AN INVESTIGATION OF LUCAS SEQUENCES

DUSTIN E. HINKEL

Contents

1. Introduction

In 1878 Édouard Lucas published a paper which summarized much of his research into the theory of what he called "simply periodic numerical functions." These were pairs of integer sequences which exhibit many properties characteristic of the trigonometric functions, and which he put to use in developing a new type of primality test.* One way to define these "Lucas sequences" is to pick some integers P and Q, and put: $(V_0, U_0) = (2, 0)$, $(V_1, U_1) = (P, 1)$, and (for $n > 1$)

$$(V_{n+1}, U_{n+1}) = P(V_n, U_n) - Q(V_{n-1}, U_{n-1}) \tag{1}$$

(here addition and scalar multiplication are performed component-wise). From these definitions one can show (and we will see) that

$$V_{n+m} = \frac{1}{2}(V_n V_m + \Delta U_n U_m),$$
$$U_{n+m} = \frac{1}{2}(V_n U_m + U_n V_m)$$

(where $\Delta = P^2 - 4Q$). If we put $V'_n = \frac{1}{2}V_n$ and $\Delta' = \frac{\Delta}{4}$, these become

$$V'_{n+m} = V'_n V'_m + \Delta' U_n U_m,$$
$$U_{n+m} = V'_n U_m + U_n V'_m.$$

Compare this last pair of identities with the trigonometric identities

$$\cos(n\theta + m\theta) = \cos n\theta \cos m\theta + (-1)\sin n\theta \sin m\theta,$$
$$\sin(n\theta + m\theta) = \cos n\theta \sin m\theta + \sin n\theta \cos m\theta.$$

That the sequence $\{(V_n, U_n)\}$ behaves like $\{(\cos n\theta, \sin n\theta)\}$ is very surprising. In fact, we will see that the sequence $\{(\cos n\theta, \sin n\theta)\}$ satisfies a recurrence similar to (1). Lucas sequences also have a curious connection to the solutions of the Pell equation $x^2 - Dy^2 = 1$ (actually this is a generalization of the connection to the trig functions): when $Q = 1$ (recall that $\{U_n\}$ and $\{V_n\}$ are defined in terms of P and Q), we have

$$V_n^2 - \Delta U_n^2 = 4$$

for all n; and in groups of the form $\{(x, y) : x^2 - Dy^2 \equiv_m 4\}$, every cyclic subgroup can be described in terms of $\{U_n\}$ and $\{V_n\}$. Another interesting fact regarding these sequences is that $\{U_n\}$ shares many of the same divisibility properties as the sequence of integers. For example $U_m | U_n$ iff $m | n$; moreover $\gcd(U_m, U_n) = U_{\gcd(m,n)}$. Combining this last fact with the fact that almost every prime q (that is, all odd primes not dividing $P^2 - 4Q$) divides either U_{q-1} or U_{q+1}, we have the following:

Theorem. Let $m > 1$ be odd and $\varepsilon = \pm 1$. If $m | U_{m-\varepsilon}$ but $m \not| U_{\frac{m-\varepsilon}{q}}$ for each prime q dividing $m - \varepsilon$, then m is prime.

In certain cases the converse also holds, as with the Lucas-Lehmer test for Mersenne primes.

The main goal of this paper is to demystify the unexpected connections among Lucas sequences, trigonometric functions, and solutions of the Pell equation. A secondary goal is to explore the number theoretic properties of Lucas sequences. In Section 2 we examine the solution set of the "Pellian equation" $x^2 - Dy^2 = r^2$ (where $r, D \in \mathbb{R}$). In particular we show that from a single solution (X, Y), we can define two linear recurrence sequences $\{x_n\}$ and $\{y_n\}$ such that $x_n^2 - Dy_n^2 = r^2$ for all n. In Section 3 we see that the solution set of $x^2 - Dy^2 = r^2$ (where x, y, D, r, r^{-1} are in an arbitrary commutative ring R) forms an abelian group, and we consider the group structure when $R = \mathbb{Z}$, $R = \mathbb{Z}_q$ (q an odd prime), and when $R = \mathbb{Z}_{pq}$ (p and q odd primes). In Section 4 we consider second-order linear recurrence sequences. We derive a closed form expression for the nth term of such a sequence, and see why Lucas' $\{U_n\}$ is a natural sequence to study. Section 5 is devoted to the Lucas sequences: their similarity to the trigonometric functions, and their number theoretic properties.

---

*We take the following from [Wil2] (page 57) to illustrate the significance of this new test:

> This is Lucas' great achievement: he discovered that primality testing for certain integers could be effected without having recourse to a table of primes, or what is more to the point, without having to perform a very large number of trial divisions. This was a completely new and original discovery, one which has exercised the most profound influence on the practice of primality proving ever since.

In Section 6 we explore how these properties can be used to discover information about a given integer's prime factors (in particular, whether that number is prime), and we examine a modification of the RSA cryptosystem. We also include an appendix on material tangential to our main discussion: in Appendix A we examine nonsimple continued fractions, and give a necessary and sufficient condition for their convergence.

The main sources for the theory of Lucas sequences are Lucas' 1878 paper[†] and Derrick Lehmer's *An Extended Theory of Lucas' Functions* (his 1930 Ph.D. thesis). These two papers develop enough of the theory to develop primality criteria and theorems regarding the prime factors of U_n. Another excellent source is Hugh Williams' book *Édouard Lucas and Primality Testing*.

We conclude this section with an example. In it we see a hint of the links among linear recurrence sequences, trigonometric functions, and solutions of the Pell equation.

Example. Recall that the matrix $\begin{pmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{pmatrix}$ represents a rotation in $\mathbb{R}^2$ about the origin by the angle θ. Applying this rotation to the point $\begin{pmatrix}\cos\theta' \\ \sin\theta'\end{pmatrix}$, we have the angle sum formulas for cosine and sine:

$$\begin{pmatrix}\cos(\theta'+\theta) \\ \sin(\theta'+\theta)\end{pmatrix} = \begin{pmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}\cos\theta' \\ \sin\theta'\end{pmatrix}. \tag{2}$$

The angle subtraction formulas can be expressed as

$$\begin{pmatrix}\cos(\theta'-\theta) \\ \sin(\theta'-\theta)\end{pmatrix} = \begin{pmatrix}\cos\theta & \sin\theta \\ -\sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}\cos\theta' \\ \sin\theta'\end{pmatrix}. \tag{3}$$

Now add equations (2) and (3):

$$\begin{aligned}\begin{pmatrix}\cos(\theta'+\theta) \\ \sin(\theta'+\theta)\end{pmatrix} + \begin{pmatrix}\cos(\theta'-\theta) \\ \sin(\theta'-\theta)\end{pmatrix} &= \left(\begin{pmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{pmatrix} + \begin{pmatrix}\cos\theta & \sin\theta \\ -\sin\theta & \cos\theta\end{pmatrix}\right)\begin{pmatrix}\cos\theta' \\ \sin\theta'\end{pmatrix} \\ &= \begin{pmatrix}2\cos\theta & 0 \\ 0 & 2\cos\theta\end{pmatrix}\begin{pmatrix}\cos\theta' \\ \sin\theta'\end{pmatrix} \\ &= 2\cos\theta\begin{pmatrix}\cos\theta' \\ \sin\theta'\end{pmatrix}.\end{aligned}$$

Fix some θ and (for all integers n) put $c_n = \cos(n\theta)$, $s_n = \sin(n\theta)$. Then we have

$$\begin{pmatrix}\cos(n\theta+\theta) \\ \sin(n\theta+\theta)\end{pmatrix} + \begin{pmatrix}\cos(n\theta-\theta) \\ \sin(n\theta-\theta)\end{pmatrix} = 2\cos\theta\begin{pmatrix}\cos(n\theta) \\ \sin(n\theta)\end{pmatrix},$$

or

$$\begin{pmatrix}c_{n+1} \\ s_{n+1}\end{pmatrix} = 2c_1\begin{pmatrix}c_n \\ s_n\end{pmatrix} - \begin{pmatrix}c_{n-1} \\ s_{n-1}\end{pmatrix}.$$

Now each pair $(\cos(n\theta), \sin(n\theta)) = (c_n, s_n)$ is a solution of $x^2 - Dy^2 = 1$ for $D = -1$. So from the solution (c_1, s_1), we can generate (possibly infinitely many) more solutions by means of this linear recurrence. Since each rational point on the unit circle corresponds to a pythagorean triple, we see that from one triple, this recurrence will generate more triples (infinitely many more, unless $c_1 = 0$ or $s_1 = 0$). To illustrate, we consider that $5^2 + 12^2 = 13^2$, and put $c_1 = \frac{5}{13}$ and $s_1 = \frac{12}{13}$ (so that $c_1^2 + s_1^2 = 1$). With $(c_0, s_0) = (1, 0)$, we get

$$\begin{aligned}(c_2, s_2) &= 2c_1(c_1, s_1) - (c_0, s_0) = \tfrac{10}{13}(\tfrac{5}{13}, \tfrac{12}{13}) - (1, 0) = (\tfrac{50}{169} - 1, \tfrac{120}{169}) = (-\tfrac{119}{169}, \tfrac{120}{169}), \\ (c_3, s_3) &= \tfrac{10}{13}(-\tfrac{119}{169}, \tfrac{120}{169}) - (\tfrac{5}{13}, \tfrac{12}{13}) = (-\tfrac{1190}{2197} - \tfrac{5}{13}, \tfrac{1200}{2197} - \tfrac{12}{13}) = (-\tfrac{2035}{2197}, -\tfrac{828}{2197}),\end{aligned}$$

[†]The title of this is *Théorie des Fonctions Numériques Simplement Périodiques*. 'Simply periodic' refers to the fact that these sequences admit expression in terms of sine and cosine; Lucas had planned on extending the theory to include sequences which behave like doubly periodic functions. Allegedly, Lucas thought that these latter sequences could be used to prove Fermat's Last Theorem. See [Wil2] (pages 71 and 72).

and likewise we find that

$$(c_4, s_4) = (-\tfrac{239}{28561}, -\tfrac{28560}{28561}),$$
$$(c_5, s_5) = (\tfrac{341525}{371293}, -\tfrac{145668}{371293}),$$
$$(c_6, s_6) = (\tfrac{3455641}{4826809}, \tfrac{3369960}{4826809}).$$

This is a rather unwieldy method of generating pythagorean triples, and much nicer methods exist (like projecting $\mathbb{Q}$ onto the unit circle by the map $t \mapsto (\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2})$). However, it is interesting that linear recurrences can be used for this purpose.

