Journal of Integer Sequences, Vol. 8 (2005),
Article 05.2.5

Elliptic Curves and Continued Fractions

Alfred J. van der Poorten[1]
Centre for Number Theory Research
1 Bimbil Place
Killara, Sydney NSW 2071
Australia
alf@math.mq.edu.au

Abstract

We detail the continued fraction expansion of the square root of the general monic quartic polynomial, noting that each line of the expansion corresponds to addition of the divisor at infinity. We analyse the data yielded by the general expansion. In that way we obtain "elliptic sequences" satisfying Somos relations. I mention several new results on such sequences. The paper includes a detailed "reminder exposition" on continued fractions of quadratic irrationals in function fields.

1 Introduction

A delightful essay [18] by Don Zagier explains why the sequence $(B_h)_{h\in\mathbb{Z}}$, defined by $B_{-2}=1$, $B_{-1}=1$, $B_0=1$, $B_1=1$, $B_2=1$ and the recursion

$$B_{h-2}B_{h+3}=B_{h+2}B_{h-1}+B_{h+1}B_h\,, \tag{1}$$

consists only of integers. Zagier comments that "the proof comes from the theory of elliptic curves, and can be expressed either in terms of the denominators of the co-ordinates of the multiples of a particular point on a particular elliptic curve, or in terms of special values of certain Jacobi theta functions."

In the present note I study the continued fraction expansion of the square root of a quartic polynomial, *inter alia* obtaining sequences generated by recursions such as (1). Here, however, it is clear that I have also constructed the co-ordinates of the shifted multiples of a point on a cubic curve and it is it fairly plain how to relate the surprising integer sequences and the curves from which they arise.

A brief reminder exposition on continued fractions in quadratic function fields appears as §6, starting at page 14 below.

It turns out that several of my results related to sequences à la (1) also appear in the recent thesis [16] of Christine Swart; for further comment see §5. Michael Somos, see [5], had *inter alia* asked for the inner meaning of the behaviour of the sequences (B_h), above, and of (C_h) defined by $C_{h-2}C_{h+2} = C_{h-1}C_{h+1} + C_h^2$ and $C_{-2} = 1$, $C_{-1} = 1$, $C_0 = 1$, $C_1 = 1$: of the sequences 5-Somos A006720 and 4-Somos A006721. More generally, of course, one may both vary the initial values and coefficients and generalise the "gap" to $2m$ or $2m+1$ by studying Somos $2m$ sequences, respectively Somos $2m+1$ sequences, namely sequences satisfying the respective recursions

$$D_{h-m}D_{h+m} = \sum_{i=1}^{m} \kappa_i D_{h-m+i}D_{h+m-i} \text{ or } D_{h-m}D_{h+m+1} = \sum_{i=1}^{m} \kappa_i D_{h-m+i}D_{h+m-i+1}\,.$$

I show in passing, a footnote on page 6, that a Somos 4 is always a Somos 6, while Theorem 2 points out it is always a Somos 5. After seeing [16], I added a somewhat painful proof, Theorem 4 on page 14, that it is also always a Somos 8. For example, 4-Somos satisfies all of

$$\begin{aligned} C_{h-3}C_{h+3} &= C_{h-1}C_{h+1} + 5C_h^2\,,\\ C_{h-2}C_{h+3} &= -C_{h-1}C_{h+2} + 5C_hC_{h+1}\,,\\ C_{h-4}C_{h+4} &= 25C_{h-1}C_{h+1} - 4C_h^2\,. \end{aligned}$$

In the light of such results one can be confident, see §3.3, page 6, that in general if (A_h) is a Somos 4, equivalently an elliptic sequence satisfying $A_{h-2}A_{h+2} = W_2^2A_{h-1}A_{h+1} - W_1W_3A_h^2$, then for all m both

$$W_1W_2A_{h-m}A_{h+m+1} = W_mW_{m+1}A_{h-1}A_{h+2} - W_{m-1}W_{m+2}A_hA_{h+1}$$

and

$$W_1^2A_{h-m}A_{h+m} = W_m^2A_{h-1}A_{h+1} - W_{m-1}W_{m+1}A_h^2\,.$$

2 Continued Fraction Expansion of the Square Root of a Quartic

We suppose the base field $\mathbb{F}$ is not of characteristic 2 because that case requires nontrivial changes throughout the exposition and not of characteristic 3 because that requires some trivial changes to parts of the exposition. We study the continued fraction expansion of a quartic polynomial $D \in \mathbb{F}[X]$; where D is not a square and for convenience suppose D to be monic and with zero trace. Set

$$\mathcal{C} : Y^2 = D(X) := (X^2 + f)^2 + 4v(X - w), \tag{2}$$

and for brevity write $A = X^2 + f$ and $R = v(X - w)$. Set $\overline{Y}$ for the conjugate $-Y$ of Y. For $h = 0, 1, 2, \ldots$ we denote the complete quotients of Y_0 by

$$Y_h = (Y + A + 2e_h)/v_h(X - w_h)\,, \tag{3}$$

noting that[2] the Y_h all are reduced, namely $\deg Y_h > 0$ but $\deg \overline{Y}_h < 0$. The upshot is that the h-th line of the continued fraction expansion of Y_0 satisfies

$$\frac{Y + A + 2e_h}{v_h(X - w_h)} = \frac{2(X + w_h)}{v_h} - \frac{\overline{Y} + A + 2e_{h+1}}{v_h(X - w_h)}.$$

Thus evident recursion formulas, see (32) at page 14, yield

$$f + e_h + e_{h+1} = -w_h^2 \tag{4}$$

and $-v_h v_{h+1}(X - w_h)(X - w_{h+1}) = (Y + A + 2e_{h+1})(\overline{Y} + A + 2e_{h+1})$. Hence

$$v_h v_{h+1}(X - w_h)(X - w_{h+1}) = -4(X^2 + f + e_{h+1})e_{h+1} + 4v(X - w). \tag{5}$$

Equating coefficients in (5), and then dividing by $-4e_{h+1} = v_h v_{h+1}$, we get

$$\begin{aligned} -4e_{h+1} &= v_h v_{h+1}\,; && X^2: \\ v/e_{h+1} &= w_h + w_{h+1}\,; && X^1: \\ f + e_{h+1} + vw/e_{h+1} &= w_h w_{h+1}\,. && X^0: \end{aligned}$$

The five displayed equations immediately above readily lead by several routes to

$$e_h e_{h+1} = v(w - w_h)\,. \tag{6}$$

For example, apply the remainder theorem to the right hand side of (5) after noting it is divisible by $X - w_h$, and recall (4).

Theorem 1. *Denote the two points at infinity on the quartic curve (2) by S and O, with O the zero of its group law. The points $M_{h+1} := (w_h, e_h - e_{h+1})$ all lie on $\mathcal{C}$. Set $M_1 = M$, and $M_{h+1} =: M + S_h$. Then $S_h = hS$.*

Proof. The points M_{h+1} lie on the curve $\mathcal{C} : Y^2 = D(X)$ because

$$\begin{aligned}(e_h - e_{h+1})^2 - (w_h^2 + f)^2 = \bigl(e_h - (e_{h+1} + w_h^2 + f)\bigr)\bigl((e_h + w_h^2 + f) - e_{h+1}\bigr)& \\ = -4e_h e_{h+1} = 4v(w_h - w)\,.&\end{aligned}$$

The birational transformations

$$X = \bigl(V - v\bigr)/U, \qquad Y = 2U - (X^2 + f)\,; \tag{7}$$

conversely,

$$2U = Y + X^2 + f\,, \qquad 2V = XY + X^3 + fX + 2v\,, \tag{8}$$

[2]Conditions apply. For instance we must choose the constants e_0 and w_0 so that $X - w_0$ divides the norm $(Y + A + 2e_0)(\overline{Y} + A + 2e_0)$ of its numerator. See page 14 for more details.

