r-w} N_1$ for $r \geq w$ and $\pi_r = N_r \tau_r$ for all $r \in \mathbb{N}$, Theorem 8.5.5 leads to our main result in this section:

Theorem 8.5.6. *Let (h_n) be a Somos 4 sequence and let p be a regular odd prime with gap $N_1 \geq 5$. Let $r \in \mathbb{N}$ such that p^r divides some term of (h_n), and let p^w be the highest power of p which divides every multiple of p in (h_n). Then there exists a positive integer $u \leq w$ such that $(h_n \bmod p^r)$ has period*

$$\pi_r = \begin{cases} \pi_1 & \text{if } r \leq u, \text{ and} \\ p^{r-u}\,\pi_1 & \text{if } r \geq u. \end{cases}$$

So if some term of (h_n) is divisible by p^{w+1}, then as r increases from 1 the period of $(h_n \bmod p^r)$ remains the same until r reaches some value $u \leq w$, and then increases by a factor of p each time.

Since $N_1 \geq 5$ for every prime p dividing some term of Somos(4), this proves Robinson's Conjecture 6.7.5 for primes with $w = 1$. It remains open whether u is always 1 for Somos(4), but we have found other Somos 4 sequences and primes for which $u > 1$; an example is the EDS (h_n) given after Theorem 5.6.7, which has $u = 2$ for $p = 5$.

In the next chapter we extend this result to the case where p does *not* divide any term of (h_n).

Since $\pi_1 = \tau_1 N_1$, by Corollary 8.5.4 we have

Corollary 8.5.7. *If p is an odd prime, then π_1 divides $(p-1)N_1$ if and only if either N_1 is even or μ is a quadratic residue modulo p. Otherwise π_1 divides $2(p-1)N_1$.*

So the period of $(h_n \bmod p)$ is a multiple of N_1, and a divisor of $2(p-1)N_r$. Robinson's Conjecture 6.7.6 is that if (h_n) is Somos(4) then the period of $(h_n \bmod p)$ is a multiple of N_1, and a divisor of $(p-1)N_1$. The following example shows that there are Somos 4 sequences with $\tau_1 \nmid (p-1)$; however, it remains open whether Somos(4) is one of them for every prime p dividing some term. (Since Somos(4) is symmetric around the initial values, N_1 is always odd; so the question is whether μ is always a quadratic residue modulo p.)

Example: Let (h_n) be the Somos 4 sequence with $\lambda_1 = -1$, $\lambda_2 = 3$ and initial values $1, 4, 6, 6$, and let $p = 7$.

Then $N_1 = 5$ and the reduced sequence $(h_n \bmod 7)$ is

$$\ldots, 0, 1, 4, 6, 6, 0, 6, 1, 4, 6, 0, 3, 6, 1, 4, 0, 1, 3, 6, 1, 0, 1, 1, 3, 6, 0, 3, 1, 1, 3,$$
$$0, 6, 3, 1, 1, 0, 1, 6, 3, 1, 0, 4, 1, 6, 3, 0, 6, 4, 1, 6, 0, 6, 6, 4, 1, 0, 4, 6, 6, 4,$$
$$0, 1, 4, 6, 6, 0, \ldots$$

So $\pi = 60$, i.e., the sequence repeats after $12 = 2(p-1)$ zeroes modulo 7.

Notice that in this sequence $N_1 = 5$ is odd, and if we use $k = 5$ then $\mu = -\frac{h_6}{h_4} = -\frac{6}{6} = -1$, which is a quadratic non-residue in $\mathbb{Z}_7$. So $\pi_1 \nmid (p-1)N_1$ is as expected from Corollary 8.5.7.

Finally, we note that we have not proved any results about the periodicity of the sequence $(h_n \bmod p^r)$ in the case where all multiples of p in (h_n) are divisible by exactly the same power p^w of p and $r > w$. We have also not covered prime powers p^r where $r \geq 2$ and p is irregular or has gap $N_1 \leq 4$, and we have obtained only partial results for $p = 2$.

Chapter 9

Primes not appearing in a Somos 4 sequence

We now use the relationship between Somos 4 sequences and elliptic curves to extend our results on the symmetry and periodicity of $(h_n \bmod p^r)$ to the case where p does not divide λ_1 or any term of (h_n).

9.1 Preliminaries

In this chapter let (h_n) be a Somos 4 sequence with coefficients λ_1, λ_2, and let θ be the squarefree part of λ_1. Let p be a reasonable prime not dividing any term of (h_n) or λ_1, such that (h_n) is not equivalent to any sequence which is constant and non-zero modulo p. (So (h_n) has no zero terms and is not equivalent to the constant sequence $\ldots, 1, 1, 1, \ldots$)

Then by Theorem 7.6.1 there exists an elliptic curve $E/\mathbb{Q}$ containing rational points $P = (0, 0)$ and $Q = (x_0, y_0)$ such that the coefficients a_i of E are p-integers, P and Q are non-singular modulo p, and the x-coordinate of $Q + [n]P$ is

$$x_n \equiv -\theta \cdot \frac{h_{n-1}\, h_{n+1}}{h_n^2} \bmod p^r \quad \text{for all } n \in \mathbb{Z}.$$

Since p does not divide any term of (h_n), $Q \bmod p \notin \langle P \bmod p \rangle$ in $E(\mathbb{Z}_p)$, i.e., $Q + [n]P \not\equiv \mathcal{O}_p \bmod p$ for any $n \in \mathbb{Z}$.