Notation

By $\mathbb{Z}$, $\mathbb{R}$, and $\mathbb{C}$ we mean (respectively) the integers, the reals, and the complex numbers. We take $\mathbb{N}$ to be the positive integers. The variables i, j, k, l, m, n, and N will denote integers, and q will denote a positive prime integer (we do not consider 1 to be prime). 'Commutative ring' will mean 'commutative ring with identity.' For a ring R, $R^\times$ is the set of invertible elements of R, considered as a group under multiplication. For an integer m, $\mathbb{Z}_m$ will mean $\mathbb{Z}/m\mathbb{Z}$. $\mathbb{F}_{q^k}$ is the field of q^k elements (so when q is a prime, we will occasionally use $\mathbb{F}_q$ and $\mathbb{Z}_q$ interchangeably). We will write $a \equiv_m b$ to mean $a \equiv b \pmod m$. When it is understood that we are working in the ring $\mathbb{Z}_m$, we will make no distinction between an integer and its congruence class modulo m. A quadratic residue (mod m) is a nonzero $a \in \mathbb{Z}_m$ such that $a \equiv_m g^2$ for some $g \in \mathbb{Z}_m$. A quadratic non-residue is a nonzero $a \in \mathbb{Z}_m$ that is not a quadratic residue. When q is an odd prime and $a \in \mathbb{Z}_q$, we define the Legendre symbol (a/q) as

$$(a/q) = \begin{cases} 1, & a \text{ is a quadratic residue,} \\ -1, & a \text{ is a quadratic non-residue,} \\ 0, & a \equiv_q 0. \end{cases}$$

For integers n and m, $n|m$ means that n divides m (that is, there is some integer k such that $m = kn$). In this paper we make use of several standard facts from algebra and number theory, such as Fermat's Little Theorem, Lagrange's Theorem, Euler's Criterion, quadratic reciprocity, the structure of $\mathbb{Z}_m$, the Chinese Remainder Theorem. The proofs of these (and any other facts not explicitly worked out) can be found among [Bur], [Ire-Ros], [Gau], [Hun].

2. The Equation $x^2 - Dy^2 = r^2$

Let $r, D \in \mathbb{R}$ be nonzero. Suppose $X, Y \in \mathbb{R}$ are such that $X^2 - DY^2 = r^2$. Equivalently,

$$\left(\frac{X}{r}\right)^2 - D\left(\frac{Y}{r}\right)^2 = 1.$$

Then for all n,

$$\left(\frac{X}{r} + \sqrt{D}\frac{Y}{r}\right)^n \left(\frac{X}{r} - \sqrt{D}\frac{Y}{r}\right)^n = \left(\left(\frac{X}{r}\right)^2 - D\left(\frac{Y}{r}\right)^2\right)^n = 1$$

(if $D < 0$, we take $\sqrt{D}$ to be $i\sqrt{|D|}$). For $n = 0, 1, 2, \ldots$, put

$$\begin{aligned} x_n &= \frac{r}{2}\left(\left(\frac{X}{r} + \sqrt{D}\frac{Y}{r}\right)^n + \left(\frac{X}{r} - \sqrt{D}\frac{Y}{r}\right)^n\right), \\ y_n &= \frac{r}{2\sqrt{D}}\left(\left(\frac{X}{r} + \sqrt{D}\frac{Y}{r}\right)^n - \left(\frac{X}{r} - \sqrt{D}\frac{Y}{r}\right)^n\right). \end{aligned} \tag{4}$$

Then

$$\begin{aligned} x_n + \sqrt{D}y_n &= r\left(\frac{X}{r} + \sqrt{D}\frac{Y}{r}\right)^n, \\ x_n - \sqrt{D}y_n &= r\left(\frac{X}{r} - \sqrt{D}\frac{Y}{r}\right)^n, \end{aligned} \tag{5}$$

and

$$x_n^2 - Dy_n^2 = r^2 \left(\frac{X}{r} + \sqrt{D}\frac{Y}{r}\right)^n \left(\frac{X}{r} - \sqrt{D}\frac{Y}{r}\right)^n = r^2.$$

By (5), we have (for $k, n \geq 0$)

$$\begin{aligned} x_{n+k} + \sqrt{D}y_{n+k} &= r\left(\frac{X}{r} + \sqrt{D}\frac{Y}{r}\right)^n \left(\frac{X}{r} + \sqrt{D}\frac{Y}{r}\right)^k \\ &= (x_n + \sqrt{D}y_n)\left(\frac{x_k + \sqrt{D}y_k}{r}\right), \end{aligned}$$

or

$$x_{n+k} + \sqrt{D}y_{n+k} = \frac{1}{r}\left((x_n x_k + Dy_n y_k) + \sqrt{D}(x_n y_k + y_n x_k)\right). \tag{6}$$

Both $(x_n x_k + Dy_n y_k)$ and $(x_n y_k + y_n x_k)$ are inner products:

$$x_n x_k + Dy_n y_k = \begin{pmatrix} x_k \\ Dy_k \end{pmatrix}^T \begin{pmatrix} x_n \\ y_n \end{pmatrix}, \quad \text{and}$$

$$x_n y_k + y_n x_k = \begin{pmatrix} y_k \\ x_k \end{pmatrix}^T \begin{pmatrix} x_n \\ y_n \end{pmatrix}.$$

So

$$x_{n+k} + \sqrt{D}y_{n+k} = \frac{1}{r}\left(\begin{pmatrix} x_k \\ Dy_k \end{pmatrix}^T \begin{pmatrix} x_n \\ y_n \end{pmatrix} + \sqrt{D}\begin{pmatrix} y_k \\ x_k \end{pmatrix}^T \begin{pmatrix} x_n \\ y_n \end{pmatrix}\right),$$

or

$$\begin{pmatrix} x_{n+k} \\ y_{n+k} \end{pmatrix} = \frac{1}{r}\begin{pmatrix} x_k & Dy_k \\ y_k & x_k \end{pmatrix}\begin{pmatrix} x_n \\ y_n \end{pmatrix}. \tag{7}$$

The determinant of $\frac{1}{r}\begin{pmatrix} x_k & Dy_k \\ y_k & x_k \end{pmatrix}$ is $\frac{1}{r^2}(x_k^2 - Dy_k^2) = 1$, so

$$\begin{pmatrix} x_n \\ y_n \end{pmatrix} = \frac{1}{r}\begin{pmatrix} x_k & -Dy_k \\ -y_k & x_k \end{pmatrix}\begin{pmatrix} x_{n+k} \\ y_{n+k} \end{pmatrix},$$

or (when $n \geq k$)

$$\begin{pmatrix} x_{n-k} \\ y_{n-k} \end{pmatrix} = \frac{1}{r}\begin{pmatrix} x_k & -Dy_k \\ -y_k & x_k \end{pmatrix}\begin{pmatrix} x_n \\ y_n \end{pmatrix}. \tag{8}$$

Adding (7) and (8), we have

$$\begin{pmatrix} x_{n+k} \\ y_{n+k} \end{pmatrix} + \begin{pmatrix} x_{n-k} \\ y_{n-k} \end{pmatrix} = \frac{1}{r}\left(\begin{pmatrix} x_k & Dy_k \\ y_k & x_k \end{pmatrix} + \begin{pmatrix} x_k & -Dy_k \\ -y_k & x_k \end{pmatrix}\right)\begin{pmatrix} x_n \\ y_n \end{pmatrix} = \frac{2x_k}{r}\begin{pmatrix} x_n \\ y_n \end{pmatrix}. \tag{9}$$

In particular,

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} + \begin{pmatrix} x_{n-1} \\ y_{n-1} \end{pmatrix} = \frac{2X}{r}\begin{pmatrix} x_n \\ y_n \end{pmatrix} \tag{10}$$

for $n > 0$.* Using (10), or by putting $n = 0$ in (9), we can define x_k and y_k for negative k (using that $(x_0, y_0) = (r, 0)$):

$$\begin{aligned} x_{-k} &= \frac{2x_k}{r} \cdot x_0 - x_k = \frac{2x_k}{r} \cdot r - x_k = x_k, \\ y_{-k} &= \frac{2x_k}{r} \cdot y_0 - y_k = \frac{2x_k}{r} \cdot 0 - y_k = -y_k. \end{aligned} \tag{11}$$

Even though the definitions of x_n and y_n sometimes require (as when $D < 0$) the use of complex numbers, the fact that $x_0, x_1, y_0, y_1 \in \mathbb{R}$ (in conjunction with the recurrence (10)) gives us that $x_n, y_n \in \mathbb{R}$ for all $n \in \mathbb{Z}$.

---

*In his "Disquisitiones Arithmeticae" (article 200), Gauss defines two sequences as in (4) and almost immediately thereafter declares that "it is easy to confirm that [equation (10) holds]." Let it be noted that figuring out how to "easily" confirm this was, in fact, *not* easy.

2.1. **Example.** Fix some $\theta \in \mathbb{R}$. In the example from the introduction we put $c_n = \cos(n\theta)$, $s_n = \sin(n\theta)$ and saw that for all n,

$$\begin{pmatrix} c_{n+1} \\ s_{n+1} \end{pmatrix} = 2c_1 \begin{pmatrix} c_n \\ s_n \end{pmatrix} - \begin{pmatrix} c_{n-1} \\ s_{n-1} \end{pmatrix}.$$

Put $D = -1$, $r = 1$, $X = \cos\theta$, and $Y = \sin\theta$. Then

$$X^2 - DY^2 = \cos^2\theta + \sin^2\theta = 1 = r^2.$$

So with x_n and y_n as in (4), we have

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = 2X \begin{pmatrix} x_n \\ y_n \end{pmatrix} - \begin{pmatrix} x_{n-1} \\ y_{n-1} \end{pmatrix} = 2\cos\theta \begin{pmatrix} x_n \\ y_n \end{pmatrix} - \begin{pmatrix} x_{n-1} \\ y_{n-1} \end{pmatrix}.$$

Now $x_0 = 1 = \cos 0 = c_0$, $y_0 = 0 = \sin 0 = s_0$, $x_1 = \cos\theta = c_1$, and $y_1 = \sin\theta = s_1$. Since the sequences $\{x_n\}$ and $\{c_n\}$ agree for $n = 0$ and $n = 1$ and since they satisfy the same recurrence, $x_n = c_n$ for all n. Likewise, $y_n = s_n$ for all n. Then in (5) we see De Moivre's formula:

$$\cos(n\theta) + \sqrt{-1}\sin(n\theta) = c_n + \sqrt{D}s_n = x_n + \sqrt{D}y_n = (X + \sqrt{D}Y)^n = (\cos\theta + \sqrt{-1}\sin\theta)^n.$$

2.2. **Remark.** Recall the derivation of equations (4) through (11): we let $r, D \in \mathbb{R}^\times$; then for $X, Y \in \mathbb{R}$ such that $X^2 - DY^2 = r^2$, we defined $x_n, y_n \in \mathbb{R}$ (or in $\mathbb{R}(i\sqrt{-D})$, if $\sqrt{D} \notin \mathbb{R}$) by equation (4). Then we have (5), from which (10) follows. This recurrence gives us that all x_n and y_n are real (despite possibly having been defined as linear combinations of nonreals). Notice that we never used $\mathbb{R}$'s order structure, nor that $\mathbb{R}$ is a metric space, nor that $\mathbb{R}$ is uncountable (or even infinite), and we never used that $\mathbb{R}$ is a field. We *did* use that $\mathbb{R}$ is a commutative ring, and (in (4)) that $2 \in \mathbb{R}^\times$. So we can generalize the results from this section by considering the solutions of $x^2 - Dy^2 = r^2$ in an arbitrary commutative ring R, where $r, D, 2 \in R^\times$. Now if $\sqrt{D} \notin R$, then many of the steps to derive an analog of (10) in R require considering $R(\sqrt{D})$. However, we show that (10) follows from an alternate definition of x_n and y_n (which does not involve $\sqrt{D}$), in which we can also get rid of the requirements that $2, D \in R^\times$. Note that (for $X, Y \in R$, and still assuming $2, D \in R^\times$)

$$\begin{aligned} x_n = \frac{r^{1-n}}{2}((X + \sqrt{D}Y)^n + (X - \sqrt{D}Y)^n) &= \frac{r^{1-n}}{2}\sum_{i=0}^{n} \binom{n}{i} X^{n-i}(\sqrt{D}Y)^i(1 + (-1)^n) \\ &= r^{1-n}\sum_{\substack{i=0 \\ i \text{ even}}}^{n} \binom{n}{i} X^{n-i}D^{\frac{i}{2}}Y^i, \end{aligned} \tag{12}$$

and that

$$\begin{aligned} y_n = \frac{r^{1-n}}{2\sqrt{D}}((X + \sqrt{D}Y)^n - (X - \sqrt{D}Y)^n) &= \frac{r^{1-n}}{2\sqrt{D}}\sum_{i=0}^{n} \binom{n}{i} X^{n-i}(\sqrt{D}Y)^i(1 - (-1)^n) \\ &= r^{1-n}\sum_{\substack{i=0 \\ i \text{ odd}}}^{n} \binom{n}{i} X^{n-i}D^{\frac{i-1}{2}}Y^i \end{aligned} \tag{13}$$