Note also that I presume that the v_h, equivalently the e_h, all are nonzero. See comments in §3.3, page 6 on the "singular case".

move the point S to $(0,0)$, leave O at infinity, and change the quartic model to a Weierstrass model

$$\mathcal{W} : V^2 - vV = U^3 - fU^2 + vwU\,. \tag{9}$$

Specifically, one sees that $U(M_{h+1}) = -e_{h+1}$, and $V(M_{h+1}) = v - w_h e_{h+1}$. We also note that $U(-M_{h+1}) = -e_{h+1}$, $V(-M_{h+1}) = w_h e_{h+1}$.

To check $S + (M + S_{h-1}) = M + S_h$ on $\mathcal{W}$ it suffices for us to show that the three points $(0,0)$, $(-e_h, v - w_{h-1}e_h)$, and $(-e_{h+1}, w_h e_{h+1})$ lie on a straight line. But that is $(v - w_{h-1}e_h)/e_h = w_h$. So $w_{h-1} + w_h = v/e_h$ proves the claim. □

One might view the preceding discussion culminating in Theorem 1 as making explicit the argument of Adams and Razar [1].

3 Elliptic sequences

Theorem 2. *Let (A_h) be the sequence defined by the "initial" values A_0, A_1 and the recursive definition*

$$A_{h-1}A_{h+1} = e_h A_h^2\,. \tag{10}$$

Then, given A_0, A_1, A_2, A_3, A_4 satisfying (10), the recursive definition

$$A_{h-2}A_{h+2} = v^2 A_{h-1}A_{h+1} + v^2(f + w^2)A_h^2 \tag{11}$$

defines the same sequence as does (10). Just so, also

$$A_{h-2}A_{h+3} = -v^2(f + w^2)A_{h-1}A_{h+2} + v^3\big(v + 2w(f + w^2)\big)A_h A_{h+1} \tag{12}$$

defines that sequence.

Proof. By (6) we obtain

$$\begin{aligned} e_{h-1}e_h^2 e_{h+1} &= v^2(w - w_{h-1})(w - w_h) \\ &= v^2(w_{h-1}w_h - w(w_{h-1} + w_h) + w^2) = v^2\left((f + e_h + vw/e_h) - w\cdot(v/e_h) + w^2\right). \end{aligned}$$

Thus

$$e_{h-1}e_h^2 e_{h+1} = v^2\left(e_h + (f + w^2)\right). \tag{13}$$

However, $A_{h-1}A_{h+1} = e_h A_h^2$ entails

$$A_{h-2}A_h A_{h-1}A_{h+1}A_h A_{h+2} = e_{h-1}e_h e_{h+1}A_{h-1}^2 A_h^2 A_{h+1}^2\,,$$

and so $A_{h-2}A_{h+2} = e_{h-1}e_h e_{h+1}A_{h-1}A_{h+1}$, which is

$$A_{h-2}A_{h+2} = e_{h-1}e_h^2 e_{h+1}A_h^2\,. \tag{14}$$

On multiplying (13) by A_h^2 we obtain (11).

Similarly (10) yields $A_{h-1}A_{h+1}A_h A_{h+2} = e_h e_{h+1}A_h^2 A_{h+1}^2$, and so

$$A_{h-1}A_{h+2} = e_h e_{h+1}A_h A_{h+1}\,. \tag{15}$$

It follows readily that

$$A_{h-2}A_{h+3} = e_{h-1}e_h^2e_{h+1}^2e_{h+2}A_hA_{h+1}\,. \tag{16}$$

Moreover, (13) implies that

$$e_{h-1}e_h^3e_{h+1}^3e_{h+2} = v^4\left(e_he_{h+1} + (f+w^2)(e_h+e_{h+1}) + (f+w^2)^2\right)\,.$$

However, by (4) we know that $v^2(e_h+e_{h+1}+f+w^2) = v^2(w^2-w_h^2)$. Here $v(w-w_h) = e_he_{h+1}$ and $v(w+w_h) = -v(w-w_h)+2vw = -e_he_{h+1}+2vw$. So

$$e_{h-1}e_h^2e_{h+1}^2e_{h+2} = v^2\left(-(f+w^2)e_he_{h+1} + v^2 + 2vw(f+w^2)\right), \tag{17}$$

which immediately allows us to see that also (12) yields the sequence (A_h). □

The extraordinary feature of the identities (13) and (17) is their independence of the translation M: thus of the initial data v_0, w_0, and e_0.

3.1 Two-sided infinite sequences

It is plain that the various definitions of the *elliptic sequence* (A_h) encourage one to think of it as bidirectional infinite. Indeed, albeit that one does feel a need to start a continued fraction expansion — so one conventionally begins it at Y_0, one is not stopped from thinking of the tableau listing the lines of the expansion as being two-sided infinite; note the remark at the end of § 6.1, page 14. In summary: we may and should view the various sequences (e_h), ..., defined above, as two-sided infinite sequences.

3.2 Vanishing

If say $v_k = 0$, then line k of the continued fraction expansion of Y_0 makes no sense both because the denominator $Q_k(X) := v_k(X-w_k)$ of the complete quotient Y_k seems to vanish identically and because the alleged partial quotient $a_k := 2(X+w_k)/v_k$ blows up.

The second difficulty is real. The vanishing of v_k entails a partial quotient blowing up to higher degree. We deal with vanishing by refusing to look at it. We move the point of impact of the issue by dismissing most of the data we have obtained, including the continued fraction tableau, and keep only a part of the sequence (e_h). That makes the first difficulty moot.[3]

Remark 3. There is no loss of generality in taking $k=0$. Then, up to an otherwise irrelevant normalisation, $Y_0 = Y + A$. If more than one of the v_h vanish then it is a simple exercise to confirm that the continued fraction expansion of $Y+A$ necessarily is purely periodic, see the discussion at page 17. If Y_0 does not have a periodic continued fraction expansion then there is some h_0, namely $h_0 = 0$, so that, for all $h > h_0$, line h of the expansion of Y_0 does make sense.

Except of course when dealing explicitly with periodicity, we suppose in the sequel that if $v_k = 0$ then $k = 0$; we refer to this case as the *singular* case.

[3] In any case, the first apparent difficulty is just an artifact of our notation. If, from the start, we had written $Q_h = v_hX + y_h$, as we might well have done at the cost of nasty fractions in our formulas, we would not have entertained the thought that $v_k = 0$ entails $y_k = 0$. Plainly, we must allow $v_k = 0$ yet $v_kw_k \neq 0$. Note, exercise, that $v_k(X-w_k)$ never does vanish identically.

3.3 The singular case

We remark that in the singular case the sequence $(e_h)_{h\geq 1}$ defines antisymmetric double-sided sequences (W_h), that is with $W_{-h}=-W_h$, by $W_{h-1}W_{h+1}=e_hW_h^2$ and so that, for all integers h, m, and n,

$$W_{h-m}W_{h+m}W_n^2+W_{n-h}W_{n+h}W_m^2+W_{m-n}W_{m+n}W_h^2=0\,. \tag{18$'$}$$

Actually, one may find it preferable to forego an insistence on antisymmetry in favour of rewriting (18′) less elegantly as

$$W_{h-m}W_{h+m}W_n^2=W_{h-n}W_{h+n}W_m^2-W_{m-n}W_{m+n}W_h^2\,, \tag{18}$$

just for $h\geq m\geq n$. In any case, (18) seems more dramatic than it is. An easy exercise confirms that, if $W_1=1$, (18) is equivalent to just

$$W_{h-m}W_{h+m}=W_m^2W_{h-1}W_{h+1}-W_{m-1}W_{m+1}W_h^2 \tag{19}$$

for all integers $h\geq m$. Indeed, (19) is just a special case of (18). However, given (19), obvious substitutions in (18) quickly show one may return from (19) to the apparently more general (18).