Various products of consecutive x_i can be usefully expressed in terms of the h_n:

Theorem 9.1.1. *For $n > 0$ and $t \in \mathbb{Z}$,*

$$\prod_{j=0}^{n-1} x_{t+j} = (-\theta)^n \, \frac{h_{t-1}}{h_t} \cdot \frac{h_{t+n}}{h_{t+n-1}}, \tag{9.1a}$$

$$\prod_{j=0}^{n-1} x_{t+j}^{n-j} = (-\theta)^{\frac{1}{2}n(n+1)} \, \frac{1}{h_t} \left(\frac{h_{t-1}}{h_t}\right)^n h_{t+n}, \tag{9.1b}$$

$$\prod_{j=0}^{n-1} x_{t-j}^{n-j} = (-\theta)^{\frac{1}{2}n(n+1)} \, \frac{1}{h_t} \left(\frac{h_{t+1}}{h_t}\right)^n h_{t-n}. \tag{9.1c}$$

Proof: Substituting $x_n = -\theta \cdot \frac{h_{n-1}\, h_{n+1}}{h_n^2}$ for all $n \in \mathbb{Z}$, we get cancellation between neighbouring terms in the products:

$$\begin{aligned}
\frac{1}{\theta^n} \prod_{j=0}^{n-1} x_{t+j} &= \prod_{j=0}^{n-1} \left(\frac{x_{t+j}}{\theta}\right) \\
&= \left(-\frac{h_{t-1}\, h_{t+1}}{h_t^2}\right) \cdot \left(-\frac{h_t\, h_{t+2}}{h_{t+1}^2}\right) \cdots \left(-\frac{h_{t+n-3}\, h_{t+n-1}}{h_{t+n-2}^2}\right) \cdot \left(-\frac{h_{t+n-2}\, h_{t+n}}{h_{t+n-1}^2}\right) \\
&= (-1)^n \, \frac{h_{t-1}}{h_t} \cdot \frac{h_{t+n}}{h_{t+n-1}}.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\frac{1}{\theta^{\frac{1}{2}n(n+1)}} \prod_{j=0}^{n-1} x_{t+j}^{n-j} &= \prod_{j=0}^{n-1} \left(\frac{x_{t+j}}{\theta}\right)^{n-j} \\
&= \left(-\frac{h_{t-1}\, h_{t+1}}{h_t^2}\right)^n \cdot \left(-\frac{h_t\, h_{t+2}}{h_{t+1}^2}\right)^{n-1} \cdot \left(-\frac{h_{t+1}\, h_{t+3}}{h_{t+2}^2}\right)^{n-2} \\
&\qquad \cdots \left(-\frac{h_{t+n-3}\, h_{t+n-1}}{h_{t+n-2}^2}\right)^2 \cdot \left(-\frac{h_{t+n-2}\, h_{t+n}}{h_{t+n-1}^2}\right) \\
&= (-1)^{\frac{1}{2}n(n+1)} \, \frac{1}{h_t} \left(\frac{h_{t-1}}{h_t}\right)^n h_{t+n}
\end{aligned}$$

and

$$\frac{1}{\theta^{\frac{1}{2}n(n+1)}} \prod_{j=0}^{n-1} x_{t-j}^{n-j} = \prod_{j=0}^{n-1} \left(\frac{x_{t-j}}{\theta}\right)^{n-j} = (-1)^{\frac{1}{2}n(n+1)} \, \frac{1}{h_t} \left(\frac{h_{t+1}}{h_t}\right)^n h_{t-n}.$$

The result follows. □

Theorem 9.1.1 with $t = 0$ allows us to express each term of (h_n) in terms of θ, the x_i and h_{-1}, h_0:

Corollary 9.1.2. *For $n \in \mathbb{N}$,*

$$h_n = \left(-\frac{1}{\theta}\right)^{\frac{1}{2}n(n+1)} x_{n-1}\, x_{n-2}^2 \,\ldots\, x_1^{n-1}\, x_0^n\, h_0 \left(\frac{h_0}{h_{-1}}\right)^n,$$

and

$$h_{-n} = \left(-\frac{1}{\theta}\right)^{\frac{1}{2}n(n-1)} x_{-n+1}\, x_{-n+2}^2 \,\ldots\, x_{-1}^{n-1}\, h_0 \left(\frac{h_{-1}}{h_0}\right)^n.$$

Proof: Setting $t = 0$ in (9.1b) gives

$$\prod_{j=0}^{n-1} x_j^{n-j} = (-\theta)^{\frac{1}{2}n(n+1)} \frac{1}{h_0} \left(\frac{h_{-1}}{h_0}\right)^n h_n,$$

i.e.,

$$h_n = (-\theta)^{-\frac{1}{2}n(n+1)} \prod_{j=0}^{n-1} x_j^{n-j}\, h_0 \left(\frac{h_0}{h_{-1}}\right)^n.$$

Setting $t = 0$ in (9.1c) gives

$$\prod_{j=0}^{n-1} x_{-j}^{n-j} = (-\theta)^{\frac{1}{2}n(n+1)} \frac{1}{h_0} \left(\frac{h_1}{h_0}\right)^n h_{-n},$$

i.e.,

$$h_{-n} = (-\theta)^{-\frac{1}{2}n(n+1)}\, x_{-n+1}\, x_{-n+2}^2 \,\ldots\, x_{-1}^{n-1}\, x_0^n\; h_0 \left(\frac{h_0}{h_1}\right)^n.$$

But $x_0^n = \left(-\theta \frac{h_{-1} h_1}{h_0^2}\right)^n$, so

$$h_{-n} = (-\theta)^{-\frac{1}{2}n(n+1)-n}\, x_{-n+1}\, x_{-n+2}^2 \,\ldots\, x_{-1}^{n-1}\, h_0 \left(\frac{h_{-1}}{h_0}\right)^n.$$

The result follows. □

9.2 The constants γ_r and β_r

For $r \in \mathbb{N}$, let N_r be the order of P in $E(\mathbb{Z}_{p^r})$. So

$$x_{N_r+t} \equiv x_t \bmod p^r \quad \text{for all } t \in \mathbb{Z}.$$

Let p^w be the highest power of p dividing the weighted Z-coordinate of $[N_1]P$ in $E(\mathbb{Q})$, i.e., such that $[N_1]P \equiv \mathcal{O}_{p^w} \bmod p^w$. Then by Theorem 3.9.5, we have

$$N_r = \begin{cases} N_1 & \text{for } r \leq w, \text{ or} \\ p^{r-w} N_1 & \text{for } r \geq w. \end{cases}$$

We define two constants related to (h_n), E, P and Q:

Definition: For $r \in \mathbb{N}$,

$$\gamma_r = \prod_{j=0}^{N_r-1} \left(-\frac{x_j}{\theta}\right) \bmod p^r,$$

and

$$\beta_r = -\frac{h_{N_r}}{h_0} \bmod p^r.$$

In other words, γ_r is $\left(-\frac{1}{\theta}\right)^{N_r}$ times the product of the x-coordinates of the N_r distinct points $Q + [n]P \bmod p^r$ in $E(\mathbb{Z}_{p^r})$, and $-\beta_r$ is the inverse ratio of the starting term of (h_n) to the term N_r positions later, both reduced modulo p^r.