(if necessary, we are working in $R(\sqrt{D})$). If we define x_n and y_n by (12) and (13), we get that

$$\begin{aligned} r\left(\frac{X}{r} + \sqrt{D}\frac{Y}{r}\right)^n &= r^{1-n}(X + \sqrt{D}Y)^n \\ &= r^{1-n}\sum_{i=0}^{n} \binom{n}{i} X^{n-i}(\sqrt{D}Y)^i \\ &= r^{1-n}\sum_{\substack{i=0 \\ i \text{ even}}}^{n} \binom{n}{i} X^{n-i}D^{\frac{i}{2}}Y^i + \sqrt{D}r^{1-n}\sum_{\substack{i=0 \\ i \text{ odd}}}^{n} \binom{n}{i} X^{n-i}D^{\frac{i-1}{2}}Y^i \\ &= x_n + \sqrt{D}y_n, \end{aligned}$$

and (similarly) $x_n - \sqrt{D}y_n = r(\frac{X}{r} - \sqrt{D}\frac{Y}{r})^n$. It follows that

$$x_{n+k} + \sqrt{D}y_{n+k} = \frac{1}{r}\left((x_nx_k + Dy_ny_k) + \sqrt{D}(x_ny_k + y_nx_k)\right), \tag{14}$$

from which we can derive (10). Note that defining x_n and y_n by (12) and (13) gives us that $x_n, y_n \in R$ regardless of whether $\sqrt{D} \in R$.

3. Groups of the Form $\{(x, y) : x^2 - Dy^2 = r^2\}$

The Group $\{(x_n, y_n) : n \in \mathbb{Z}\}$. Let $r, D \in R$ (with $r \in R^\times$) and let $X, Y \in R$ be such that $X^2 - DY^2 = r^2$. Put $G = \{(x_n, y_n) : n \in \mathbb{Z}\}$ (where x_n and y_n are as in (12) and (13)) and define an operation $*$ on G by

$$(x_k, y_k) * (x_l, y_l) = \left(\frac{1}{r}(x_k x_l + D y_k y_l), \frac{1}{r}(x_k y_l + y_k x_l)\right). \tag{15}$$

By (14) we see that this is (x_{k+l}, y_{k+l}).*

3.1. **Remark.** Notice that $(x_n, y_n) = (X, Y)^n$ for all $n \geq 0$.

3.2. **Claim.** $(G, *)$ is an abelian group, with identity (x_0, y_0).

Proof. Indeed, for $j, k, l \in \mathbb{Z}$,

$$\begin{aligned}
(x_k, y_k) * (x_l, y_l) = (x_{k+l}, y_{k+l}) = (x_{l+k}, y_{l+k}) = (x_l, y_l) * (x_k, y_k),\\
(x_0, y_0) * (x_k, y_k) = (x_k, y_k) * (x_0, y_0) = (x_{k+0}, y_{k+0}) = (x_k, y_k),\\
(x_k, y_k) * (x_{-k}, y_{-k}) = (x_{k-k}, y_{k-k}) = (x_0, y_0),
\end{aligned}$$

$$\begin{aligned}
((x_j, y_j) * (x_k, y_k)) * (x_l, y_l)) &= (x_{j+k}, y_{j+k}) * (x_l, y_l)\\
&= (x_{j+k+l}, y_{j+k+l})\\
&= (x_j, y_j) * (x_{k+l}, y_{k+l})\\
&= (x_j, y_j) * ((x_k, y_k) * (x_l, y_l)).
\end{aligned}$$

□

This operation $*$ allows us to combine two solutions of the equation $x^2 - Dy^2 = r^2$ to obtain a third solution (as long as the two solutions can each be generated by the same solution). But it might be the case that two arbitrarily chosen solutions are not generated by the same solution. In that case it is worthwhile to extend $*$ to be defined on larger sets.

3.3. **Definition.** Let R be a commutative ring with identity and let $r, D \in R$. Let

$$\mathcal{L}_R = \{(x, y) \in R \times R : x^2 - Dy^2 = r^2\}.$$

When R is understood from the context, $\mathcal{L}$ will denote $\mathcal{L}_R$.

3.4. **Definition.** If r is invertible, define an operation $*$ on $\mathcal{L}$ by

$$(x, y) * (x', y') = (r^{-1}(xx' + Dyy'), r^{-1}(xy' + x'y)). \tag{16}$$

3.5. **Claim.** $(\mathcal{L}_R, *)$ is an abelian group, with identity $(r, 0)$.

Proof. $\mathcal{L}_R$ is nonempty—it contains both $(r, 0)$ and $(-r, 0)$. Since R is commutative,

$$xx' + Dyy' = x'x + Dy'y$$

(and since R is a ring, $xy' + x'y = x'y + xy'$). Then $(r, 0) * (x, y) = (x, y) * (r, 0)$, which is

$$(r^{-1}(xr), r^{-1}(ry)) = (x, y).$$

If $(x, y) \in \mathcal{L}$, then also $(x, -y) \in \mathcal{L}$ (since $x^2 - D(-y)^2 = x^2 - Dy^2$), and

$$(x, y) * (x, -y) = (r^{-1}(x^2 - Dy^2), r^{-1} \cdot 0) = (r, 0). \tag{17}$$

To see that $*$ is an associative binary operation on $\mathcal{L}$, first notice that each $(x, y) \in \mathcal{L}$ corresponds to a unique matrix in $R^{2\times 2}$ whose determinant is r^2 (namely, $\begin{pmatrix} x & Dy \\ y & x \end{pmatrix}$). Also notice that left $*$-multiplication in $\mathcal{L}$ corresponds to left multiplication in $R^{2\times 2}$ by the matrix

$$\begin{pmatrix} r^{-1}x & r^{-1}Dy \\ r^{-1}y & r^{-1}x \end{pmatrix} = r^{-1} \begin{pmatrix} x & Dy \\ y & x \end{pmatrix}.$$

---

*Actually some care needs to be taken here. When $\sqrt{D} \notin R$, $\{1, \sqrt{D}\}$ is linearly independent over R, so we can equate the like terms in (14). However, when $\sqrt{D} \in R$, we use the recurrence to show that $(x_{n+1}, y_{n+1}) = (x_n, y_n) * (x_1, y_1)$. Then by induction we have $(x_{k+l}, y_{k+l}) = (x_k, y_k) * (x_l, y_l)$.

Suppose $(x,y),(x',y'),(x'',y'') \in \mathcal{L}$. Then

$$r^{-1}\begin{pmatrix} x & Dy \\ y & x \end{pmatrix}\begin{pmatrix} x' & Dy' \\ y' & x' \end{pmatrix}$$

has determinant r^2 (so $(x,y)*(x',y') \in \mathcal{L}$), and

$$r^{-1}\left(r^{-1}\begin{pmatrix} x & Dy \\ y & x \end{pmatrix}\begin{pmatrix} x' & Dy' \\ y' & x' \end{pmatrix}\right)\begin{pmatrix} x'' & Dy'' \\ y'' & x'' \end{pmatrix} = r^{-1}\begin{pmatrix} x & Dy \\ y & x \end{pmatrix}\left(r^{-1}\begin{pmatrix} x' & Dy' \\ y' & x' \end{pmatrix}\begin{pmatrix} x'' & Dy'' \\ y'' & x'' \end{pmatrix}\right),$$

so

$$((x,y)*(x',y'))*(x'',y'') = (x,y)*((x',y')*(x'',y'')). \qquad \square$$

The Group $\mathcal{L}_{\mathbb{Z}}$. By the previous discussion, $\mathcal{L}_{\mathbb{Z}}$ is an abelian group. When D is a negative integer, $\mathcal{L}$ is finite. Although this case is interesting,[†] we examine $\mathcal{L}_{\mathbb{Z}}$ only for D a positive integer.

3.6. **Theorem.** If $D \in \mathbb{N}$ is not a square and $r = 1$, then $\mathcal{L}_{\mathbb{Z}} \cong \mathbb{Z}_2 \times \mathbb{Z}$.

Before proving this, we make a reduction. For any $(a,b) \in \mathcal{L}$, we have $(a,b) = (-1,0)*(-a,b)$. Then $(1,0) = (-1,0)*(-1,0)$. So any $(a,b) \in \mathcal{L}$ can be expressed as $(-1,0)^i * (|a|,b)$ (where i is even if $a > 0$ and i is odd if $a < 0$). Put $\mathcal{L}' = \{(x,y) \in \mathcal{L} : x > 0\}$. Now $\{(1,0)^n\} \cong \mathbb{Z}_2$, so to show $\mathcal{L}_{\mathbb{Z}} \cong \mathbb{Z}_2 \times \mathbb{Z}$ it suffices to show $\mathcal{L}' \cong \mathbb{Z}$. We prove this by showing that (a) if $\mathcal{L}'\backslash\{(1,0)\}$ is nonempty, then $\mathcal{L}' \cong \mathbb{Z}$; and (b) $\mathcal{L}'\backslash\{(1,0)\}$ is nonempty. To prove the theorem, we use the following:

3.7. **Lemma.** Suppose $a,b,x,y,r \in \mathbb{N}$, $a^2 - Db^2 = x^2 - Dy^2 = r^2$, $r < x < a$, and $0 < y < b$. Then

$$0 < ax - Dby,$$

and

$$0 < bx - ay < br.$$

Proof of Lemma. Now

$$\frac{a}{\sqrt{D}b} = \sqrt{\frac{a^2}{Db^2}} = \sqrt{1 + \frac{r^2}{Db^2}} > 1,$$
$$\frac{x}{\sqrt{D}y} = \sqrt{\frac{x^2}{Dy^2}} = \sqrt{1 + \frac{r^2}{Dy^2}} > 1,$$

so $\frac{ax}{Dby} > 1$ (or $ax - Dby > 0$). Since $b > y$,

$$\frac{a^2}{b^2} = D + \frac{r^2}{b^2} < D + \frac{r^2}{y^2} = \frac{x^2}{y^2}.$$

Then $\frac{a}{b} < \frac{x}{y}$, and $bx - ay > 0$. If we suppose (for the sake of contradiction) that $br \le bx - ay$, then

$$b(x-r) \ge ay > 0,$$
$$b^2(x^2 - 2xr + r^2) \ge a^2y^2,$$
$$b^2(Dy^2 + r^2 - 2xr + r^2) \ge (Db^2 + r^2)y^2,$$
$$b^2(2r^2 - 2xr) \ge r^2y^2,$$
$$2rb^2(r-x) \ge r^2y^2 > 0.$$