But there is a drama here. As already remarked in a near identical situation, the recurrence relation $W_{h-2}W_{h+2}=W_2^2W_{h-1}W_{h+1}-W_1W_3W_h^2$, and four nonzero initial values, already suffices to produce (W_h). Thus (19) for all m is apparently entailed by its special case $m=2$.

I can show this directly[4], by way of new relations on the e_h, for $m=3$. But the case $m=4$ already did not seem worth the effort. Whatever, my approach gave me no hint as to how to concoct an inductive argument leading to general m. Plan B, to look it up, fared little better. In her thesis [12], Rachel Shipsey shyly refers the reader back to Morgan Ward's opus [17]; but Ward does not comment on the matter at all, having *defined* his sequences by (19). Well, perhaps Ward does comment. The issue is whether (19) is coherent: do different m yield the one sequence? Ward notes that if σ is the Weierstraß σ-function then a sequence $\big(\sigma(hu)/\sigma(u)^{h^2}\big)$ satisfies (19) for all m. Whatever, a much more direct argument would be much more satisfying.[5]

[4]Plainly $e_{h-2}e_{h-1}^2e_h^3e_{h+1}^2e_{h+2}\cdot e_h=v^4\big(e_{h-1}+(f+w^2)\big)\big(e_{h+1}+(f+w^2)\big)e_h^2$. Now notice that $(e_{h-1}e_h+e_he_{h+1})e_h=v(w-w_{h-1}+w-w_h)e_h=2vwe_h-v^2$ and recall that $e_{h-1}e_h^2e_{h+1}=v^2\big(e_h+(f+w^2)\big)$. The upshot is a miraculous cancellation yielding

$$e_{h-2}e_{h-1}^2e_h^3e_{h+1}^2e_{h+2}\cdot e_h=v^4\big((f+w^2)^2e_h^2+v(v+2w(f+w^2))e_h\big)$$

and allowing us to divide by the auxiliary e_h. Thus the bottom line is

$$A_{h-3}A_{h+3}=v^4\big((f+w^2)^2A_{h-1}A_{h+1}+v(v+2w(f+w^2))A_h^2\big),$$

which is $A_{h-3}A_{h+3}=W_3^2A_{h-1}A_{h+1}-W_2W_4A_h^2$.

[5]For additional remarks, and a dissatisfying proof for the case $m=4$, see §5.2 on page 12.

Note Added in Proof: Christine Swart and the author have recently succeeded in applying the ideas of her thesis and of this paper to obtain a succinct inductive proof (thus, a much more satisfying direct argument) of the conjectures hinted at in §3.3, and stated at the end of §1.

Theorem 2 shows that certainly $W_{h-2}W_{h+2}=W_2^2W_{h-1}W_{h+1}-W_1W_3W_h^2$ for $h=1, 2, \ldots$, in which case (19) apparently follows by arguments in [17] and anti-symmetry; (18) is then just an easy exercise.

Up to multiplying the expansion by a constant, the singular case is initiated by $v_1=4v$, $w_1=w$, $e_1=0$, $e_2=-(f+w^2)$. For temporary convenience set $x=v/(f+w^2)$. From the original continued fraction expansion of $Y+A$ or, better, the recursion formulas of page 3, we fairly readily obtain $v_2=1/x$, $w_2=w-x$, $e_3=-x(x+2w)$, $e_4=v(x^2(x+2w)-v)/x^2(x+2w)^2$.

We are now free to choose, say $W_1=1$, $W_2=v$, leading to $W_3=-v^2(f+w^2)$, $W_4=-v^4\bigl(v+2w(f+w^2)\bigr)$, $W_5=-v^6\bigl(v(v+2w(f+w^2))-(f+w^2)^3\bigr)$, $\ldots$.

That allows us to notice that (12) apparently is

$$W_1W_2A_{h-2}A_{h+3}=W_2W_3A_{h-1}A_{h+2}-W_1W_4A_hA_{h+1}$$

and that (11) of course is

$$A_{h-2}A_{h+2}=W_2^2A_{h-1}A_{h+1}-W_1W_3A_h^2\,.$$

Given that also $A_{h-3}A_{h+3}=W_3^2A_{h-1}A_{h+1}-W_2W_4A_h^2$, it is of course tempting to guess that more is true. Certainly, more *is* true in the special case $(A_h)=(W_h)$, that's the point of the discussion above. Moreover, the same "more" is true, see for example [16, Theorem 8.1.2, p. 191], for sequence translates: thus $(A_h)=(W_{h+k})$.

3.4 Elliptic divisibility sequences

Recall that in the singular case and for $h=1, 2, \ldots$ the $-e_h$ are in fact the U co-ordinates of the multiples hS of the point $S=(0,0)$ on the curve $V^2-vV=U^3-fU^2+vwU$.

Suppose we are working in the ring $Z=\mathbb{Z}[f,v,vw]$ of "integers". If $\gcd(v,vw)=1$, so the exact denominator of the "rational" w is v, then our choices $W_1=1$, $W_2=v$ lead the definition $W_{h-1}W_{h+1}=e_hW_h^2$ to be such that W_h^2 is always the exact denominator of the "rational" e_h. It is this that is shown in detail by Rachel Shipsey [12]. In particular it follows that (W_h) is an elliptic divisibility sequence as described by Ward [17]. A convenient recent introductory reference is Chapter 10 of the book [3].

Set $hS=(U_h/W_h^2, V_h/W_h^3)$, thus defining sequences (U_h), (V_h), and (W_h) of integers, with W_h chosen minimally. Shipsey notices, provided that indeed $\gcd(v,vw)=1$, that $\gcd(U_h,V_h)=W_{h-1}$ and $W_{h-1}W_{h+1}=-U_h$. Here, I have used this last fact to define the sequence (W_h).

Starting, in effect, from the definition (18), Ward [17] shows that with $W_0=0$, $W_1=1$, and $W_2\big|W_4$, the sequence (W_h) is a *divisibility sequence*; that is, if $a\big|b$ then $W_a\big|W_b$. A little more is true. If also $\gcd(W_3,W_4)=1$ then in fact $\gcd(W_a,W_b)=W_{\gcd(a,b)}$. On the other hand, a prime dividing both W_3 and W_4 divides W_h for all $h\ge 3$.

3.5 Quasi-periodicity of the continued fraction expansion

Suppose now that the sequence (W_h) has a zero additional to its zero $W_0=0$. From the continued fraction expansion and, say, [8], we find that $v=0$ (but $w'=vw\neq 0$ if our curve is to be elliptic) is the case of the continued fraction having quasi-period $r = 1$ and the divisor at infinity on the curve having torsion $m = 2$. Just so, $f + w^2 = 0$, thus $W_3 = 0$, signals $r = 2$ and $m = 3$, and $x + 2w = 0$, or $W_4 = 0$, is $r = 3$ and $m = 4$. And so on; for more see [8]. In summary, $m > 0$ is minimal with $W_m = 0$ if and only if the continued fraction expansion of, say, $Z = \frac{1}{2}(Y + A)$ has a minimal quasi-period of length $r = m - 1$.