Note that we are using the same notation as we used for the constants

$$\gamma_r = \frac{h_{N_r+k-1}}{h_{k-1}} \cdot \frac{h_{k-2}}{h_{N_r+k-2}} \bmod p^r,$$

and

$$\beta_r = -\gamma_r^{-k+1} \cdot \frac{h_{k+N_r-1}}{h_{k-1}} \bmod p^r$$

defined in chapter 8 for the case where p^r divides h_k. We will prove in Corollary 9.3.5 that for *any* $k \in \mathbb{Z}$ our new constants γ_r and β_r also satisfy these two congruences.

We now show that, for each $r \geq w+2$, we can write γ_r and β_r in terms of γ_{r-1} and β_{r-1}.

Of course, for $r \leq w$ we have $N_r = N_1$, so $\gamma_r = \prod_{j=0}^{N_1-1} \left(-\frac{x_j}{\theta}\right) \equiv \gamma_w \bmod p^r$ and $\beta_r \equiv -\frac{h_{N_1}}{h_0} \equiv \beta_w \bmod p^r$.

Theorem 9.2.1. *If p is odd, then for $r > w$,*

$$\gamma_r \equiv \gamma_{r-1}^p \bmod p^r.$$

Proof: Since $r > w$, $N_r = p\,N_{r-1}$. So

$$\begin{aligned}\gamma_r &= (-\theta)^{-p\,N_{r-1}} \prod_{j=0}^{p\,N_{r-1}-1} x_j \\ &= (-\theta)^{-p\,N_{r-1}} \left(\prod_{j=0}^{N_{r-1}-1} x_j\right)\cdot\left(\prod_{j=0}^{N_{r-1}-1} x_{N_{r-1}+j}\right)\cdots\left(\prod_{j=0}^{N_{r-1}-1} x_{(p-1)N_{r-1}+j}\right).\end{aligned}$$

Grouping the products in a different order, we get

$$\gamma_r = (-\theta)^{-p\,N_{r-1}} \prod_{j=0}^{N_{r-1}-1} \Big(x_j \cdot x_{N_{r-1}+j}\cdots x_{(p-1)N_{r-1}+j}\Big).$$

Note that, if $B_i = x_{iN_{r-1}+j} \bmod p^r$ for $i = 0, \ldots, p-1$ and $B = x_j \bmod p^r$, then the B_i are p integers which are distinct modulo p^r but all congruent to B modulo p^{r-1}. Hence by Theorem 2.2.8,

$$\prod_{i=0}^{p-1} x_{iN_{r-1}+j} \equiv x_j^p \bmod p^r.$$

So we have

$$\begin{aligned}\gamma_r &\equiv (-\theta)^{-p\,N_{r-1}} \prod_{j=0}^{N_{r-1}-1} x_j^p \bmod p^r \\ &\equiv \left((-\theta)^{-N_{r-1}} \prod_{j=0}^{N_{r-1}-1} x_j\right)^p \bmod p^r \\ &\equiv \gamma_{r-1}^p \bmod p^r.\end{aligned}$$

□

Theorem 9.2.2. *If p is odd, then for $r > w+1$,*

$$\beta_r \equiv \left(\gamma_{r-1}^{\frac{1}{2}(p-1)\,N_{r-1}}\,\beta_{r-1}\right)^p \bmod p^r.$$

Proof: Since $r > w+1$, $N_r = p\,N_{r-1}$ and $N_{r-1} = p\,N_{r-2}$. So by Corollary 9.1.2,

$$\begin{aligned}\beta_r &\equiv -\frac{h_{N_r}}{h_0} \bmod p^r \\ &\equiv -(-\theta)^{-\frac{1}{2}N_r(N_r-1)} \left(\prod_{j=0}^{p\,N_{r-1}-1} x_j^{p\,N_{r-1}-j}\right)\left(\frac{h_0}{h_{-1}}\right)^{N_r} \bmod p^r \\ &\equiv -(-\theta)^{-\frac{1}{2}N_r(N_r-1)} \left(\prod_{n=0}^{p-1}\prod_{j=0}^{N_{r-1}-1} x_{j+nN_{r-1}}^{pN_{r-1}-(j+nN_{r-1})}\right)\left(\frac{h_0}{h_{-1}}\right)^{N_r} \bmod p^r.\end{aligned}$$

Grouping the products in a different order, we get

$$\begin{aligned}\beta_r &\equiv -(-\theta)^{-\frac{1}{2}N_r(N_r-1)} \left(\prod_{j=0}^{N_{r-1}-1}\prod_{n=0}^{p-1} x_{j+nN_{r-1}}^{(p-1-n)\,N_{r-1}} \cdot x_{j+nN_{r-1}}^{N_{r-1}-j}\right)\left(\frac{h_0}{h_{-1}}\right)^{N_r} \bmod p^r \\ &\equiv -(-\theta)^{-\frac{1}{2}N_r(N_r-1)} \cdot \prod_{j=0}^{N_{r-1}-1}\left(\prod_{n=0}^{p-1} x_{j+nN_{r-1}}^{p-1-n}\right)^{N_{r-1}} \\ &\qquad\qquad \cdot \prod_{j=0}^{N_{r-1}-1}\left(\prod_{n=0}^{p-1} x_{j+nN_{r-1}}\right)^{N_{r-1}-j}\left(\frac{h_0}{h_{-1}}\right)^{N_r} \bmod p^r.\end{aligned}$$