But $r - x < 0$, while $2rb^2 > 0$. So $bx - ay < br$. $\square$

Proof of (a). Choose $(X,Y) \in \mathcal{L}'$ which is different from $(1,0)$. Assume that (X,Y) is chosen so that $|Y|$ is minimal. Without loss of generality, we assume $Y > 0$ (if $Y < 0$, then $(X,-Y) \in \mathcal{L}'$ and $-Y > 0$). Define the sequences $\{x_n\}$ and $\{y_n\}$ as in (4) (and recall that the sequences can be extended to negative indices by (10) or (11)). We claim that $\mathcal{L}' = \{(x_n,y_n) : n \in \mathbb{Z}\}$. Suppose $(a,b) \in \mathcal{L}'$ but $(a,b) \notin \{(x_n,y_n)\}$. Assume that (a,b) is chosen so that $|b|$ is minimal. Now $(a,-b) \in \mathcal{L}'$ but $(a,-b) \notin \{(x_n,y_n)\}$ (if $(a,-b) = (x_k,y_k)$,

[†]It is related to the (nontrivial) problem of determining which numbers are of the form $x^2 + y^2$ (or of the form $x^2 + ny^2$ for any natural number n).

then $(x_{-k}, y_{-k}) = (a, b)$), so we can assume without loss of generality that $b \geq 0$. Also b is nonzero, since $(x_0, y_0) = (1, 0)$. Since $\{y_n\}$ is increasing and since $b > y_0 = 0$, there is some k such that

$$y_k < b < y_{k+1}.$$

Consider

$$(x', y') = (a, b) * (x_k, -y_k) = (ax_k - Dby_k, bx_k - ay_k).$$

By the lemma, x' and y' are positive, so $(x', y') \notin \{(x_n, y_n)\}$ (otherwise $(a, b) = (x', y') * (x_k, y_k) \in \{(x_n, y_n)\}$). But (with $r = 1$ in the lemma)

$$y' = bx_k - ay_k < b,$$

contradicting the minimality of $|b|$. □

Proof of (b). By Corollary (A.8) of Appendix A, we have that there are infinitely many coprime pairs of integers (x, y) such that

$$\left| \frac{x}{y} - \sqrt{D} \right| < \frac{1}{y^2}$$

(or $|x - \sqrt{D}y| < \frac{1}{y}$). Either $\frac{x}{y} < \sqrt{D}$ or $\frac{x}{y} > \sqrt{D}$. If $\frac{x}{y} < \sqrt{D}$, then $|x + \sqrt{D}y| < 2\sqrt{D}y$, and

$$|x^2 - Dy^2| = |x - \sqrt{D}y||x + \sqrt{D}y| < \frac{1}{y} \cdot 2\sqrt{D}y = 2\sqrt{D}.$$

If $\sqrt{D} < \frac{x}{y}$, then $|x + \sqrt{D}y| < 2x$, so

$$|x^2 - Dy^2| = |x - \sqrt{D}y||x + \sqrt{D}y| < \frac{1}{y} \cdot 2x = 2\left(\frac{x}{y}\right) < 2\left(\sqrt{D} + \frac{1}{y^2}\right) \leq 2(\sqrt{D} + 1)$$

(we have $\frac{x}{y} < \sqrt{D} + \frac{1}{y^2}$ since $|\frac{x}{y} - \sqrt{D}| < \frac{1}{y^2}$). In either case, $|x^2 - Dy^2| < 2\sqrt{D} + 2$. The set

$$\{(x, y) \in \mathbb{N} \times \mathbb{N} : |x^2 - Dy^2| < 2\sqrt{D} + 2\}$$

is infinite, but there are finitely many integers between $-2\sqrt{D} - 2$ and $2\sqrt{D} + 2$. So for some m (with $|m| < 2\sqrt{D} + 2$), the set

$$\{(x, y) \in \mathbb{N} \times \mathbb{N} : x^2 - Dy^2 = m\}$$

is infinite (note that $m \neq 0$). Partition this set into equivalence classes: say $(x, y) \sim (x', y')$ iff $x \equiv_m x'$ and $y \equiv_m y'$. Since there are finitely many equivalence classes, there are distinct pairs of positive integers (a_1, b_1) and (a_2, b_2) ($a_1 \leq a_2$ without loss of generality; note that this is actually a strict inequality, since $a_1 \equiv a_2$ implies $b_1 \equiv b_2$) such that

(a) $a_1^2 - Db_1^2 = a_2^2 - Db_2^2 = m$,
(b) $a_1 - a_2 \equiv_m b_1 - b_2 \equiv_m 0$.

Then

$$a_1a_2 - Db_1b_2 \equiv_m a_1^2 - Db_1^2 \equiv_m 0$$

and

$$a_1b_2 - b_1a_2 \equiv_m a_1b_1 - b_1a_1 = 0.$$

Put $x' = \frac{1}{m}(a_1a_2 - Db_1b_2)$ and $y' = \frac{1}{m}(a_1b_2 - b_1a_2)$ (and note that these are integers). Since $a_1 < a_2$, we have

$$b_1^2 = \frac{a_1^2 - m}{D} < \frac{a_2^2 - m}{D} = b_2^2,$$

and so $b_1 < b_2$. By the lemma, x' and y' are positive. And

$$(x')^2 - D(y')^2 = \frac{1}{r^2}((a_1^2a_2^2 - 2Da_1a_2b_1b_2 + D^2b_1^2b_2^2) - D(a_1^2b_2^2 - 2a_1a_2b_1b_2 + b_1^2a_2^2)),$$

which equals

$$\frac{1}{r^2}(a_2^2(a_1^2 - Db_1^2) - Db_2^2(a_1^2 - Db_1^2)) = \frac{1}{r^2}(m(a_2^2 - Db_2^2)) = 1.$$

□

The Group $\mathcal{L}_{\mathbb{F}_q}$.

3.8. **Notation.** Let q be an odd prime. Recall that $\mathbb{F}_q$ is the field of q elements. We take $\mathbb{F}_q$ to be the set

$$\{0, 1, \ldots, q-1\}.$$

We assume that $D \in \mathbb{F}_q$ is nonzero. For a given $r \neq 0$, each solution to $x^2 - Dy^2 = 1$ corresponds uniquely to a solution of $(x')^2 - D(y')^2 = r^2$ (by the map $(x, y) \mapsto (xr, yr)$). For the remainder of this section we assume $r = 1$, but we note (in anticipation of the later case wherein $r = 2$) that the structure of $\mathcal{L}$ is independent of the choice of r (as long as $r \neq 0$). Recall that (D/q) is the Legendre symbol.

3.9. **Theorem.** The group $\mathcal{L}_{\mathbb{F}_q}$ has order $q - (D/q)$.

Proof. Case 1: $(D/q) = 1$. Let $d \in \mathbb{F}_q$ be such that $d^2 = D$. We claim that there is a one-to-one correspondence between $\mathcal{L}$ and

$$\mathcal{H} = \{(a, b) \in \mathbb{F}_q^\times \times \mathbb{F}_q^\times : ab = 1\}.$$

Consider the map $(x, y) \mapsto (x + dy, x - dy)$. If $(x, y) \in \mathcal{L}$, then

$$(x + dy)(x - dy) = x^2 - Dy^2 = 1,$$

so we see that the map takes $\mathcal{L}$ to $\mathcal{H}$. If $x + dy = x' + dy'$ and $x - dy = x' - dy'$, then we get (by adding the equalities $2x = 2x'$ and (by subtracting) $2dy = 2dy'$. Then $(x, y) = (x', y')$. To see that this map is surjective, put (for $(a, b) \in \mathcal{H}$) $x = \frac{a+b}{2}$ and $y = \frac{a-b}{2d}$ (since $(D/q) = 1$, $d \neq 0$). Then

$$x^2 - Dy^2 = \frac{(a+b)^2}{4} - D\frac{(a-b)^2}{4D} = \frac{4ab}{4} = 1.$$

This shows that

$$|\mathcal{L}| = |\{(a, b) \in \mathbb{F}_q^\times \times \mathbb{F}_q^\times : ab = 1\}| = |\mathbb{F}_q^\times| = q - 1 = q - (D/q).$$

Case 2: $(D/q) = -1$. In this case we consider the set $\mathcal{H}' = \{(a, b) \in \mathbb{F}_q^\times \times \mathbb{F}_q^\times : ab = D\}$ and show that

$$|\mathcal{L}| = |\mathcal{H}'| + 2 = |\mathbb{F}_q^\times| + 2 = (q - 1) + 2 = q + 1 = q - (D/q)$$

(noting that since $(a, b) \in \mathcal{H}'$ is uniquely determined by $a \in \mathbb{F}_q^\times$, $|\mathcal{H}'| = |\mathbb{F}_q^\times| = q - 1$). If $(x, y) \in \mathcal{L}$ and $y \neq 0$, then

$$\frac{x+1}{y} \cdot \frac{x-1}{y} = \frac{x^2 - 1}{y^2} = D,$$

so consider the map

$$(x, y) \mapsto \left(\frac{x+1}{y}, \frac{x-1}{y}\right) : \mathcal{L} \backslash \{(\pm 1, 0)\} \to \mathcal{H}'.$$

Now if $\frac{x+1}{y} = \frac{x'+1}{y'}$ and $\frac{x-1}{y} = \frac{x'-1}{y'}$, then (by subtracting the equalities) $\frac{2}{y} = \frac{2}{y'}$ (or $y = y'$). Adding the equalities, we would have $\frac{2x}{y} = \frac{2x'}{y'}$, or $x = x'$. This map is also onto: if $(a, b) \in \mathcal{H}'$, then $a \neq b$ (otherwise we would have $a^2 = D$; but $(D/q) = -1$), so we can put $y = \frac{2}{a-b}$ and $x = \frac{a+b}{a-b}$, and then

$$x^2 - Dy^2 = \frac{(a+b)^2 - 4D}{(a-b)^2} = \frac{a^2 + 2ab + b^2 - 4ab}{(a-b)^2} = \frac{(a-b)^2}{(a-b)^2} = 1.$$

So $\mathcal{L}$ is in one-to-one correspondence with $\mathcal{H}' \cup \{(1, 0), (-1, 0)\}$, and

$$|\mathcal{L}| = |\mathcal{H}'| + 2 = q - (D/q).$$

□

3.10. **Theorem.** The group $\mathcal{L}_{\mathbb{F}_q}$ is cyclic.

In this proof, we use a fact about finite fields (namely, that if $\mathbb{F}$ is a finite field, $\mathbb{F}^\times$ is cyclic). See Appendix B for a discussion of finite fields and a proof that they are cyclic.

Proof. Case 1: $(D/q) = 1$. Suppose $d^2 = D$. Let a be a generator of $\mathbb{F}_q^\times$ (in number theoretic terms, a is a primitive root mod q), and put $b = a^{-1}$. Then with $X = \frac{a+b}{2}$ and $Y = \frac{a-b}{2d}$, $X^2 - DY^2 = 1$. For $n = 1, 2, \ldots, q-1$, put (noting that $X + dY = a$ and $X - dY = b$)

$$x_n = \frac{1}{2}((X + dY)^n + (X - dY)^n) = \frac{a^n + b^n}{2},$$
$$y_n = \frac{1}{2d}((X + dY)^n - (X - dY)^n) = \frac{a^n - b^n}{2d}.$$

As in Remark 3.1,

$$(X,Y)^n = (x_n, y_n) = \left(\frac{a^n+b^n}{2}, \frac{a^n-b^n}{2d}\right)$$

for all n. If

$$\left(\frac{a^n+b^n}{2}, \frac{a^n-b^n}{2d}\right) = \left(\frac{a^m+b^m}{2}, \frac{a^m-b^m}{2d}\right)$$

(where $1 \le n, m \le q-1$), then $a^n + b^n = a^m + b^m$ and $a^n - b^n = a^m - b^m$. Adding the two equalities, we have $2a^n = 2a^m$, or $a^n = a^m$. Since a generates $\mathbb{F}_q^\times$ and $1 \le n, m \le q-1$, $n = m$. So the pairs (x_n, y_n) are distinct for $n = 1, 2, \ldots, q-1$. That is, $\left(\frac{a+b}{2}, \frac{a-b}{2d}\right)$ generates $\mathcal{L}$.