Note 1. It has been suggested in my hearing that "Mathematics is the study of degeneracy". Given that slogan, it is equivalent to $W_m = 0$ that the sequence (e_h) be periodic with period m. However, recall that $W_0 = 0$ and $W_{\pm 1} = \pm 1$ entails e_0 must be infinite. Just so then, $W_m = 0$ entails e_m infinite and $e_{m+1} = 0$. It also follows, see §6, that the word $e_1, e_2, \cdots, e_{m-1}$ is a palindrome. Note that, in effect, we define (W_h) by way of antisymmetry, its initial values $W_1 = 1$, W_2, $W_1W_3 = e_2W_2^2$, $W_2W_4 = e_3W_3^2$, and the recurrence relation $W_{h-2}W_{h+2} = W_2^2W_{h-1}W_{h+1} - W_1W_3W_h^2$ — plainly that relation allows us to "jump over" zeros of the sequence. Note that, in contrast, Christine Swart [16] declares her elliptic sequences as undefined beyond a 0.

As for the singular continued fraction expansion, our notation has us set $Z_1 := -Z/v(X - w)$. In consequence, we are committed to the line $Z = A - \overline{Z}$ in effect functioning as the *pair of* lines $m - 1$ and m; just so then it must also be the pair of lines $\overline{1}$ and 0. These matters are no great issue here; but they will matter in generalising the present work to hyperelliptic curves of higher genus.

The periodicity of (e_h) is necessary, but not at all sufficient for the periodicity of (W_h). Indeed, Ward [17] shows and one fairly readily confirms that precisely the periods 1, 2, 3, 4, 5, 6, 8, or 10 are possible for an integral elliptic divisibility sequence defined by (19).

4 Examples

4.1

Consider the curve $\mathcal{C} : Y^2 = (X^2 - 29)^2 - 4 \cdot 48(X + 5)$; here a corresponding cubic model is $\mathcal{E} : V^2 + 48V = U^3 + 29U^2 + 240U$. Set $A = X^2 - 29$. The first several preceding and succeeding steps in the continued fraction expansion of $Y_0 = (Y + A + 16)/8(X + 3)$ are[6]

$$
\begin{aligned}
\frac{Y + A + 18}{16(X+2)/3} &= \frac{X-2}{8/3} - \frac{\overline{Y} + A + 32}{16(X+2)/3} && \text{line } \overline{3}: \\
\frac{Y + A + 32}{12(X+1)} &= \frac{X-1}{6} - \frac{\overline{Y} + A + 24}{12(X+1)} && \text{line } \overline{2}: \\
\frac{Y + A + 24}{4(X+3)} &= \frac{X-3}{2} - \frac{\overline{Y} + A + 16}{4(X+3)} && \text{line } \overline{1}: \\
\frac{Y + A + 16}{8(X+3)} &= \frac{X-3}{4} - \frac{\overline{Y} + A + 24}{8(X+3)} && \text{line } 0: \\
\frac{Y + A + 24}{6(X+1)} &= \frac{X-1}{3} - \frac{\overline{Y} + A + 32}{6(X+1)} && \text{line } 1: \\
\frac{Y + A + 32}{32(X+2)/3} &= \frac{X-2}{16/3} - \frac{\overline{Y} + A + 18}{32(X+2)/3} && \text{line } 2: \\
\frac{Y + A + 18}{9(3X+10)/8} &= \cdots
\end{aligned}
$$

where elegance has suggested we write "line $\overline{h}$" as short for "line $-h$".

The feature motivating this example is the six integral points $(-2, \pm 7)$, $(-1, \pm 4)$, and $(-3, \pm 4)$ on $\mathcal{C}$. With $M_{\mathcal{C}} = (-3, 4)$ and $S_{\mathcal{C}}$ the "other" point at infinity these are in fact the six points $M_{\mathcal{C}} + hS_{\mathcal{C}}$ for $h = -3,\ -2,\ -1,\ 0,\ 1,$ and 2.

Correspondingly, on $\mathcal{E}$ we have the integral points $M + 2S = (-16, -32)$ and $M - 2S = (-16, -16)$, $M - S = (-12, -36)$ and $M + S = (-12, -12)$; here $M = M_{\mathcal{E}} = (-8, -24)$; $S = S_{\mathcal{E}} = (0, 0)$. Of course $\mathcal{E}$ is not minimal; nor, for that matter was $\mathcal{C}$. In fact the replacements $X \leftarrow 2X + 1$, $Y \leftarrow 4Y$ yield

$$Y^2 = (X^2 + X - 7)^2 - 4 \cdot 6(X + 3)\,, \tag{20}$$

correctly suggesting we need a more general treatment than that presented in the discussion above. It turns out to be enough for present purposes to replace $e_h \leftarrow 4e_h$ obtaining

$$\ldots,\ e_{-3} = \tfrac{9}{4},\ e_{-2} = 4,\ e_{-1} = 3,\ e_0 = 2,\ e_1 = 3,\ e_2 = 4,\ e_3 = \tfrac{9}{4},\ \ldots\,.$$

Then $A_0 = 1$, $A_1 = 1$ and

$$A_{h-1}A_{h+1} = e_h A_h^2$$

yields the sequence $\ldots$, $A_{-4} = 2^5 3^5$, $A_{-3} = 2^5 3^2$, $A_{-2} = 2^3 3$, $A_{-1} = 2$, $A_0 = 1$, $A_1 = 1$, $A_2 = 3$, $A_3 = 2^2 3^2$, $A_4 = 2^2 3^5$, $\ldots$. Notice that we're hit for six[7] by increasingly high powers of primes dividing 6 appearing as factors of the A_h. However, we know that (12) derives from (17). With the original e_hs divided by 4 that yields

$$A_{h-2}A_{h+3} = 6^2 A_{h-1}A_{h+2} + 6^3 A_h A_{h+1}\,.$$

[6]Here my choice of $v_0 = 8$ is arbitrary but not at random.

[7]My remark is guided by knowing that $V^2 + UV + 6V = U^3 + 7U^2 + 12U$ is a minimal model for $\mathcal{E}$, and noticing that $\gcd(6, 12) = 6$. Notice too that 6^2 divides the discriminant of this model.

Remarkably, one may remove the effect of the 6 by renormalising to a sequence (B_h) of integers satisfying

$$B_{h-2}B_{h+3} = B_{h-1}B_{h+2} + B_hB_{h+1}\,.$$

Specifically, $\dots,\ B_{-4}=3,\ B_{-3}=2,\ B_{-2}=1,\ B_{-1}=1,\ B_0=1,\ B_1=1,\ B_2=1,\ B_3=2,$ $B_4=3,\ B_5=5,\ B_6=11,\ B_7=37,\ B_8=83,\ \dots,$ and the sequence is symmetric about $B=0$. Interestingly, the choice of each B_h as a divisor of A_h is forced, in the present case by the data $A_0=A_1=1$ and the decision that the coefficient of B_hB_{h+1} be 1. Of course it is straightforward to verify that $A_{h-2}A_{h+3}$ is always divisible by 6^3 and $A_{h-1}A_{h+2}$ always by 6. For a different treatment see §4.4 below.

4.2

Take $v=\pm1$ and $f+w^2=1$. Thus, by (13), $e_{h-1}e_h^2e_{h+1}=e_h+1$ and so $e_0=1$, $e_1=1$ yields the sequence $\dots,\ 2,\ 1,\ 1,\ 2,\ 3/4,\ 14/9,\ \dots,$ of values of e_h. As explained above, with $C_0=1$ and $C_1=1$, the definition $C_{h-1}C_{h+1}=e_hC_h^2$ yields the symmetric sequence $\dots,\ 2,\ 1,\ 1,\ 1,\ 1,\ 2,\ 3,\ 7,\ 23,\ 59,\ \dots,$ of values of C_h satisfying the recursion

$$C_{h-2}C_{h+2} = C_{h-1}C_{h+1} + C_h^2\,.$$

Set $Y^2=A^2+4v(X-w)$, where $A=X^2+f$. Then the recursion for (C_h) entails $v^2=1$ and $f+w^2=1$. Plainly, one can get four consecutive values 1 in a sequence (C_h) only by having two consecutive values 1 in the corresponding sequence (e_h). Thus (4) yields $f+2=-w_0^2$ and then $f+w^2=1$ entails $w^2-w_0^2=3$. Hence $w^2=4$ and $f=-3$. The identity (6) implies vw is positive.