Note that, if $B_n = x_{nN_{r-1}+j} \bmod p^r$ for $n = 0, \ldots, p-1$ and $B = x_j \bmod p^r$, then the B_i are p integers which are distinct modulo p^r but all congruent to B modulo p^{r-1}. Hence by Theorem 2.2.8

$$\prod_{n=0}^{p-1} x_{j+nN_{r-1}} \equiv x_j^p \bmod p^r.$$

Furthermore, since

$$\prod_{n=0}^{p-1} x_{j+nN_{r-1}}^{p-1-n} \equiv \prod_{n=0}^{p-1} x_j^{p-1-n} \equiv x_j^{\frac{1}{2}p(p-1)} \bmod p^{r-1}$$

and N_{r-1} is divisible by p, by Theorem 2.2.5 we have

$$\left(\prod_{n=0}^{p-1} x_{j+nN_{r-1}}^{p-1-n}\right)^{N_{r-1}} \equiv \left(x_j^{\frac{1}{2}p(p-1)}\right)^{N_{r-1}} \bmod p^r.$$

Finally, we note that

$$\begin{aligned}\frac{1}{2}N_r(N_r-1) &= \frac{1}{2}pN_{r-1}(pN_{r-1}-N_{r-1}+N_{r-1}+1) \\ &= \frac{1}{2}p(p-1)N_{r-1}^2 + p\cdot\frac{1}{2}N_{r-1}(N_{r-1}-1).\end{aligned}$$

So we have

$$\begin{aligned}
\beta_r &\equiv -(-\theta)^{-\frac{1}{2}N_r(N_r-1)} \cdot \prod_{j=0}^{N_{r-1}-1} \left(x_j^{\frac{1}{2}p(p-1)}\right)^{N_{r-1}} \\
&\qquad \cdot \prod_{j=0}^{N_{r-1}-1} \left(x_j^p\right)^{N_{r-1}-j} \cdot \left(\frac{h_0}{h_{-1}}\right)^{p\,N_{r-1}} \bmod p^r \\
&\equiv -\left((-\theta)^{-N_{r-1}} \prod_{j=0}^{N_{r-1}-1} x_j\right)^{\frac{1}{2}p(p-1)N_{r-1}} \\
&\qquad \cdot \left((-\theta)^{-\frac{1}{2}N_{r-1}(N_{r-1}-1)} \prod_{j=0}^{N_{r-1}-1} x_j^{N_{r-1}-j} \cdot \left(\frac{h_0}{h_{-1}}\right)^{N_{r-1}}\right)^p \bmod p^r \\
&\equiv \gamma_{r-1}^{\frac{1}{2}p(p-1)N_{r-1}} \cdot \left(-\frac{h_{N_{r-1}}}{h_0}\right)^p \bmod p^r \\
&\equiv \left(\gamma_{r-1}^{\frac{1}{2}(p-1)\,N_{r-1}}\,\beta_{r-1}\right)^p \bmod p^r.
\end{aligned}$$

This completes the proof. □

The above proof requires N_{r-1} to be divisible by p, but if $p \nmid N_{r-1}$ then all congruences in the proof still hold modulo p^{r-1}. This gives us the following weaker result for $r = w + 1$:

Theorem 9.2.3. *If p is odd, then*

$$\beta_{w+1} \equiv \left(\gamma_w^{\frac{1}{2}(p-1)\,N_1}\,\beta_w\right)^p \bmod p^w.$$

9.3 The symmetry formula

In this section we extend Theorem 8.4.2 to the case where p does not divide any term of (h_n), by using elliptic curves to prove the same symmetry formula for $(h_n \bmod p^r)$.

Lemma 9.3.1. *For all $t \in \mathbb{Z}$,*

$$\frac{h_{t+N_r}}{h_t} \equiv \gamma_r^t\,(-\beta_r) \bmod p^r.$$

Proof: We let $t > 0$, but the proof for $t < 0$ is similar. By Corollary 9.1.2,

$$\begin{aligned} h_{t+N_r} &= (-\theta)^{-\frac{1}{2}(t+N_r)(t+N_r+1)} \left(x_{t+N_r-1}\, x_{t+N_r-2}^2 \cdots x_{N_r}^t \right) \\ &\qquad \cdot \left(x_{N_r-1}^{t+1}\, x_{N_r-2}^{t+2} \cdots x_1^{t+N_r-1}\, x_0^{t+N_r} \right) h_0 \left(\frac{h_0}{h_{-1}} \right)^{t+N_r} \\ &= (-\theta)^{-\frac{1}{2}(t+N_r)(t+N_r+1)} \left(\prod_{j=0}^{t-1} x_{N_r+j}^{t-j} \right) \\ &\qquad \cdot \left(\prod_{j=0}^{N_r-1} x_j \right)^t \left(\prod_{j=0}^{N_r-1} x_j^{N_r-j} \right) h_0 \left(\frac{h_0}{h_{-1}} \right)^{N_r} \left(\frac{h_0}{h_{-1}} \right)^t. \end{aligned}$$

Note that

$$\frac{1}{2}(t+N_r)(t+N_r+1) = \frac{1}{2}N_r(N_r+1) + \frac{1}{2}t(t+1) + t\,N_r.$$

Substituting $x_{N_r+j} \equiv x_j \bmod p^r$ for $j = 0, 1, \ldots, t-1$ and rearranging, we get

$$\begin{aligned} h_{t+N_r} &\equiv \left((-\theta)^{-\frac{1}{2}N_r(N_r-1)} \left(\prod_{j=0}^{N_r-1} x_j^{N_r-j} \right) \left(\frac{h_0}{h_{-1}} \right)^{N_r} \right) \cdot \\ &\qquad \left((-\theta)^{-N_r} \prod_{j=0}^{N_r-1} x_j \right)^t \cdot \left((-\theta)^{-\frac{1}{2}t(t+1)} \left(\prod_{j=0}^{t-1} x_j^{t-j} \right) h_0 \left(\frac{h_0}{h_{-1}} \right)^t \right) \bmod p^r \\ &\equiv \left(\frac{h_{N_r}}{h_0} \right) \gamma_r^t\, h_t \bmod p^r \\ &\equiv -\beta_r\, \gamma_r^t\, h_t \bmod p^r. \end{aligned}$$

□

Theorem 9.3.2. *For all integers t and s,*

$$\frac{h_{t+sN_r}}{h_t} \equiv \left(\gamma_r^{N_r}\right)^{\frac{1}{2}s(s-1)} \gamma_r^{st} (-\beta_r)^s \bmod p^r.$$

Proof: We use induction on s. By Lemma 9.3.1,

$$\frac{h_{t+N_r}}{h_t} \equiv \gamma_r^t (-\beta_r) \bmod p^r \quad \text{for all } t \in \mathbb{Z}, \tag{9.2}$$

so the result holds for $s = 1$, and it holds trivially for $s = 0$. Suppose it holds for $0, 1, 2, \ldots, s-1$ for some $s > 0$. We prove it holds for s.