Case 2: $(D/q) = -1$. Put $\mathbb{F} = \mathbb{F}_q[x]/(x^2 - D)$ and let $d \in \mathbb{F}$ be such that $d^2 = D$. The polynomial $x^2 - D$ is irreducible in $\mathbb{F}_q[x]$, so $\mathbb{F} \cong \mathbb{F}_{q^2}$ (and so $\mathbb{F}^\times$ is cyclic). Define $N : \mathbb{F} \to \mathbb{F}_q$ by

$$N(a + db) = a^2 - Db^2.$$

Now N is multiplicative (that is, for any $g_1, g_2 \in \mathbb{F}$, $N(g_1)N(g_2) = N(g_1g_2)$). To see this, consider the map (for $a, b \in \mathbb{F}_q$) $a + db \mapsto \begin{pmatrix} a & Db \\ b & a \end{pmatrix}$ (call this map ϕ). Then, with $g_1 = a_1 + db_1$ and $g_2 = a_2 + db_2$,

$$\begin{aligned}
\phi(g_1)\phi(g_2) &= \begin{pmatrix} a_1 & Db_1 \\ b_1 & a_1 \end{pmatrix} \begin{pmatrix} a_2 & Db_2 \\ b_2 & a_2 \end{pmatrix} \\
&= \begin{pmatrix} a_1a_2 + Db_1b_2 & D(a_1b_2 + b_1a_2) \\ a_1b_2 + b_1a_2 & a_1a_2 + Db_1b_2 \end{pmatrix} \\
&= \phi(g_1g_2),
\end{aligned}$$

so

$$N(g_1)N(g_2) = \det(\phi(g_1))\det(\phi(g_2)) = \det(\phi(g_1)\phi(g_2)) = \det(\phi(g_1g_2)) = N(g_1g_2).$$

Since N is multiplicative, the set $G = \{g \in \mathbb{F} : N(g) = 1\}$ is a subgroup of $\mathbb{F}^\times$. Then because $\mathbb{F}^\times$ is cyclic, G is also cyclic. So if we show that $\mathcal{L}$ is isomorphic to G, we have shown that $\mathcal{L}$ is cyclic. Consider the map $(a, b) \mapsto (a + db) : \mathcal{L} \to G$. This is surjective (if $N(a + db) = 1$, then $(a, b) \in \mathcal{L}$), and has trivial kernel (if $(a, b) \mapsto 1 = id_G$, then $(a, b) = (1, 0) = id_{\mathcal{L}}$), so it is an isomorphism. □

The Group $\mathcal{L}_{\mathbb{Z}_{pq}}$. In Section 6 we consider a cryptosystem based on the group $\mathcal{L}_{\mathbb{Z}_{pq}}$, where p and q are odd primes. The purpose of this section is to obtain this group's structure.

3.11. **Theorem.** For p and q odd primes, $\mathcal{L}_{\mathbb{Z}_{pq}} \cong \mathcal{L}_{\mathbb{Z}_p} \times \mathcal{L}_{\mathbb{Z}_q}$.

Proof. Recall that $\mathbb{Z}_{pq} \cong \mathbb{Z}_p \times \mathbb{Z}_q$. Let x_p (respectively x_q) denote the projection of $x \in \mathbb{Z}_{pq}$ onto $\mathbb{Z}_p$ (respectively $\mathbb{Z}_q$). Consider the projection $\rho : \mathbb{Z}_{pq}^2 \to \mathbb{Z}_p^2 \times \mathbb{Z}_q^2$, defined by

$$\rho(a, b) = ((a_p, b_p), (a_q, b_q)).$$

We claim that ρ is an isomorphism between $\mathcal{L}_{\mathbb{Z}_{pq}}$ and $\mathcal{L}_{\mathbb{Z}_p} \times \mathcal{L}_{\mathbb{Z}_q}$. Let $((a', b'), (a'', b'')) \in \mathcal{L}_{\mathbb{Z}_p} \times \mathcal{L}_{\mathbb{Z}_q}$. By the Chinese Remainder Theorem,[‡] there is a unique $(a, b) \in \mathbb{Z}_{pq}^2$ such that $a \equiv_p a'$, $a \equiv_q a''$, $b \equiv_p b'$, and $b \equiv_q b''$. If $a^2 - Db^2 \not\equiv_{pq} 1$, then either $a^2 - Db^2 \not\equiv_p 1$ or $a^2 - Db^2 \not\equiv_q 1$, which would contradict the choice of (a', b') and (a'', b''). So $(a, b) \in \mathcal{L}_{\mathbb{Z}_{pq}}$, and thus ρ is surjective. It is also injective, because the Chinese Remainder Theorem gives us that (a, b) is unique. □

3.12. **Corollary.** $|\mathcal{L}_{\mathbb{Z}_{pq}}| = (p - (D/p))(q - (D/q))$.

[‡]The Chinese Remainder Theorem states that if $n_1, n_2, \ldots, n_k$ are pairwise coprime, and if $a_1, a_2, \ldots, a_k$ are all integers, then there is a unique positive $x \le n_1n_2 \cdots n_k$ such that $x \equiv_{n_i} a_i$ for each i. See [Bur] (Theorem 4.8) for a proof and some examples.

4. Second-Order Linear Recurrences

Recall how we defined the Lucas sequences $\{U_n\}$ and $\{V_n\}$. We let $P, Q \in \mathbb{Z}$, put $(V_0, U_0) = (2, 0)$, $(V_1, U_1) = (P, 1)$, and

$$(V_{n+1}, U_{n+1}) = P(V_n, U_n) - Q(V_{n-1}, U_{n-1})$$

for $n > 1$. In this section we examine linear recurrence sequences and derive a formula for the general term of such a sequence.

4.1. **Example.** Consider the sequence (call it $\{t_n\}$)

$$1,\ 2,\ 5,\ 12,\ 29,\ 70,\ 169,\ 408,\ 985,\ 2378,\ 5741,\ 13860,\ \ldots,$$

where $t_1 = 1$, $t_2 = 2$, and $t_{n+2} = 2t_{n+1} + t_n$ for $n \geq 1$. One thing to notice is that $\frac{t_{n+1}}{t_n}$ (call this ratio r_n) is at least 2, so (for $n > 1$)

$$2 \leq r_n = 2 + \frac{1}{r_{n-1}} \leq 2 + \frac{1}{2} = \frac{5}{2}$$

(and these are strict inequalities for $n > 2$). Hence $t_n = \frac{t_n}{t_{n-1}} \frac{t_{n-1}}{t_{n-2}} \cdots \frac{t_2}{t_1}$ is bounded below by 2^{n-1} and bounded above by $\left(\frac{5}{2}\right)^{n-1}$. The first few terms of $\{r_n\}$ are

$$2,\ 2.5,\ 2.4,\ 2.416\ldots,\ 2.4137\ldots,\ 2.41428\ldots,\ 2.414201\ldots,\ 2.414215\ldots;$$

further into the sequence, we have

$$\begin{aligned}
r_{12} &= 2.414213564\ldots,\\
r_{18} &= 2.4142135623731\ldots,\\
r_{24} &= 2.414213562373095049\ldots,\\
r_{30} &= 2.41421356237309504880017\ldots,\\
r_{36} &= 2.41421356237309504880168872 49\ldots,\\
r_{42} &= 2.41421356237309504880168872420 97\ldots,\\
r_{48} &= 2.41421356237309504880168872420969 8079\ldots,
\end{aligned}$$

which strongly suggests that these r_n's are converging. The nature of this (probable) convergence can be seen from the first several values of $r_{n+1} - r_n$:

$$\frac{1}{2}, -\frac{1}{10}, \frac{1}{60}, -\frac{1}{348}, \frac{1}{2030}, -\frac{1}{11830}, \frac{1}{68952}, -\frac{1}{401880}, \frac{1}{2342330}, -\frac{1}{13652098}, \frac{1}{79570260}.$$

It appears that, if r is the limit of $\{r_n\}$, the sequence $\{r_n - r\}$ is alternating (and converges like a geometric sequence). If we suppose that $\{r_n\}$ *does* converge to r, then

$$\begin{aligned}
r &= \lim \frac{t_{n+1}}{t_n} = \lim \left(2 + \frac{t_n}{t_{n-1}}\right) = 2 + \frac{1}{r},\\
r^2 &= 2r + 1,\\
r &= 1 \pm \sqrt{2}.
\end{aligned}$$

Now $1 - \sqrt{2}$ is negative, but each r_n is positive (since each t_n is positive). So if $\{r_n\}$ converges, its limit is $1 + \sqrt{2}$. We will show that $\{r_n\}$ does converge, and since t_n is of the order* of $(1 + \sqrt{2})^n$, the sequence $\{(1 + \sqrt{2})^n\}$ should be of interest in any search for a closed form for t_n. Because $1 + \sqrt{2}$ is a solution of $x^2 = 2x + 1$, we have

$$\begin{aligned}
(1 + \sqrt{2})^{n+2} = (1 + \sqrt{2})^n (1 + \sqrt{2})^2 &= (1 + \sqrt{2})^n (2(1 + \sqrt{2}) + 1)\\
&= 2(1 + \sqrt{2})^{n+1} + (1 + \sqrt{2})^n.
\end{aligned}$$

That is, the sequence $\{(1+\sqrt{2})^n\}$ satisfies the same recurrence as $\{t_n\}$ (the same is also true of $\{(1-\sqrt{2})^n\}$). Later we will see that we can express t_n as a function of both $(1 + \sqrt{2})^n$ and $(1 - \sqrt{2})^n$.

---

*There are constants c and C and some integer N such that for all $n > N$, $c(1 + \sqrt{2})^n < t_n < C(1 + \sqrt{2})^n$. In asymptotic notation, $t_n = \Theta((1 + \sqrt{2})^n)$.