So $v=\pm1$, $w=\pm2$, $f=-3$. Up to $X\leftarrow -X$, the sequence (C_h) is given by the curve $\mathcal{C}: Y^2=(X^2-3)^2+4(X-2)$ and its points $M_{\mathcal{C}}+hS_{\mathcal{C}}$, $M_{\mathcal{C}}=(1,0)$, $S_{\mathcal{C}}$ the "other point" at infinity; equivalently by

$$\mathcal{E}: V^2-V=U^3+3U^2+2U \quad \text{with } M=(-1,1),\ S=(0,0).$$

Indeed, $M+S=(-1,0)$, $M+2S=(-2,1)$, $M+3S=(-3/4,3/8)$, $\dots$. Note that it is impossible to have three consecutive values 1 in the sequence (e_h) if also $v=\pm1$, except for trivial periodic cases, so the hoo-ha of the example at §4.1 above is in a sense unavoidable.

4.3 Remarks

The two examples get a rather woolly treatment in [15] and its preceding discussion; see [5] for context. Before seeing [16] I had also remarked that "the observation that a twist $V^2-vV=dU^3-fU^2+vwU$ becomes $V^2-dvV=U^3-fU^2+dvwU$ by $U\leftarrow dU$, $V\leftarrow dV$ allows one to presume $v=\pm1$. A suitable choice of e_0, e_1 and A_0, A_1 should now allow one to duplicate the result claimed in [15] in somewhat less brutal form." Namely, one wishes to obtain elliptic curves yielding a sequence (A_h) with nominated A_{-1}, A_0, A_1, A_2 and such that $A_{h-2}A_{h+2}=\kappa A_{h-1}A_{h+1}+\lambda A_h^2$; only the cases κ not a square are at issue. In fact, this issue is dealt with by Christine Swart at [16, p. 153*ff*] in more straightforward fashion than I had in mind. In very brief, if $A_{h-1}A_{h+1}=e_hA_h^2$, then

$$B_h=\kappa^{\frac12 h(h+1)}A_h \quad \text{entails} \quad B_{h-1}B_{h+1}=\kappa e_hB_h^2\,,$$

and so $B_{h-2}B_{h+2} = (1/\kappa)^2 B_{h-1}B_{h+1} + (\lambda/\kappa^4)B_h^2$. Yet more simply, one may remark that there always is an elliptic curve defined over the base field with $\sqrt{\kappa}$ adjoined; that is, over a quadratic twist — exactly as I had mooted.

4.4 Reprise

It seems appropriate to return to the example of §4.1 so as to *discover* the elliptic curve giving rise to $(B_h) = (\dots,1,1,1,1,1,\dots)$, given that $B_{h-2}B_{h+3} = B_{h-1}B_{h+2} + B_hB_{h+1}$. Recall we expect the squares of the integers B_h to be the precise denominators of the points $M + hS$ on the minimal Weierstraß model $\mathcal{W}$ of the curve; here M is some point on that model and $S = (0,0)$.

Suppose e_{-2}, e_{-1}, e_0, e_1, e_2 supply the five integer co-ordinates yielding B_{-2}, B_{-1}, B_0, B_1, B_2. Because no more than two of these e_i can be 1 we must have

$$e_0B_{-1}B_1 = e_0B_0^2\,, \quad e_1B_0B_2 = e_1B_1^2\,, \quad \tfrac{1}{2}e_2B_1B_3 = e_2B_2^2\,,$$

since of course the recursion for (B_h) entails $B_3 = 2$. Suppose in general that

$$c_hB_{h-1}B_{h+1} = e_hB_h^2\,.$$

Then the identity (17),

$$e_{h-1}e_h^2e_{h+1}^2e_{h+2} = v^2\left(-(f+w^2)e_he_{h+1} + v^2 + 2vw(f+w^2)\right),$$

and $B_{h-2}B_{h+3} = B_{h-1}B_{h+2} + B_hB_{h+1}$ entail

$$c_{h-1}c_h^2c_{h+1}^2c_{h+2} = -v^2(f+w^2)c_hc_{h+1} = v^3\big(v + 2w(f+w^2)\big).$$

Thus $c_hc_{h+1} = kv$, say, is independent of h and we have

$$k^2 = -(f+w^2) \quad \text{and} \quad k^3 + 2wk^2 - v = 0\,.$$

Note that if $f+w^2$, or $2w$ and v, are integers, also k must be an integer. Also,

$$e_0e_1 = vk \quad \text{and} \quad e_1e_2 = 2vk\,. \tag{21}$$

Remark 1. However, $e_{h-1}e_h^2e_{h+1} = v^2e_h + v^2(f+w^2)$ implies

$$k^2B_{h-2}B_{h+2} = c_hB_{h-1}B_{h+1} + (f+w^2)B_h^2\,,$$

without the coefficients necessarily being independent[8] of h. In particular, $k^2 = -(f+w^2)$ entails $c_0 = e_0 = 2k^2$ and $c_1 = e_1 = 3k^2$.

[8] Both Christine Swart and Andy Hone have pointed out to me that c_hc_{h+1} constant, and thus both c_{2h} and c_{2h+1} constant, of course boils down to a Somos 5 corresponding to a pair of interlinked Somos 4.

On the other hand, the identity (6) now reports that $k = w - w_0$ and $2k = w - w_1$. By (5) we then have

$$w_0 + w_1 = v/e_1 = 2w - 3k$$

whilst by (4) we see that $f + e_0 + e_1 = -(w-k)^2$, $f + e_1 + e_2 = -(w-2k)^2$, so, recalling that $e_2 = 2e_0$, also $e_0 = (2w-3k)k$. Hence

$$(2w-3k)k = 2k^2 \quad \text{and so} \quad 2w = 5k\,. \tag{22}$$

In summary, we then can quickly conclude that also

$$v = 6k^3\,, \quad 4f = -29k^2\,, \quad \text{and} \quad 2w_0 = 3k\,. \tag{23}$$

The normalisation $k = -2$ retrieves the continued fraction expansion in §4.1 on page 8. As shown in §7 on page 17 the corresponding minimal Weierstraß model is $V^2 + UV + 6V = U^3 + 7U^2 + 12U$; and $M = (-2,-2)$ is a point of order two.

5 Somos Sequences

5.1 Christine Swart's thesis [16]

Much of the work reported by me here is *sui generis* with original intent to make explicit the ideas of Adams and Razar [1] and to rediscover the rational elliptic torsion surfaces (thus, the "pencils" of rational elliptic curves with, say $(0,0)$, a torsion point of given order m) by a new method, see [8] and its references. Eventually, I learned of Michael Somos's sequences, see [5], and realised how they arise from my data. I had sort of heard of Christine Swart's work from Nelson Stephens in 2003 but, fortunately as it turned out[9], did not have access to her thesis [16] until very recently when this paper was already essentially complete; see [9].

Christine Swart's discussion of the interrelationship between elliptic sequences and elliptic curves is more detailed and complete than mine. Among many other things, she is careful to recognise that the formulas neither know nor care whether the given elliptic curve is in fact elliptic: thus, for example, also my quartics may have multiple zeros. If so, extra comment — mostly, quite straightforward — is required at a number of points above; but is neglected by me. Further, Christine Swart reports *inter alia* that Nelson Stephens (personal communication to her) had noticed, by completely different methods, identities equivalent to (13) and (11), see page 4; these are her Theorems 3.5.1 and 7.1.2 [16, p. 29 and p. 153].