Replacing t by $t+(s-1)N_r$ in (9.2) and then using the induction hypothesis, we get

$$\begin{aligned} h_{t+sN_r} &\equiv \gamma_r^{t+(s-1)N_r}\,(-\beta_r)\,h_{t+(s-1)N_r} \bmod p^r \\ &\equiv \gamma_r^{t}\,\left(\gamma_r^{N_r}\right)^{s-1}\,(-\beta_r)\cdot\left(\gamma_r^{N_r}\right)^{\frac{1}{2}(s-1)(s-2)}\,\gamma_r^{(s-1)t}\,(-\beta_r)^{s-1}\,h_t \bmod p^r \\ &\equiv \left(\gamma_r^{N_r}\right)^{\frac{1}{2}s(s-1)}\,\gamma_r^{st}\,(-\beta_r)^{s}\,h_t \bmod p^r, \end{aligned}$$

as required. This establishes the result for $s > 0$.

To prove the result for $s < 0$, assume it holds for $0, -1, -2 \ldots, s+1$. Replacing t by $t + sN_r$ in (9.2), we get

$$h_{t+(s+1)N_r} \equiv \gamma_r^{t+sN_r}\,(-\beta_r)\,h_{t+sN_r} \bmod p^r.$$

Rearranging and using the induction hypothesis, we get

$$\begin{aligned} h_{t+sN_r} &\equiv \gamma_r^{-t-sN_r}\,(-\beta_r)^{-1}\,h_{t+(s+1)N_r} \bmod p^r \\ &\equiv \gamma_r^{-t}\,\left(\gamma_r^{N_r}\right)^{-s}\,(-\beta_r)^{-1}\cdot\left(\gamma_r^{N_r}\right)^{\frac{1}{2}(s+1)(s)}\,\gamma_r^{(s+1)t}\,(-\beta_r)^{s+1}\,h_t \bmod p^r \\ &\equiv \left(\gamma_r^{N_r}\right)^{\frac{1}{2}s(s-1)}\,\gamma_r^{st}\,(-\beta_r)^{s}\,h_t \bmod p^r, \end{aligned}$$

as required. □

It is easy to prove, by setting in turn $t = -N_r$, $t = 0$ and $t = -\frac{N_r}{2}$ in Lemma 9.3.1, that we have

Corollary 9.3.3. *The constants γ_r and β_r are related by*

$$\beta_r^2 \equiv \gamma_r^{N_r}\,\left(\frac{h_{N_r}}{h_{-N_r}}\right) \bmod p^r.$$

In fact, if N_r is even,

$$\beta_r \equiv \gamma_r^{\frac{N_r}{2}}\,\left(\frac{h_{\frac{N_r}{2}}}{h_{-\frac{N_r}{2}}}\right) \bmod p^r.$$

We also have

Corollary 9.3.4. *For all $t \in \mathbb{Z}$, the constants γ_r and β_r satisfy*

$$\frac{h_{t+N_r}\,h_{t-N_r}}{h_t^2} \equiv \gamma_r^{N_r} \bmod p^r \qquad \textit{and} \qquad \frac{h_{t+N_r}}{h_t}\cdot\frac{h_{-t+N_r}}{h_{-t}} \equiv \beta_r^2 \bmod p^r.$$

We can now justify using the same notation for our new constants β_r, γ_r as we used for the constants β_r, γ_r we defined in the last chapter:

Corollary 9.3.5. *For any $k \in \mathbb{Z}$, we have*

$$\frac{h_{N_r+k-1}}{h_{k-1}} \cdot \frac{h_{k-2}}{h_{N_r+k-2}} \equiv \gamma_r \bmod p^r,$$

and

$$-\gamma_r^{-k+1} \cdot \frac{h_{k+N_r-1}}{h_{k-1}} \equiv \beta_r \bmod p^r.$$

9.4 Periodicity in $(h_n \bmod p^r)$

In this section we prove that the sequence $(h_n \bmod p^r)$ is periodic for all $n \in \mathbb{Z}$, and describe how the period increases with r.

Definition: We define a constant τ_r to be the smallest positive integer such that

$$\gamma_r^{\tau_r} \equiv 1 \bmod p^r$$

and

$$(-\beta_r)^{\tau_r} \equiv \gamma_r^{-\frac{1}{2}N_r\tau_r(\tau_r-1)} \bmod p^r.$$

Theorem 9.4.1. *$(h_n \bmod p^r)$ is periodic with period $\pi_r = \tau_r N_r$.*

Proof: By Theorem 9.3.2, for all integers t,

$$\begin{aligned} h_{t+\tau_r N_r} &\equiv \left(\gamma_r^{N_r}\right)^{\frac{1}{2}\tau_r(\tau_r-1)} \left(\gamma_r^{\tau_r}\right)^t (-\beta_r)^{\tau_r} h_t \bmod p^r \\ &\equiv h_t \bmod p^r. \end{aligned}$$

It follows that $(h_n \bmod p^r)$ is periodic, with period π_r dividing $\tau_r N_r$.

Since $h_{t+\pi_r} \equiv h_t \bmod p^r$ and $x_t = -\theta \cdot \frac{h_{t-1}\,h_{t+1}}{h_t^2}$ for all $t \in \mathbb{Z}$, it follows that

$$x_{t+\pi_r} \equiv x_t \bmod p^r \quad \text{for all } t \in \mathbb{Z}.$$

So π_r must be a multiple of N_r, the order of $P \bmod p^r$ in $E(\mathbb{Z}_{p^r})$ — say $\pi_r = v\,N_r$.