Proof that $\{r_n\}$ converges. We use the fact that

$$t_{n+2}t_n - t_{n+1}^2 = (-1)^{n+1} \tag{18}$$

for all n. To see this, first note that

$$\begin{pmatrix} t_{n+2} & t_{n+1} \\ t_{n+1} & t_n \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} t_{n+3} & t_{n+2} \\ t_{n+2} & t_{n+1} \end{pmatrix}.$$

Put $T = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}$, and (for $n \geq 1$) put $M_n = \begin{pmatrix} t_{n+2} & t_{n+1} \\ t_{n+1} & t_n \end{pmatrix}$. Then

$$M_n = M_{n-1}T = M_{n-2}T^2 = \cdots = M_1 T^{n-1},$$

so (since $\det M_1 = 1$ and $\det T = -1$)

$$t_{n+2}t_n - t_{n+1}^2 = \det M_n = (\det M_1)(\det T^{n-1}) = (-1)^{n-1} = (-1)^{n+1}.$$

Let $n > 2$. Then $t_n > 2^{n-1}$, so

$$|r_{n+1} - r_n| = \left|\frac{t_{n+2}}{t_{n+1}} - \frac{t_{n+1}}{t_n}\right| = \left|\frac{t_{n+2}t_n - t_{n+1}^2}{t_n t_{n+2}}\right| = \frac{1}{t_n t_{n+2}} < (1/2)^{2n}.$$

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

$$\begin{aligned} |r_{n+k+1} - r_n| &\leq |r_{n+k+1} - r_{n+k}| + |r_{n+k} - r_{n+k-1}| + \cdots + |r_{n+1} - r_n| \\ &< (1/2)^{2n+2k} + (1/2)^{2n+2k-1} + \cdots + (1/2)^{2n+1} + (1/2)^{2n} \\ &= (1/2)^{2n}((1/2)^{2k} + (1/2)^{2k-1} + \cdots + (1/2) + 1) \\ &< (1/2)^{2n}(1 + (1/2) + (1/2)^2 + \cdots) \\ &= (1/2)^{2n-1}. \end{aligned}$$

Now let $n \to \infty$. □

Notice the similarity between the derivation of (18) in this proof and the derivation of equation (40) from Appendix A. Using continued fractions (and Theorem A.5) we can show that $\{\frac{w_{n+1}}{w_n}\}$ converges when $\{w_n\}$ satisfies (19) and $(w_1, w_2) = (1, A)$. To simplify our exploring this more general case, we introduce some notation.

4.2. **Notation.** For nonzero A, $B \in \mathbb{R}$, let $\mathcal{R}_{A,B}$ denote the set of sequences $\{w_n\}$ in $\mathbb{R}$ satisfying the recurrence

$$w_{n+2} = Aw_{n+1} + Bw_n \tag{19}$$

for all integers n.[†]

4.3. **Definition.** Let $\{W_n(A,B)\}$ be the sequence in $\mathcal{R}_{A,B}$ with $W_0(A,B) = 0$ and $W_1(A,B) = 1$. When A and B are unambiguous from the context, $\{W_n\}$ will denote $\{W_n(A,B)\}$.

4.4. **Claim.** For all nonzero $A, B \in \mathbb{R}$, $\frac{W_{n+1}(A,B)}{W_n(A,B)}$ converges.

4.5. **Remark.** As in the example $\{t_n\} = \{W_n(2,1)\}$, where the limit of $\frac{t_{n+1}}{t_n}$ was one of the solutions of $x^2 = 2x + 1$, the limit of $\frac{W_{n+1}(A,B)}{W_n(A,B)}$ is one of the solutions of $x^2 = Ax + B$.

Proof of Claim 4.4. Case 1: $A, B > 0$. Consider the continued fraction

$$A + \frac{B}{A+}\frac{B}{A+}\cdots$$

(see Appendix A for an explanation of this notation). Let the sequences $\{p_n\}$ and $\{q_n\}$ be defined as in Appendix A. Now $\{p_n\}$ and $\{q_n\}$ both satisfy the same recurrence as $\{W_n\}$. And $q_0 = 1 = W_0$, $q_1 = p_0 = A = W_1$,

[†]The sequence $\{t_n\}$ was only defined for $n \geq 1$. But we can extend $\{t_n\}$ so that it is defined for $n < 1$ by the rule

$$t_{n-1} = t_{n+1} - 2t_n.$$

Similarly, a sequence $\{w_n\}$ satisfying (19) can be extended by the rule

$$w_{n-1} = \frac{w_{n+1} - Aw_n}{B}.$$

So we can assume that w_n is defined for all integers n whenever $\{w_n\} \in \mathcal{R}_{A,B}$.

and $p_1 = A^2 + B = W_2$; so $q_n = W_n$ and $p_n = W_{n+1}$ for $n > 0$. By Theorem A.5, $\{\frac{p_n}{q_n}\} = \{\frac{W_{n+1}}{W_n}\}$ converges.

Case 2: $A > 0$, $B < 0$. We show that $W_n(A,B) = \frac{1}{C}W_{2n}(C,D)$ for some $C, D > 0$. By Case 1, $\frac{W_{n+1}(C,D)}{W_n(C,D)}$ would converge, and

$$\begin{aligned}\lim_{n\to\infty}\frac{W_{n+1}(A,B)}{W_n(A,B)} = \lim_{n\to\infty}\frac{\frac{1}{C}W_{2n+2}(C,D)}{\frac{1}{C}W_{2n}(C,D)} &= \lim_{n\to\infty}\left(\frac{W_{2n+2}(C,D)}{W_{2n+1}(C,D)}\frac{W_{2n+1}(C,D)}{W_{2n}(C,D)}\right)\\ &= \left(\lim_{n\to\infty}\frac{W_{n+1}(C,D)}{W_n(C,D)}\right)^2.\end{aligned}$$

Suppose $C, D > 0$ and for the moment let W_n denote $W_n(C,D)$. Now

$$\begin{aligned}W_{n+2} &= CW_{n+1} + DW_n\\ &= C(CW_n + DW_{n-1}) + DW_n\\ &= (C^2 + D)W_n + CDW_{n-1},\end{aligned}$$

and

$$CDW_{n-1} = D(CW_{n-1}) = D(W_n - DW_{n-2}),$$

so

$$\begin{aligned}W_{n+2} &= (C^2 + D)W_n + (DW_n - D^2W_{n-2})\\ &= (C^2 + 2D)W_n - D^2W_{n-2}.\end{aligned}$$

In particular, $\{W_{2n}(C,D)\} \in \mathcal{R}_{C',D'}$, where $C' = C^2 + 2D$ and $D' = -D^2$. Now $W_0(C,D) = 0$ and $W_2(C,D) = C$, so

$$\frac{1}{C}W_0(C,D) = 0 = W_0(C^2 + 2D, -D^2)$$

and

$$\frac{1}{C}W_2(C,D) = 1 = W_1(C^2 + 2D, -D^2).$$

Since the sequences $\{\frac{1}{C}W_{2n}(C,D)\}$ and $\{W_n(C^2 + 2D, -D^2)\}$ agree for $n = 0$ and $n = 1$, and since they satisfy the same recurrence, they agree for all n. If we put $D = \sqrt{-B}$ and $C = \sqrt{A - 2\sqrt{-B}}$, then

$$\begin{aligned}C^2 + 2D &= A - 2\sqrt{-B} + 2\sqrt{-B} = A,\\ -D^2 &= B,\end{aligned}$$

and so

$$W_n(A,B) = W_n(C^2 + 2D, -D^2) = \frac{1}{C}W_{2n}(C,D)$$

for all n.

Case 3: $A < 0$. By Cases 1 and 2, the sequence $\{\frac{W_{n+1}(-A,B)}{W_n(-A,B)}\}$ converges. We show by induction that

$$W_n(A,B) = (-1)^n W_n(-A,B)$$

for all n. Now

$$\begin{aligned}(-1)^0 W_0(-A,B) = 0 = W_0(A,B), \text{ and}\\ (-1)^1 W_1(-A,B) = A = W_1(A,B).\end{aligned}$$

If $k > 0$ and $W_j(A,B) = (-1)^j W_j(-A,B)$ for all $j < k$, then

$$\begin{aligned}(-1)^{k+1}W_{k+1}(-A,B) &= (-1)^{k+1}(-AW_k(-A,B) + BW_{k-1}(-A,B))\\ &= (-1)^2(A\cdot(-1)^k W_k(-A,B) + B\cdot(-1)^{k-1}W_{k-1}(-A,B))\\ &= AW_k(A,B) + BW_{k-1}(A,B)\\ &= W_{k+1}(A,B).\end{aligned}$$

So

$$\lim_{n\to\infty}\frac{W_{n+1}(A,B)}{W_n(A,B)} = \lim_{n\to\infty} -\frac{W_{n+1}(-A,B)}{W_n(-A,B)} \in \mathbb{R}. \qquad \square$$

4.6. **Remark.** Each sequence $\{w_n\}$ in $\mathcal{R}_{A,B}$ is completely determined by w_0 and w_1 (because we can find any other w_k by the recurrence relation). So we can identify $\mathcal{R}_{A,B}$ with the set of ordered pairs of reals. Moreover, we have the following for all λ, μ in $\mathbb{R}$ and $\{a_n\}$, $\{b_n\}$ in $\mathcal{R}_{A,B}$:

$$\begin{aligned}\lambda a_{n+2} - \mu b_{n+2} &= \lambda(Aa_{n+1} + Ba_n) - \mu(Ab_{n+1} + Bb_n)\\ &= A(\lambda a_{n+1} - \mu b_{n+1}) + B(\lambda a_n - \mu b_n).\end{aligned}$$

That is, $\mathcal{R}_{A,B}$ is closed under addition and closed under multiplication by scalars (and the zero sequence is certainly in $\mathcal{R}_{A,B}$). So $\mathcal{R}_{A,B}$ is a two-dimensional vector space over $\mathbb{R}$.

A Closed Form Expression for t_n. Previously we observed that $\{t_n\}$, $\{(1+\sqrt{2})^n\}$, and $\{(1-\sqrt{2})^n\}$ are in $\mathcal{R}_{2,1}$. We can identify these sequences with the vectors $\begin{pmatrix}0\\1\end{pmatrix}$, $\begin{pmatrix}1\\1+\sqrt{2}\end{pmatrix}$, and $\begin{pmatrix}1\\1-\sqrt{2}\end{pmatrix}$, respectively. Now $\begin{pmatrix}1\\1+\sqrt{2}\end{pmatrix}$ and $\begin{pmatrix}1\\1-\sqrt{2}\end{pmatrix}$ are linearly independent, so

$$\{t_n\} = \lambda\{(1+\sqrt{2})^n\} + \mu\{(1-\sqrt{2})^n\},$$

where λ and μ solve

$$\begin{pmatrix}0\\1\end{pmatrix} = \lambda\begin{pmatrix}1\\1+\sqrt{2}\end{pmatrix} + \mu\begin{pmatrix}1\\1-\sqrt{2}\end{pmatrix}.$$

Then $\lambda = \frac{1}{2\sqrt{2}}$, $\mu = \frac{-1}{2\sqrt{2}}$, and

$$t_n = \frac{(1+\sqrt{2})^n - (1-\sqrt{2})^n}{\sqrt{2}}.$$

The More General Case. Put $\alpha = \frac{A}{2} + \frac{\sqrt{A^2+4B}}{2}$ and $\beta = \frac{A}{2} - \frac{\sqrt{A^2+4B}}{2}$ (so α and β solve $x^2 = Ax + B$). Then the sequences $\{\alpha^n\}$ and $\{\beta^n\}$ are in $\mathcal{R}_{A,B}$. As long as $A^2 + 4B \neq 0$, these sequences are linearly independent. So if $\{w_n\} \in \mathcal{R}_{A,B}$, there are constants λ and μ such that

$$\{w_n\} = \lambda\{\alpha^n\} + \mu\{\beta^n\}.$$

We can find λ and μ by solving

$$\begin{pmatrix}w_0\\w_1\end{pmatrix} = \begin{pmatrix}\alpha^0 & \beta^0\\ \alpha^1 & \beta^1\end{pmatrix}\begin{pmatrix}\lambda\\\mu\end{pmatrix} = \begin{pmatrix}1 & 1\\ \alpha & \beta\end{pmatrix}\begin{pmatrix}\lambda\\\mu\end{pmatrix}.$$