5.2 Much more satisfying?

I can in fact show that an elliptic sequence (A_h) is also given by $A_{h-4}A_{h+4} = W_4^2 A_{h-1}A_{h+1} - W_3W_5A_h^2$. However, I consider my argument as typical of the sort of thing that gives mathematics its bad name: and regret to have to admit that this sort of nonsense seemingly does generalise to proving that a Somos 4 also always is a three-term Somos $(4+n)$.

[9] It was fun to puzzle out not just the answers to questions, but also to attempt to guess what the questions ought to be.

Specifically, we know that

$$A_{h-2}A_{h+2} = W_2^2 A_{h-1}A_{h+1} - W_1W_3A_h^2\,, \quad \text{by definition;} \tag{24}$$
$$A_{h-3}A_{h+3} = W_3^2 A_{h-1}A_{h+1} - W_2W_4A_h^2\,, \quad \text{see footnote page 6,} \tag{25}$$

and intend to show that

$$A_{h-4}A_{h+4} = W_4^2 A_{h-1}A_{h+1} - W_3W_5A_h^2\,. \tag{26}$$

Notice that, in particular, $W_1W_5 = W_2^2W_2W_4 - W_1W_3W_3^2$; so because $W_1 = 1$, $W_3^4 = W_2^3W_3W_4 - W_3W_5$. Now observe that

$$A_{h-4}A_{h+2}A_{h-2}A_{h+4} = (W_3^2A_{h-2}A_h - W_2W_4A_{h-1}^2)(W_3^2A_hA_{h+2} - W_2W_4A_{h+1}^2)\,.$$

In the product on the right, the first term is

$$(W_2^3W_3W_4A_h^2 - W_3W_5A_h^2)A_{h-2}A_{h+2}$$

and half of it contributes half of (26). Similarly, half the final term of the product, thus of

$$W_4^2A_{h-1}A_{h+1}\cdot W_2^2A_{h-1}A_{h+1} = W_4^2A_{h-1}A_{h+1}(A_{h-2}A_{h+2} + W_3A_h^2)\,,$$

provides the other half of (26). Thus it's ugly but true that we have proved that (26) holds if and only if $W_3W_4 = 0$ or

$$W_2^3A_{h-2}A_h^2A_{h+2} - W_2W_3(A_{h-2}A_hA_{h+1}^2 + A_{h-1}^2A_hA_{h+2}) + W_4A_{h-1}A_h^2A_{h+1} = 0\,.$$

I now compound this brutality by fiercely replacing the two occurrences of A_{h-2} by the evident relation

$$A_{h-2} = (W_2^2A_{h-1}A_{h+1} - W_1W_3A_h^2)/A_{h+2}\,.$$

That necessitates our then multiplying by A_{h+2}. Fortunately, we can compensate for this cruelty by dividing by A_h. We are left with needing to show that

$$\begin{aligned}W_2^5A_{h-1}A_hA_{h+1}A_{h+2} - W_3^2W_3A_h^3A_{h+2} - W_2^3W_3A_{h-1}A_{h+1}^3&\\ - W_2W_3A_{h-1}^2A_{h+2}^2 + W_2W_3^2A_h^2A_{h+1}^2 + W_4A_{h-1}A_hA_{h+1}A_{h+2} &= 0\,.\end{aligned} \tag{27}$$

It's now natural to despair, and to start looking for a Plan B. However, one might notice, on page 7, that $W_4 = -v^4\bigl(v + 2w(f+w^2)\bigr)$; and $W_2 = v$. Moreover $W_3 = -v^2(f+w^2)$. Thus, conveniently,

$$W_2^5 + W_4 = -2v^4w(f+w^2) = W_2^2W_3\,.$$

Hence, just as our result is trivial if $W_3W_4 = 0$, so also it is trivial if $W_2 = 0$. All this is a sign that we may not as yet have made an error. We may divide (27) by W_2W_3. Better yet, let's also divide by $A_h^2A_{h+1}^2$ by using the definitions

$$A_{h-1}A_{h+1} = e_hA_h^2\,, \quad \text{whence also} \quad A_{h-1}A_{h+2} = e_he_{h+1}A_hA_{h+1}\,.$$

Then all that remains is a confirmation that

$$2vwe_he_{h+1} - v^2(e_h + e_{h+1}) - e_h^2e_{h+1}^2 - v^2(f+w^2) = 0\,. \tag{28}$$

However (13), page 4, is $e_he_{h+1} = v(w - w_h)$, while $e_h + e_{h+1} = -f - w_h^2$ is (4), page 3. Astonishingly, the claim (28) follows immediately.

Theorem 4. *If (A_h) is a Somos 4 then it is a Somos 8 of the shape*

$$A_{h-4}A_{h+4} = \kappa A_{h-1}A_{h+1} + \lambda A_h^2\,.$$

Proof. Given the argument above, it suffices to note that any Somos 4 is equivalent to an elliptic sequence. $\square$

6 Rappels

6.1 Continued fraction expansion of a quadratic irrational

Let $Y = Y(X)$ be a quadratic irrational integral element of the field $\mathbb{F}((X^{-1}))$ of Laurent series

$$\sum_{h=-d}^{\infty} f_{-h}X^{-h}, \quad \text{some } d \in \mathbb{Z} \tag{29}$$

defined over some given base field $\mathbb{F}$; that is, there are polynomials T and D defined over $\mathbb{F}$ so that

$$Y^2 = T(X)Y + D(X)\,. \tag{30}$$

Plainly, by translating Y by a polynomial if necessary, we may suppose that $\deg D \geq 2\deg T + 2$, with $\deg D = 2g+2$, say, and $\deg T \leq g$; then $\deg Y = g+1$. Recall here that a Laurent series (29) with $f_d \neq 0$ has degree d.

Set $Y_0 = (Y+P_0)/Q_0$ where P_0 and Q_0 are polynomials so that Q_0 divides the norm $(Y+P_0)(\overline{Y}+P_0)$; notice here that an $\mathbb{F}[X]$-module $\langle Q, Y+P\rangle$ is an ideal in $\mathbb{F}[X,Y]$ if and only if $Q\big|(Y+P)(\overline{Y}+P)$.

Further, suppose that $\deg Y_0 > 0$ and $\deg \overline{Y}_0 < 0$; that is, Y_0 is *reduced*. Then the continued fraction expansion of Y_0 is given by a sequence of lines, of which the h-th is

$$Y_h := (Y+P_h)/Q_h = a_h - (\overline{Y}+P_{h+1})/Q_h\,; \quad \text{in brief} \quad Y_h = a_h - \overline{B}_h\,. \tag{31}$$

Here the polynomial a_h is a *partial quotient*, and the next *complete quotient* Y_{h+1} is the reciprocal of the preceding *remainder* $-(\overline{Y}+P_{h+1})/Q_h$. Plainly the sequences of polynomials (P_h) and (Q_h) are given by the recursion formulas

$$P_h + P_{h+1} + (Y+\overline{Y}\,) = a_hQ_h \text{ and } Y\overline{Y} + (Y+\overline{Y})P_{h+1} + P_{h+1}^2 = -Q_hQ_{h+1}\,. \tag{32}$$

It is easy to see by induction on h that Q_h divides the norm $(Y+P_h)(\overline{Y}+P_h)$.

We observe also that we have a conjugate expansion with h-th line

$$B_h := (Y+P_{h+1})/Q_h = a_h - (\overline{Y}+P_h)/Q_h\,, \quad \text{that is,} \quad B_h = a_h - \overline{Y}_h\,. \tag{33}$$

Note that the next line of this expansion is the conjugate of the previous line of its conjugate expansion: conjugation reverses a continued fraction tableau. Because the conjugate of line 0 is the last line of its tableau we can extend the expansion forming the conjugate tableau, leading to lines $h = 1,\ 2,\ \dots$

$$(Y+P_{-h+1})/Q_{-h} = a_{-h} - (Y+P_{-h})/Q_{-h}\,; \quad \text{that is,} \quad B_{-h} = a_{-h} - \overline{Y}_{-h}\,.$$

Plainly the original continued fraction tableau also is two-sided infinite and our thinking of it as "starting" at Y_0 is just convention.