Now since

$$h_{t+v\,N_r} \equiv \left(\gamma_r^{N_r}\right)^{\frac{1}{2}v(v-1)} \gamma_r^{v\,t}\,(-\beta_r)^v\,h_t \equiv h_t \bmod p^r$$

for all integers t, we have

$$(-\beta_r)^v \left(\gamma_r^{N_r}\right)^{\frac{1}{2}v(v-1)} \gamma_r^{v\,t} \equiv 1 \bmod p^r$$

for all integers t. It follows (setting $t = 2$ and $t = 1$ and dividing) that $\gamma_r^v \equiv 1 \bmod p^r$, and hence that $(-\beta_r)^v \left(\gamma_r^{N_r}\right)^{\frac{1}{2}v(v-1)} \equiv 1 \bmod p^r$. Hence $v \geq \tau_r$, and so $v = \tau_r$. □

So the sequence $(h_n \bmod p^r)$ starts repeating when the sequence of points $Q + [n]P$ has already repeated τ_r times.

By the same argument used in the proof of Lemma 8.5.2, we have

Lemma 9.4.2. *The integers y which satisfy*

$$\gamma_r^y \equiv 1 \bmod p^r$$

and

$$(-\beta_r)^y \equiv \gamma_r^{-\frac{1}{2}y(y-1)N_r} \bmod p^r$$

are precisely the multiples of τ_r.

If p is odd, then we can write τ_r in terms of the order of γ_r and $-\beta_r$:

Theorem 9.4.3. *Let p be an odd prime, and let n_r and ν_r be the orders of γ_r and $(-\beta_r)$ respectively modulo p^r. Then*

$$\tau_r = \begin{cases} \frac{1}{2}\operatorname{lcm}(n_r, \nu_r) & \text{if } N_1 \text{ is odd and } \frac{\nu_r}{2} \text{ and } n_r \text{ are even} \\ & \text{and divisible by exactly the same power of } 2, \text{ or} \\ 2\operatorname{lcm}(n_r, \nu_r) & \text{if } N_1 \text{ is odd and } n_r \text{ is divisible by} \\ & \text{at least as high a power of } 2 \text{ as } \nu_r \text{ is, or} \\ \operatorname{lcm}(n_r, \nu_r) & \text{otherwise.} \end{cases}$$

Again, this is proved in the same way as Theorem 8.5.3 in the previous chapter.

Corollary 9.4.4. *If p is an odd prime, then τ_1 divides $p-1$ if and only if N_1 is even or $\frac{h_{-N_1}}{h_{N_1}}$ is a quadratic residue modulo p. Otherwise τ_1 divides $2\,(p-1)$.*

Proof: Since $\mathrm{lcm}(n_1, \nu_1) \mid (p-1)$, it follows from Theorem 9.4.3 that $\tau_1 \nmid (p-1)$ if and only if N_1 is odd and n_1 is divisible by the same power of 2 as $p-1$, i.e., if and only if N_1 is odd and γ_1 is a quadratic non-residue modulo p. Since $\gamma_1^{N_1} \equiv \beta_1^2 \left(\frac{h_{-N_1}}{h_{N_1}}\right) \bmod p$, this is true if and only if N_1 is odd and $\frac{h_{-N_1}}{h_{N_1}}$ is a quadratic non-residue modulo p. □

9.4.1 How the period of $(h_n \bmod p^r)$ increases with r

We now consider the behaviour of τ_r as r increases.

Theorem 9.4.5. *If p is an odd prime then there exists a positive integer $u \le w$ such that*

$$\tau_r = \begin{cases} \tau_1 & \text{for } r \le u \\ p^{r-u}\tau_1 & \text{for } u \le r \le w \\ p\,\tau_w,\ \tau_w \text{ or } \frac{\tau_w}{p} & \text{for } r = w+1 \\ \tau_{w+1} & \text{for } r \ge w+1. \end{cases}$$

If $p \nmid \tau_w$ (i.e., if $w = u$) then τ_{w+1} is either τ_w or $p\,\tau_w$, and if $p \parallel \tau_w$ (i.e., if $w = u+1$) then τ_{w+1} is either τ_w or $\frac{\tau_w}{p}$. Otherwise $\tau_{w+1} = \tau_w$.

Proof: We first consider the case $2 \le r \le w$. Then $N_r = N_{r-1} = N_1$, so by definition of β_r and γ_r,

$$\gamma_r \equiv \gamma_w \bmod p^r \qquad \text{and} \qquad \beta_r \equiv \beta_w \bmod p^r.$$

Let g be a generator of $\mathbb{Z}_{p^w}$; so g is also a generator of $\mathbb{Z}_{p^r}^*$ for each $r \le w$. Let $\gamma_w \equiv g^s \bmod p^w$ and $\beta_w \equiv g^t \bmod p^w$. For $r \in \mathbb{N}$ let γ_r and β_r have order n_r and ν_r respectively.

If $p \mid n_r$ then (since $\gamma_r \equiv \gamma_w \bmod p^r$ for all $r \le w$) by Theorem 2.2.5 $n_{r-1} = \frac{n_r}{p}$; else $n_{r-1} = n_r$. Similarly, $\nu_{r-1} = \frac{\nu_r}{p}$; else $\nu_{r-1} = \nu_r$. It follows from Theorem 9.4.3 that as r decreases from w we have $\tau_{r-1} = \frac{\tau_r}{p}$ until r reaches a value u for which *neither* n_r nor ν_r is divisible by p, and from then on we have $\tau_r = \tau_{r-1}$.