So

$$\begin{pmatrix}\lambda\\\mu\end{pmatrix} = \frac{1}{\beta-\alpha}\begin{pmatrix}\beta & -1\\ -\alpha & 1\end{pmatrix}\begin{pmatrix}w_0\\w_1\end{pmatrix} = \frac{1}{\beta-\alpha}\begin{pmatrix}\beta w_0 - w_1\\ -\alpha w_0 + w_1\end{pmatrix},$$

and then

$$\begin{aligned}w_n &= \frac{1}{\beta-\alpha}[(\beta w_0 - w_1)\alpha^n + (-\alpha w_0 + w_1)\beta^n]\\ &= \frac{1}{\beta-\alpha}[(\alpha\beta)w_0(\alpha^{n-1} - \beta^{n-1}) - w_1(\alpha^n - \beta^n)].\end{aligned}$$

Now $\alpha\beta = -B$ (since α and β are roots of $x^2 - Ax - B$), so

$$w_n = \frac{1}{\beta-\alpha}[-Bw_0(\alpha^{n-1} - \beta^{n-1}) - w_1(\alpha^n - \beta^n)].$$

After some rearranging, we obtain a handy closed form expression for w_n:

$$w_n = w_1\frac{\alpha^n - \beta^n}{\alpha-\beta} + Bw_0\frac{\alpha^{n-1} - \beta^{n-1}}{\alpha-\beta}. \tag{20}$$

4.7. **Example.** Recall that $W_0 = 0$ and $W_1 = 1$. Then

$$W_n = \frac{\alpha^n - \beta^n}{\alpha-\beta} = \alpha^{n-1} + \alpha^{n-2}\beta + \alpha^{n-3}\beta^2 + \cdots + \alpha\beta^{n-2} + \beta^{n-1}, \tag{21}$$

and (20) becomes

$$t_n = t_1W_n + Bt_0W_{n-1}. \tag{22}$$

The Fibonacci sequence $\{F_n\}$ is $\{W_n(1,1)\}$. The roots of $x^2 - x - 1$ are $\phi = \frac{1+\sqrt{5}}{2}$ and $-\phi^{-1} = \frac{1-\sqrt{5}}{2}$, so (21) is Binet's formula:

$$F_n = \frac{\phi^n - (-1)^n\phi^{-n}}{\sqrt{5}} = \frac{(1+\sqrt{5})^n - (1-\sqrt{5})^n}{2^n\sqrt{5}}.$$

The Lucas sequences satisfy the recurrence (19), with $A = P$ and $B = -Q$ (actually, $U_n = W_n(P, -Q)$). Put $\delta = \sqrt{P^2 - 4Q}$. Then the roots of $x^2 - Px + Q$ are $\frac{P+\delta}{2}$ and $\frac{P-\delta}{2}$. Since $U_n = W_n(P, -Q)$, we get

$$U_n = \frac{(\frac{P+\delta}{2})^n - (\frac{P-\delta}{2})^n}{\frac{P+\delta}{2} - \frac{P-\delta}{2}} = \frac{1}{\delta}\left(\left(\frac{P+\delta}{2}\right)^n - \left(\frac{P-\delta}{2}\right)^n\right).$$

Using (22), we get that $V_n = PU_n - 2QU_{n-1}$. In the next section we will see that $V_n = (\frac{P+\delta}{2})^n + (\frac{P-\delta}{2})^n$ (more precisely, we will see that defining V_n by this equality agrees with the original definition).

4.8. **Remark.** Although we have considered only second-order recurrence sequences, the previous methods can be used to find a closed form expression for the general term in a sequence satisfying a recurrence of higher order. We consider this briefly. Suppose $\{w_n\}$ is a sequence in $\mathbb{C}$ such that for all n and for some constants $c_0, c_1 \ldots, c_{k-1} \in \mathbb{C}$,

$$w_{n+k} = c_{k-1}w_{n+k-1} + c_{k-2}w_{n+k-2} + \cdots + c_1w_{n+1} + c_0w_n.$$

The set of all sequences $\{w_n\}$ satisfying this recurrence (call it V) is a k-dimensional vector space, isomorphic to $\mathbb{C}^k$. If the polynomial

$$x^k - c_{k-1}x^{k-1} - c_{k-2}x^{k-2} - \cdots - c_1x - c_0$$

has distinct roots $\alpha_1, \alpha_2, \ldots, \alpha_k$, then the sequences $\{\alpha_1^n\}, \{\alpha_2^n\}, \ldots, \{\alpha_k^n\}$ are linearly independent over $\mathbb{C}$, and thus form a basis for V. There will then be $\lambda_1, \lambda_2 \ldots, \lambda_k \in \mathbb{C}$ such that

$$w_n = \lambda_1\alpha_1^n + \lambda_2\alpha_2^n + \cdots + \lambda_k\alpha_k^n.$$

5. The Lucas Sequences

In the introduction we defined $\{U_n\}$ as the sequence satisfying the recurrence

$$U_{n+1} = PU_n - QU_{n-1}, \tag{23}$$

such that $U_0 = 0$, $U_1 = 1$. From this definition, we saw in the previous section that

$$U_n = \frac{\alpha^n - \beta^n}{\alpha - \beta}, \tag{24}$$

where α and β are the solutions of $x^2 = Px - Q$. We could instead define U_n by (24). Then since

$$\alpha^{n+2} = \alpha^n(\alpha^2) = \alpha^n(P\alpha - Q) = P\alpha^{n+1} - Q\alpha^n$$

(and similarly, $\beta^{n+2} = P\beta^{n+1} - Q\beta^n$) and since U_n is a linear combination of α^n and β^n, we would get (23). When P and Q are integers, U_0 and U_1 are integers; so by the recurrence each U_n is an integer. This is the approach taken by Lucas (and after [Luc], it has become the conventional approach).

5.1. **Remark.** In the research for this paper, the investigation of Lucas sequences began as a study of linear recurrence sequences. From this direction it is curious that the recursive definition of $\{U_n\}$ and $\{V_n\}$ contains "$-Q$" instead of "$+Q$" (considering that the most commonly encountered linear recurrence sequences satisfy a recurrence of the form

$$w_{n+2} = Aw_{n+1} + Bw_n,$$

where A and B are positive integers).* However, the "$-Q$" in (23) comes from a (natural) sufficient condition that U_n is an integer for $n \geq 0$: for $\alpha, \beta \in \mathbb{C}$, the sequences $\{\alpha^n\}$ and $\{\beta^n\}$ (and hence $\{U_n\}$) satisfy

$$w_{n+2} = (\alpha + \beta)w_{n+1} - \alpha\beta w_n;$$

so since U_0 and U_1 are integers, each U_n is an integer as long as $\alpha + \beta$ and $\alpha\beta$ are integers.

*The Fibonacci sequence is the most famous. Other sequences which show up again and again include the numerators and denominators of the convergents of certain continued fractions.

5.2. **Remark.** The most interesting properties of Lucas sequences arise when we consider P and Q to be coprime integers, but for now we take P and Q to be nonzero reals such that $P^2 \neq 4Q$. Put $\delta = \sqrt{P^2 - 4Q}$, $\Delta = P^2 - 4Q$, $\alpha = \frac{P+\delta}{2}$, and $\beta = \frac{P-\delta}{2}$ (so that $\alpha + \beta = P$, $\alpha - \beta = \delta$, and $\alpha\beta = Q$; then also α, β solve $x^2 = Px - Q$).

5.3. **Definition.** Define the sequences $\{U_n(P,Q)\}$ and $\{V_n(P,Q)\}$ by

$$U_n(P,Q) = \frac{\alpha^n - \beta^n}{\alpha - \beta}, \tag{25}$$

$$V_n(P,Q) = \alpha^n + \beta^n = \frac{U_{2n}(P,Q)}{U_n(P,Q)}. \tag{26}$$

5.4. **Remark.** By previous discussion we see that $U_n(P,Q) = U_n$. Notice that $\{V_n(P,Q)\}$ satisfies the same recurrence as $\{\alpha^n\}$, $\{\beta^n\}$, and $\{U_n\}$. Also, $V_0(P,Q) = 2 = V_0$ and $V_1(P,Q) = P = V_1$. Since $\{V_n(P,Q)\}$ agrees with $\{V_n\}$ on $n = 0$ and $n = 1$, we have that they agree on all n. From now on we take take (25) and (26) to be our definitions of $\{V_n\}$ and $\{U_n\}$.

5.5. **Remark.** From (25) and (26), we see that we can extend $U_n(P,Q)$ and $V_n(P,Q)$ to the functions

$$U(P,Q,z) = \frac{e^{z\log\alpha} - e^{z\log\beta}}{\alpha - \beta}$$

and

$$V(P,Q,z) = e^{z\log\alpha} + e^{z\log\beta},$$

which are defined for all $z \in \mathbb{C}$ (note that $\alpha - \beta \neq 0$; and if $\alpha, \beta \notin \mathbb{R}$, choose the principal branch of the logarithm).

Bounds on U_n, V_n. We are assuming $\Delta \neq 0$. If $\Delta < 0$, then $\beta = \overline{\alpha}$, so $|\alpha| = |\beta|$. If $\Delta > 0$, then

$$\alpha = \frac{P+\delta}{2} > \frac{P-\delta}{2} = \beta.$$

In either case, $|\alpha| \geq |\beta|$. So

$$|V_n| = |\alpha^n + \beta^n| \leq |\alpha|^n + |\beta|^n \leq 2|\alpha|^n,$$

and

$$\begin{aligned} |U_n| = \left|\frac{\alpha^n - \beta^n}{\alpha - \beta}\right| = \left|\sum_{i+j=n-1} \alpha^i \beta^j\right| &\leq \sum_{i+j=n-1} |\alpha|^i |\beta|^j \\ &\leq \sum_{i+j=n-1} |\alpha|^{i+j} \\ &= n|\alpha|^{n-1}. \end{aligned}$$

Identities Involving U_n and V_n. The most fundamental identities are

$$\begin{aligned} V_{n+m} &= \frac{V_n V_m + \Delta U_n U_m}{2}, \\ U_{n+m} &= \frac{V_n U_m + U_n V_m}{2}. \end{aligned} \tag{27}$$

Recall the similarity between these last equalities and the trigonometric identities

$$\begin{aligned} \cos(n\theta + m\theta) &= \cos n\theta \cos m\theta + (-1)\sin n\theta \sin m\theta, \\ \sin(n\theta + m\theta) &= \cos n\theta \sin m\theta + \sin n\theta \cos m\theta. \end{aligned}$$

Actually, *every* identity involving $\{U_n\}$ and $\{V_n\}$ corresponds to a trigonometric identity, and vice versa (see [Bel], in which these correspondences are stately explicitly). Later we will briefly explore this connection. Of the many other known identities,[†] we list those which are most relevant to our study of the sequences' number theoretic properties:

$$V_{2n} = V_n^2 - 2Q^n, \tag{28}$$

[†]In [Wil2] (pages 74 through 83) are at least 70 identities (and in many cases, pairs of identities). These include several pages dealing with Chebyshev polynomials.