6.2 Continued fractions

One writes $Y_0 = [\,a_0\,,\,a_1\,,\,a_2\,,\,\ldots]$, where formally

$$[\,a_0\,,\,a_1\,,\,a_2\,,\,\ldots\,,\,a_h\,] = a_0 + 1/\,[\,a_1\,,\,a_2\,,\,\ldots\,,\,a_{h-1}\,] \quad \text{and} \quad [\quad] = \infty\,. \tag{34}$$

It follows, again by induction on h, that the definition

$$\begin{pmatrix} a_0 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a_1 & 1 \\ 0 & 1 \end{pmatrix}\cdots\begin{pmatrix} a_h & 1 \\ 0 & 1 \end{pmatrix} =: \begin{pmatrix} x_h & x_{h-1} \\ y_h & y_{h-1} \end{pmatrix}$$

entails $[\,a_0\,,\,a_1\,,\,a_2\,,\,\ldots\,,\,a_h\,] = x_h/y_h$. This provides a correspondence between the *convergents* x_h/y_h and certain products of 2×2 matrices (more precisely, between the sequences (x_h), (y_h) of *continuants* and those matrices). It is a useful exercise to notice that $Y_0 = [\,a_0\,,\,a_1\,,\,\ldots\,,\,a_h\,,\,Y_{h+1}\,]$ implies that

$$Y_{h+1} = -(y_{h-1}Y - x_{h-1})/(y_hY - x_h)$$

and that this immediately gives

$$Y_1Y_2\cdots Y_{h+1} = (-1)^h(x_h - y_hY)^{-1}. \tag{35}$$

The quantity $-\deg(x_h - y_hY) = \deg y_{h+1}$ is a weighted sum giving a measure of the "distance" traversed by the continued fraction expansion to its $(h+1)$-st complete quotient. Taking norms yields

$$(x_h - y_hY)(x_h - y_h\overline{Y}) = (-1)^{h+1}Q_{h+1}\,. \tag{36}$$

6.3 Conjugation, symmetry, and periodicity

Each partial quotient a_h is the polynomial part of its corresponding complete quotient Y_h. Note, however, that the assertions above are independent of that conventional selection rule.

One readily shows that Y_0 being *reduced*, to wit $\deg Y_0 > 0$ and $\deg \overline{Y}_0 < 0$, implies that each complete quotient Y_h is reduced. Indeed, it also follows that $\deg B_h > 0$, while plainly $\deg \overline{B}_h < 0$ since $-\overline{B}_h$ is a remainder; so the B_h too are reduced. In particular a_h, the polynomial part of Y_h, is also the polynomial part of B_h.

Plainly, at least the first two leading terms of each polynomial P_h must coincide with the leading terms of $Y - T$. It also follows that the polynomials P_h and Q_h satisfy the bounds

$$\deg P_h = g+1 \quad \text{and} \quad \deg Q_h \le g\,. \tag{37}$$

Thus, if the base field $\mathbb{F}$ is finite the box principle entails the continued fraction expansion of Y_0 is periodic. If $\mathbb{F}$ is infinite, periodicity is just happenstance.

Suppose, however, that the function field $\mathbb{F}(X,Y)$ is exceptional in that Y_0, say, *does* have a periodic continued fraction expansion. If the continued fraction expansion of Y_0 is periodic then, by conjugation, also the expansion of B_0 is periodic. But conjugation reverses the order of the lines comprising a continued fraction tableau. Hence the conjugate of any preperiod is a "postperiod", an absurd notion. It follows that, if periodic, the two conjugate expansions are purely periodic.

Denote by A the polynomial part of Y, and recall that $Y+\overline{Y}=T$. It happens that line 0 of the continued fraction expansion of $Y+A-T$ is

$$Y+A-T=2A-T-(\overline{Y}+A-T) \tag{38}$$

and is symmetric. In general, if the expansion of Y_0 has a symmetry, and if the continued fraction expansion is periodic, its period must have a second symmetry[1]. So if Y is exceptional in having a periodic continued fraction expansion then its period is of length $2s$ and has an additional symmetry of the first kind $P_s=P_{s+1}$, or its period is of length $2s+1$ and also has a symmetry of the second kind, $Q_s=Q_{s+1}$. Conversely, this is the point, if the expansion of Y has a second symmetry then it must be periodic as just described.

6.4 Units

It is easy to apply the Dirichlet box principle to prove that an order $\mathbb{Q}[\omega]$ of a quadratic number field $\mathbb{Q}(\omega)$ contains nontrivial units. Indeed, by that principle there are infinitely many pairs of integers (p,q) so that $|q\omega-p|<1/q$, whence $|p^2-(\omega+\overline{\omega})pq+\omega\overline{\omega}q^2|<(\omega-\overline{\omega})+1$. It follows, again by the box principle, that there is an integer l with $0<|l|<(\omega-\overline{\omega})+1$ so that the equation $p^2-(\omega+\overline{\omega})pq+\omega\overline{\omega}q^2=l$ has infinitely many pairs (p,q) and (p',q') of solutions with $p\equiv p'$ and $q\equiv q' \pmod{l}$. For each such distinct pair, $xl=pp'-\omega\overline{\omega}qq'$, $yl=pq'-p'q+(\omega+\overline{\omega})qq'$, yields $(x-\omega y)(x-\overline{\omega}y)=1$.

In the function field case, we cannot apply the the box principle for a second time if the base field $\mathbb{F}$ is infinite. So the existence of a nontrivial unit $x(X)-y(y)Y(X)$ is exceptional. This should not be a surprise. By the definition of the notion "unit", such a unit $u(X)$ say, has a divisor supported only at infinity. Moreover, u is a function of the order $\mathbb{F}[X,Y]$, and is say of degree m, so the existence of u implies that the class containing the divisor at infinity is a torsion divisor on the Jacobian of the curve (30). The existence of such a torsion divisor is of course exceptional.

Suppose now that the function field $\mathbb{F}(X,Y)$ does contain a nontrivial unit u, say of norm $-\kappa$ and degree m. Then $\deg(yY-x)=-m<-\deg y$, so x/y is a convergent of Y and so some Q is $\pm\kappa$, say $Q_r=\kappa$ with r odd. That is, line r of the continued fraction expansion of $Y+A-T$ is

$$Y_r:=(Y+A-T)/\kappa=2A/\kappa-(\overline{Y}+A-T)/\kappa\,; \qquad \text{line r:}$$

here we have used the fact that $(Y+P_r)/\kappa$ is reduced to deduce that necessarily $P_r=P_{r+1}=A-T$.

By conjugation of the $(r+1)$-line tableau commencing with (38) we see that

$$Y_{2r}:=Y+A-T=2A-T-(\overline{Y}+A-T)\,, \qquad \text{line 2r:}$$

so that in any case if $Y+A-T$ has a quasi-periodic continued fraction expansion then it is periodic of period twice the quasi-period. This result of Tom Berry [2] applies to arbitrary

[1]The case of period length 1 is an exception unless we count its one line as having two symmetries; alternatively unless we deem it to have period $r=2$.

quadratic irrationals with polynomial trace. Other elements $(Y+P)/Q$ of $\mathbb{F}(X,Y)$, with Q dividing the norm $(Y+P)(\overline{Y}+P)$, may be honest-to-goodness quasi-periodic, that is, not also periodic.