We now consider the case $r \ge w+1$, i.e., $N_r = pN_{r-1}$. We first find an upper bound for τ_r. If either $r = w+1$ and $p \mid \tau_w$ or $r \ge w+2$, then using either

Theorems 9.2.3 and 2.2.5 or Theorem 9.2.2 we have

$$(-\beta_r)^{\tau_{r-1}} \equiv \left(\left(\gamma_{r-1}^{\tau_{r-1}}\right)^{\frac{1}{2}(p-1)\,N_{r-1}} \, (-\beta_{r-1})^{\tau_{r-1}} \right)^p \bmod p^r. \tag{9.3}$$

But by definition of τ_{r-1},

$$\gamma_{r-1}^{\tau_{r-1}} \equiv 1 \bmod p^{r-1} \quad \text{and} \quad (-\beta_{r-1})^{\tau_{r-1}} \equiv \left(\gamma_{r-1}^{-\frac{1}{2}\tau_{r-1}(\tau_{r-1}-1)}\right)^{N_{r-1}} \bmod p^{r-1}.$$

Since $\gamma_{r-1}^{\frac{1}{2}\tau_{r-1}(\tau_{r-1}-1)} \equiv \pm 1 \bmod p^{r-1}$ and N_r has the same parity as N_{r-1}, we can also write this as

$$\gamma_{r-1}^{\tau_{r-1}} \equiv 1 \bmod p^{r-1} \quad \text{and} \quad (-\beta_{r-1})^{\tau_{r-1}} \equiv \left(\gamma_{r-1}^{-\frac{1}{2}\tau_{r-1}(\tau_{r-1}-1)}\right)^{N_r} \bmod p^{r-1},$$

and it follows by Theorem 2.2.5 and (9.3) that

$$(-\beta_r)^{\tau_{r-1}} \equiv \left(\gamma_{r-1}^{-\frac{1}{2}N_r\tau_{r-1}(\tau_{r-1}-1)}\right)^p \bmod p^r.$$

Since by Theorem 9.2.1 $\gamma_r \equiv \gamma_{r-1}^p \bmod p^r$, it follows that

$$(-\beta_r)^{\tau_{r-1}} \equiv \gamma_r^{-\frac{1}{2}\tau_{r-1}(\tau_{r-1}-1)N_r} \bmod p^r,$$

and hence by Lemma 9.4.2 that $\tau_r \mid \tau_{r-1}$.

For $r = w+1$ and $p \nmid \tau_w$, we have instead of (9.3),

$$(-\beta_{w+1})^{p\,\tau_w} \equiv \left((\gamma_w^{\tau_w})^{\frac{1}{2}(p-1)\,N_w} \, (-\beta_w) \right)^p \bmod p^{w+1}, \tag{9.4}$$

and it follows in a similar way that $\tau_{w+1} \mid p\,\tau_w$.

We now prove a lower bound on τ_r. Note that π_r is a multiple of π_{r-1} (since if the sequence repeats modulo p^r then it must repeat modulo p^{r-1}). It follows that $p\,N_{r-1}\,\tau_r$ is a multiple of $N_{r-1}\,\tau_{r-1}$, and hence that τ_r is a multiple of $\frac{\tau_{r-1}}{p}$ if $p \mid \tau_{r-1}$, or of τ_{r-1} otherwise.

If either $r = w+1$ and $p \mid \tau_{w+1}$ or $r \geq w+2$, then we can improve this lower bound to $\tau_{r-1} \mid \tau_r$. Using either Theorems 9.2.3 and 2.2.5 or Theorem 9.2.2 we have

$$(-\beta_r)^{\tau_r} \equiv \left(\gamma_{r-1}^{\frac{1}{2}(p-1)\,N_{r-1}} \cdot (-\beta_{r-1}) \right)^{p\,\tau_r} \bmod p^r. \tag{9.5}$$

But by definition of τ_r,

$$\gamma_{r-1}^{p\,\tau_r} \equiv \gamma_r^{\tau_r} \equiv 1 \bmod p^r$$

and (since $N_r = p\,N_{r-1}$)

$$(-\beta_r)^{\tau_r} \equiv \left(\gamma_r^{-\frac{1}{2}\tau_r(\tau_r-1)}\right)^{p\,N_{r-1}} \bmod p^r.$$

It follows from (9.5) that

$$\left(\gamma_r^{-\frac{1}{2}\tau_r(\tau_r-1)}\right)^{p\,N_{r-1}} \equiv (-\beta_{r-1})^{p\,\tau_r} \bmod p^r,$$

and hence by Theorem 2.2.5 that

$$(-\beta_{r-1})^{\tau_r} \equiv \gamma_r^{-\frac{1}{2}N_{r-1}\tau_r(\tau_r-1)} \bmod p^{r-1}.$$

So by Lemma 9.4.2, $\tau_{r-1} \mid \tau_r$.

We have proved that $\tau_r = \tau_{r-1}$ for all $r \geq w+2$. Moreover, if $p \mid \tau_w$ then $\frac{\tau_w}{p} \mid \tau_{w+1} \mid \tau_w$, and if $p \mid \tau_{w+1}$ then $\tau_w \mid \tau_{w+1} \mid p\,\tau_w$. If p divides neither τ_w nor τ_{w+1} then $\tau_{w+1} = \tau_w$. It follows that if τ_w and τ_{w+1} are both divisible by p, or both coprime to p, then $\tau_{w+1} = \tau_w$. If p divides τ_w but not τ_{w+1}, then $p \parallel \tau_w$ and $\tau_{w+1} = \frac{\tau_w}{p}$. Similarly, if p divides τ_{w+1} but not τ_w, then $\tau_{w+1} = p\,\tau_w$.

Finally, we note that $\tau_1 \mid 2(p-1)$ implies $p \nmid \tau_1$, so the conditions $p \nmid \tau_w$ and $p \parallel \tau_w$ are equivalent to the conditions $w = u$ and $w = u+1$ respectively. The result follows. □

Since $N_r = N_1$ for $r \leq w$, $N_r = p^{r-w}\,N_1$ for $r \geq w$ and $\pi_r = N_r\,\tau_r$ for all $r \in \mathbb{N}$, Theorem 9.4.5 leads to the following result:

Theorem 9.4.6. *If p is an odd prime, then there exists a positive integer $1 \leq u \leq w$ such that for $r \in \mathbb{N}$ the sequence $(h_n \bmod p^r)$ has period*

$$\pi_r = \begin{cases} \pi_1 & \text{if } r \leq u, \\ p\,\pi_{r-1} & \text{if } u+1 \leq r \leq w \\ p^2\,\pi_w,\ p\,\pi_w \text{ or } \pi_w & \text{if } r = w+1 \\ p\,\pi_{r-1} & \text{if } r \geq w+2. \end{cases}$$

Moreover, $\pi_{w+1} = p\,\pi_w$ unless $\pi_w = \pi_1$ or $p\,\pi_1$. If $\pi_w = \pi_1$ (i.e., $w = u$) then π_{w+1} is either $p\,\pi_1$ or $p^2\,\pi_1$, and if $\pi_w = p\,\pi_1$ (i.e., $w = u+1$) then π_{w+1} is either $p\,\pi_w$ or π_w.

Although we have not yet been able to prove that $\pi_{w+1} = p\,\pi_w$ for all Somos 4 sequences (h_n), we suspect that this is the case; in other words, that as r increases from 1 the period of $(h_n \bmod p^r)$ remains the same until r reaches some value $u \leq w$, and then increases by a factor of p each time.

Robinson's Conjecture 6.7.5 is that if (h_n) is Somos(4) and p is an odd prime then $\pi_r = p\,\pi_{r-1}$ for every $r \geq 2$. We have proved that this is true for every odd prime p not dividing any terms of Somos(4) if and only if $\pi_2 = p\,\pi_1$ and $\pi_{w+1} = p\,\pi_w$, but we have not eliminated the possibility that $\pi_{w+1} = \pi_w$ or $p^2\,\pi_w$ for some of these primes.

Finally, we note that we have not proved any results about primes which divide λ_1 but not any term of (h_n). We have also not covered sequences in which some term is divisible by p^w but no term is divisible by p^{w+1} (as is the case for $p = 2$ in Somos(4)); since for $r > w$ no term h_k is divisible by 2^r, the results in chapter 8 do not apply, but some term is divisible by p so the results in this chapter do not apply either.

Bibliography

[1] Eric Bach and Jeffrey Shallit. *Algorithmic Number Theory, Volume 1: Efficient Algorithms*. MIT Press, 1996.

[2] I. Blake, G. Seroussi, and N. Smart. *Elliptic Curves in Cryptography*. Cambridge University Press, 1999.

[3] M. Bousquet-Mélou, J. Propp, and J.West. "Matchings graphs for Gale-Robinson recurrences". In preparation.

[4] L.S. Charlap and D.P. Robbins. "An elementary introduction to elliptic curves". Technical Report 31, Institute for Defense Analysis, Princeton, December 1988. Available at www.idaccr.org/reports/reports.html.

[5] D.V. Chudnovsky and G.V. Chudnovsky. "Sequences of numbers generated by addition in formal groups and new primality and factorization tests". *Advances in Applied Mathematics*, 7:385–434, 1986.

[6] Ian Connell. *Elliptic Curve Handbook*. Available at www.math.mcgill.ca/connell/public/ECH1/.

[7] M. Einsiedler, G. Everest, and T. Ward. "Primes in elliptic divisibility sequences". *LMS Journal of Computation and Mathematics*, 4:1–13, 2001.

[8] G. Everest, V. Miller, and N. Stephens. "Primes generated by elliptic curves". Preprint. 2003.

[9] G. Everest and T. Ward. "The canonical height of an algebraic point on an elliptic curve". *New York Journal of Mathematics*, 6:331–342, 2000.

[10] G. Everest and T. Ward. "Primes in divisibility sequences". *Cubo Matematica Educacional*, 3:245–259, 2001.

[11] Graham Everest. "Zsigmondy's Theorem for elliptic curves". Preprint. 2002.

[12] David Gale. "Somos sequence update". *Mathematical Intelligencer*, 13(4):49–50, 1991.

[13] David Gale. "The strange and surprising saga of the Somos sequences". *Mathematical Intelligencer*, 13(1):40–42, 1991.

[14] Andy Hone. "Elliptic curves and quadratic recurrence sequences". Preprint. 2003.

[15] H.W. Lenstra Jr. "Elliptic Curves and Number-Theoretic Algorithms". *Proceedings of the International Congress of Mathematicians*, pages 99–120, 1986.

[16] A. Menezes, T. Okamoto, and S. Vanstone. "Reducing Elliptic Curve Logarithms to a Finite Field". *IEEE Transaction on Information Theory*, 39:1639–1646, 1993.

[17] Alfred Menezes. *Elliptic Curve Public Key Cryptosystems*. Kluwer Academic Publishers, 1997.

[18] I. Niven, H. Zuckerman, and H. Montgomery. *An Introduction to the Theory of Numbers*. Wiley, 5th edition, 1991.

[19] James Propp. "Bilinear forum". See www.math.wisc.edu/ propp/ bilinear/welcome.

[20] James Propp. "Somos sequence site". See www.math.wisc.edu/ propp/ somos.html.

[21] Raphael M. Robinson. "Periodicity of Somos sequences". *Proceedings of the AMS*, 116(3):613–619, November 1992.

[22] H.E. Rose. *A Course in Number Theory*. Oxford Science Publications, 2nd edition, 1994.

[23] Rachel Shipsey. *Elliptic Divisibility Sequences*. PhD thesis, Goldsmiths, University of London, 2001.

[24] J. Silverman and J. Tate. *Rational Points on Elliptic Curves*. Undergraduate Texts in Mathematics. Springer, 1992.

[25] Joseph Silverman. *The Arithmetic of Elliptic Curves*. Number 106 in Graduate Texts in Mathematics. Springer-Verlag, 1986.

[26] Nigel Smart. "The discrete logarithm problem on elliptic curves of trace one". *Journal of Cryptography*, 1999.

[27] Michael Somos. "Problem 1470". *Crux Mathematicorum*, 15:208, 1989.

[28] Nelson Stephens. Personal communication. 2001-2.

[29] I. Stewart and D. Tall. *Algebraic Number Theory*. Mathematics Series. Chapman and Hall, 1987.

[30] Morgan Ward. "Memoir on Elliptic Divisibility Sequences". *American Journal of Mathematics*, 70:31–74, 1948.

[31] Morgan Ward. "The Law of Repetition of Primes in an Elliptic Divisibility Sequence". *Duke Mathematical Journal*, 15:941–946, 1948.

[32] Lawrence C. Washington. *Elliptic Curves: Number Theory and Cryptography*. Chapman and Hall, 2003.