Further, if $\kappa \neq -1$ then r *must* be odd. To see that, notice the identity

$$B[\,Ca_0\,,\,Ba_1\,,\,Ca_2\,,\,Ba_3\,,\,\ldots\,] = C[\,Ba_0\,,\,Ca_1\,,\,Ba_2\,,\,Ca_3\,,\,\ldots\,], \tag{39}$$

reminding one how to multiply a continued fraction expansion by some quantity; this cute formulation of the multiplication rule is due to Wolfgang Schmidt [11]. The "twisted symmetry" occasioned by division by κ, equivalent to the existence of a non-trivial quasi-period, is noted by Christian Friesen [4].

In summary, if the continued fraction expansion of Y is quasi-periodic it is periodic, and the expansion has the symmetries of the more familiar number field case, as well as twisted symmetries occasioned by a nontrivial κ.

One shows readily that if $x/y = [\,A\,,\,a_1\,,\,\ldots\,,\,a_{r-1}\,]$ and $x - Yy$ is a unit of the domain $\mathbb{F}[X,Y]$ then, with $a_{r-1} = \kappa a_1$, $a_{r-2} = a_2/\kappa$, $a_{r-3} = \kappa a_3$, $\ldots$,

$$[\,\overline{2A-T\,,\,a_1\,,\,\ldots\,,\,a_{r-1}\,,\,(2A-T)/\kappa\,,\,a_{r-1}\,,\,\ldots\,,\,a_1}\,]$$

is the quadratic irrational Laurent series $Y + A - T$. Alternatively, given the expansion of $Y + A - T$, and noting that therefore $\deg Q_r = 0$, the fact that the said expansion of x/y yields a unit follows directly from (36).

7 Comments

7.1

According to Gauss (*Disquisitiones Arithmeticæ*, Art. 76) ... *veritates ex notionibus potius quam ex hauriri debebant*[2]. Nonetheless, one should not underrate the importance of notation; good notation can decrease the viscosity of the flow to truth. From the foregoing it seems clear that, given $Y^2 = A^2 + 4v(X-w)$, one should study the continued fraction expansion of $Z = \frac{1}{2}(Y+A)$, as is done in [1]. Moreover, it is a mistake to be frustrated by minimal models $V^2 + UV - vV = U^3 - fU + vwU$.

Specifically, we understand that $V^2 - 8vV = U^3 - (4f-1)U^2 + 8v(2w-1)U$ yields $Y^2 = (X^2 + 4f - 1)^2 + 4\cdot 8v\big(X - (2w-1)\big)$ by way of $2U = X^2 + Y + (4f-1)$ and $(V - 8v) = XU$. Now $X \leftarrow 2X+1$, $Y \leftarrow 4Y$ means that, instead, we obtain $Y^2 = (X^2 + X + f)^2 + 4v\big(X - (w-1)\big)$. This derives from $V^2 + UV - vV = U^3 - fU + vwU$ by taking $2U = X^2 + X + Y + f$ and $V - v = XU$.

7.2

The discussion above may have some interest for its own sake, but my primary purpose is to test ideas for generalisation to higher genus g. An important difficulty when $g > 1$ is that

[2] [mathematical] truths flow from notions rather than from notations.

partial quotients may be of degree greater than one without that entailing periodicity, whence my eccentric aside on page 8. Happily, the generalisation to translating by a point $(w_0, e_0 - e_1)$ on the quartic model effected above also is a simplification in that one surely may always choose a translating divisor so as to avoid meeting singular steps in the continued fraction expansion. In that context one finds that the sequence $(\dots, 2, 1, 1, 1, 1, 1, 1, 2, 3, 4, 8, 17, 50, \dots)$ satisfying the recursion $T_{h-3}T_{h+3} = T_{h-2}T_{h+2} + T_h^2$ arises from adding multiples of the class of the divisor at infinity on the Jacobian of the curve $Y^2 = (X^3 - 4X + 1)^2 + 4(X - 2)$ of genus 2 to the class of the divisor defined by the pair of points $(\varphi, 0)$ and $(\overline{\varphi}, 0)$; here φ is the golden ratio.

References

[1] William W. Adams and Michael J. Razar, Multiples of points on elliptic curves and continued fractions, *Proc. London Math. Soc.* **41** (1980), 481–498.

[2] T. G. Berry, On periodicity of continued fractions in hyperelliptic function fields, *Arch. Math.* **55** (1990), 259–266.

[3] Graham Everest, Alf van der Poorten, Igor Shparlinski, and Thomas Ward, *Recurrence Sequences*, Mathematical Surveys and Monographs 104, American Mathematical Society 2003, 318pp.

[4] Christian Friesen, Continued fraction characterization and generic ideals, in [6], 465–474.

[5] David Gale, The strange and surprising saga of the Somos sequences, *Mathematical Intelligencer* **13**.1 (1991), 40–42; and 'Somos sequence update', *ibid.* **13**.4 (1991), 49–50. For more see Jim Propp, 'The Somos Sequence Site', http://www.math.wisc.edu/~propp/somos.html.

[6] David Goss, David R. Hayes and Michael I. Rosen eds., *The Arithmetic of Function Fields*, Proceedings of the workshop held at The Ohio State University, Columbus, Ohio, June 17–26, 1991. Ohio State University Mathematical Research Institute Publications, **2**. Walter de Gruyter & Co., Berlin, 1992. viii+482 pp.

[7] B. Mazur, Modular curves and the Eisenstein ideal, *Inst. Hautes Études Sci. Publ. Math.* **47** (1977), 33–186.

[8] Alfred J. van der Poorten, Periodic continued fractions and elliptic curves, in *High Primes and Misdemeanours*: lectures in honour of the 60th birthday of Hugh Cowie Williams, Alf van der Poorten and Andreas Stein eds., Fields Institute Communications **42**, American Mathematical Society, 2004, 353–365.

[9] ———, Elliptic curves and continued fractions, v0.9 of this manuscript: http://www.arxiv.org/math.NT/0403225.

[10] Alfred J. van der Poorten and Xuan Chuong Tran, Quasi-elliptic integrals and periodic continued fractions, *Monatshefte Math.*, 131 (2000), 155-169.

[11] Wolfgang M, Schmidt, On continued fractions and diophantine approximation in power series fields, *Acta Arith.* **95** (2000), 139–166.

[12] Rachel Shipsey, *Elliptic divisibility sequences*, Phd Thesis, Goldsmiths College, University of London, 2000 (see http://homepages.gold.ac.uk/rachel/).

[13] Rachel Shipsey, talk at "The Mathematics of Public Key Cryptography", Toronto 1999; see http://homepages.gold.ac.uk/rachel/toronto/sld006.htm.

[14] N. J. A. Sloane, The on-line encyclopedia of integer sequences, http://www.research.att.com/~njas/sequences/.

[15] David Speyer, E-mail concluding http://www.math.wisc.edu/~propp/somos/elliptic.

[16] Christine Swart, *Elliptic curves and related sequences*, PhD Thesis, Royal Holloway and Bedford New College, University of London, 2003; 226pp.

[17] Morgan Ward, Memoir on elliptic divisibility sequences, *Amer. J. Math.* **70** (1948), 31–74.

[18] Don Zagier, Problems posed at the St Andrews Colloquium, 1996, Solutions, 5th day; see http://www-groups.dcs.st-and.ac.uk/~john/Zagier/Problems.html.

2000 *Mathematics Subject Classification*: Primary 11A55, 11G05; Secondary 14H05, 14H52.
Keywords: continued fraction expansion, function field of characteristic zero, elliptic curve, Somos sequence.

(Concerned with sequences A006720 and A006721.)

Received September 20, 2004; revised version received November 18 2004. Published in *Journal of Integer Sequences*, May 16 2005.

Return to Journal of Integer Sequences home page.