(29) $$U_{2n} = U_n V_n,$$

(30) $$V_{2n+1} = \frac{PV_n^2 + \Delta V_n U_n}{2} - PQ^n,$$

(31) $$U_{2n+1} = \frac{V_n^2 + PV_n U_n}{2} - Q^n,$$

(32) $$\begin{aligned} Q^n V_{-n} &= V_n, \\ Q^n U_{-n} &= -U_n, \end{aligned}$$

(33) $$\begin{aligned} 2Q^m V_{n-m} &= V_n V_m - \Delta U_n U_m, \\ 2Q^m U_{n-m} &= -V_n U_m + U_n V_m, \end{aligned}$$

(34) $$\begin{aligned} 2^{m-1} V_{mn} &= \sum_{\substack{i=0 \\ i \text{ even}}}^{m} \binom{m}{i} V_n^{m-i} \Delta^{\frac{i}{2}} U_n^i \quad (\text{for } m > 0 \text{ odd}), \\ 2^{m-1} U_{mn} &= \sum_{\substack{i=0 \\ i \text{ odd}}}^{m} \binom{m}{i} V_n^{m-i} \Delta^{\frac{i-1}{2}} U_n^i \quad (\text{for } m > 0 \text{ odd}), \end{aligned}$$

(35) $$V_n^2 - \Delta U_n^2 = 4Q^n,$$

(36) $$V_{mn}(P, Q) = V_m(V_n(P, Q), Q^n).$$

Deriving the Identities. Our main tools are

$$V_n + \delta U_n = (\alpha^n + \beta^n) + (\alpha^n - \beta^n) = 2\alpha^n$$

and

$$V_n - \delta U_n = (\alpha^n + \beta^n) - (\alpha^n - \beta^n) = 2\beta^n.$$

Multiplying these, we get (35). They also give us

$$\alpha^{n+m} = \frac{(V_n + \delta U_n)(V_m + \delta U_m)}{4} = \frac{V_n V_m + \Delta U_n U_m}{4} + \frac{\delta(V_n U_m + U_n V_m)}{4}$$

and (similarly)

$$\beta^{n+m} = \frac{V_n V_m + \Delta U_n U_m}{4} - \frac{\delta(V_n U_m + V_n U_m)}{4},$$

which, upon addition and subtraction, yield (27):

$$V_{n+m} = \alpha^{n+m} + \beta^{n+m} = \frac{V_n V_m + \Delta U_n U_m}{2}$$

and

$$\delta U_{n+m} = (\alpha^{n+m} - \beta^{n+m}) = \frac{\delta(V_n U_m + U_n V_m)}{2}$$

(or $U_{n+m} = \frac{V_n U_m + U_n V_m}{2}$). We have (29) by virtue of (26). From (27) and (35) we get

$$V_{2n} = \frac{V_n^2 + \Delta U_n^2}{2} = \frac{V_n^2 + (V_n^2 - 4Q^n)}{2} = V_n^2 - 2Q^n.$$

Now (27), (28), and (29) give us

$$\begin{aligned} V_{2n+1} = \frac{V_{2n}V_1 + \Delta U_{2n}U_1}{2} = \frac{PV_{2n} + \Delta U_{2n}}{2} &= \frac{P(V_n^2 - 2Q^n) + \Delta V_n U_n}{2} \\ &= \frac{PV_n^2 + \Delta V_n U_n}{2} - PQ^n, \end{aligned}$$

$$U_{2n+1} = \frac{V_{2n}U_1 + U_{2n}V_1}{2} = \frac{V_{2n} + PU_{2n}}{2} = \frac{V_n^2 - 2Q^n + PV_n U_n}{2} = \frac{V_n^2 + PV_n U_n}{2} - Q^n.$$

By (25) and (26), we have

$$Q^n V_{-n} = (\alpha\beta)^n(\alpha^{-n} + \beta^{-n}) = \beta^n + \alpha^n = V_n,$$
$$Q^n U_{-n} = (\alpha\beta)^n \frac{\alpha^{-n} - \beta^{-n}}{\alpha - \beta} = \frac{\beta^n - \alpha^n}{\alpha - \beta} = -U_n.$$

Combining (27) and (32) yields

$$\begin{aligned} 2Q^m V_{n-m} &= Q^m(V_n V_{-m} + \Delta U_n U_{-m}) \\ &= V_n(Q^m V_{-m}) + \Delta U_n(Q^m U_{-m}) \\ &= V_n V_m - \Delta U_n U_m, \end{aligned}$$

$$\begin{aligned} 2Q^m U_{n-m} &= Q^m(V_n U_{-m} + U_n V_{-m}) \\ &= V_n(Q^m U_{-m}) + U_n(Q^m V_{-m}) \\ &= -V_n U_m + U_n V_m. \end{aligned}$$

To show (34), suppose $m > 0$ is odd. We have $V_n + \delta U_n = 2\alpha^n$ and $V_n - \delta U_n = 2\beta^n$, so

$$2^m \alpha^{mn} = (V_n + \delta U_n)^m = \sum_{i=0}^{m} \binom{m}{i} V_n^{m-i} \delta^i U_n^i,$$
$$2^m \beta^{mn} = (V_n - \delta U_n)^m = \sum_{i=0}^{m} \binom{m}{i} (-1)^i V_n^{m-i} \delta^i U_n^i.$$

Then

$$2^m(\alpha^{mn} + \beta^{mn}) = \sum_{i=0}^{m} \binom{m}{i} V_n^{m-i} \delta^i U_n^i (1 + (-1)^i) = 2 \sum_{\substack{i=0 \\ i \text{ even}}}^{m} \binom{m}{i} V_n^{m-i} \delta^i U_n^i,$$

or

$$2^{m-1} V_{mn} = 2^{m-1}(\alpha^{mn} + \beta^{mn}) = \sum_{\substack{i=0 \\ i \text{ even}}}^{m} \binom{m}{i} V_n^{m-i} \Delta^{\frac{i}{2}} U_n^i.$$

Similarly,

$$2^{m-1} \delta U_{mn} = 2^{m-1}(\alpha^{mn} - \beta^{mn}) = \sum_{\substack{i=0 \\ i \text{ odd}}}^{m} \binom{m}{i} V_n^{m-i} \Delta^{\frac{i}{2}} U_n^i.$$

After cancelling a factor of $\delta = \Delta^{\frac{1}{2}}$, we get

$$2^{m-1} U_{mn} = \sum_{\substack{i=0 \\ i \text{ odd}}}^{m} \binom{m}{i} V_n^{m-i} \delta^{i-1} U_n^i = \sum_{\substack{i=0 \\ i \text{ odd}}}^{m} \binom{m}{i} V_n^{m-i} \Delta^{\frac{i-1}{2}} U_n^i.$$

For (36), put $P' = \alpha^n + \beta^n = V_n(P, Q)$ and $Q' = (\alpha\beta)^n = Q^n$ (so α^n and β^n are the roots of $x^2 = P'x - Q'$). Then

$$V_m(P', Q') = (\alpha^n)^m + (\beta^n)^m = \alpha^{mn} + \beta^{mn},$$

or

$$V_m(V_n(P, Q), Q^n) = V_{mn}(P, Q).$$

Computing U_n and V_n. Equations (28) through (31), known as the duplication formulas, allow us to compute a specific V_N or U_N fairly efficiently. The process is a modification of the repeated-squaring algorithm for exponentiation. For example, to find V_N and U_N for $N = 41$, we would take $V_1 = P$ and $U_1 = 1$, and use the duplication formulas to find V_n and U_n for $n = 2$, then $n = 5$, $n = 10$, $n = 20$, and

finally $n = 41$ (where either $n_{i+1} = 2n_i$ or $n_{i+1} = 2n_i + 1$, depending on the binary expansion of N). To formalize this, define functions $f_0, g_0, f_1, g_1 : \mathbb{R}^2 \times \mathbb{N} \to \mathbb{R}$ by:

$$\begin{aligned} f_0(u, v, n) &= uv, \\ g_0(u, v, n) &= v^2 - Q^n, \\ f_1(u, v, n) &= \frac{v^2 + Puv}{2} - Q^n, \\ g_1(u, v, n) &= \frac{Pv^2 + \Delta uv}{2} - PQ^n. \end{aligned}$$

Then for $n \in \mathbb{N}$ and for $\varepsilon \in \{0, 1\}$,

$$\begin{aligned} f_\varepsilon(U_n, V_n, n) &= U_{2n+\varepsilon}, \\ g_\varepsilon(U_n, V_n, n) &= V_{2n+\varepsilon}. \end{aligned}$$

Suppose $N, k \in \mathbb{N}$ and

$$N = \sum_0^k b_i 2^{k-i}$$

(so the binary expansion of N reads $b_0 b_1 \dots b_{k-1} b_k$). Put $N_0 = b_0$, $S_0 = U_1$, and $T_0 = V_1$. For $1 \le i \le k$ let

$$\begin{aligned} N_i &= 2N_{i-1} + b_i, \\ S_i &= f_{b_i}(S_{i-1}, T_{i-1}, N_i), \\ T_i &= g_{b_i}(S_{i-1}, T_{i-1}, N_i). \end{aligned}$$

Then $S_k = U_N$ and $T_k = V_N$. To illustrate, take $N = 41$ again. In binary, this is 101001. Then we have $N_0 = 1$, $S_0 = U_1$, $T_0 = V_1$, and

$$\begin{array}{lll} N_1 = 2N_0 + 0 = 2, & S_1 = f_0(S_0, T_0, N_1) = U_2, & T_1 = g_0(S_0, T_0, N_1) = V_2, \\ N_2 = 2N_1 + 1 = 5, & S_2 = f_1(S_1, T_1, N_2) = U_5, & T_2 = g_1(S_1, T_1, N_2) = V_5, \\ N_3 = 2N_2 + 0 = 10, & S_3 = f_0(S_2, T_2, N_3) = U_{10}, & T_3 = g_0(S_2, T_2, N_3) = V_{10}, \\ N_4 = 2N_3 + 0 = 20, & S_4 = f_1(S_3, T_3, N_4) = U_{20}, & T_4 = g_1(S_3, T_3, N_4) = V_{20}, \\ N_5 = 2N_4 + 1 = 41, & S_5 = f_1(S_4, T_4, N_5) = U_{41}, & T_5 = g_1(S_4, T_4, N_5) = V_{41}. \end{array}$$

Using the duplication formulas, we can compute U_{2n}, V_{2n}, U_{2n+1}, or V_{2n+1} in $O(\log V_n \log U_n) = O(n^2)$ time. In each step, $S_i = U_{2n+\varepsilon}$ and $T_i = V_{2n+\varepsilon}$ for $\varepsilon \in \{0, 1\}$ and where $2^{i-1} < n \le 2^i$. So S_i and T_i can each be computed in $O(2^i)$ time. We are finding S_i and T_i for i from 1 to $\lceil \log_2 N \rceil$, so the total cost of computing U_N or V_N is

$$O(\sum_{i=1}^{\lceil \log_2 N \rceil} 2^i) = O(2^{\lceil \log_2 N \rceil} - 1) = O(N).$$

We will often be considering the sequences $\{U_n\}$ and $\{V_n\}$, reduced modulo some m. If m is odd, then the duplication formulas hold in $\mathbb{Z}_m$. Multiplying any $a, b \in \mathbb{Z}_m$ is done in $O(1)$ time, so by the preceding "successive-duplication" process, U_N or V_N can be computed in $O(\log N)$ time.

U_n and V_n as Trigonometric Functions. Recall that the hyperbolic cosine and sine are defined (respectively) as

$$\cosh(z) = \frac{e^z + e^{-z}}{2} \qquad \text{and} \qquad \sinh(z) = \frac{e^z - e^{-z}}{2},$$

and so these can take complex-valued arguments. Notice that for $\alpha, \beta, x \in \mathbb{R}$ (and assuming $\alpha, \beta \neq 0$),

$$(\alpha/\beta)^x + (\beta/\alpha)^x = (\alpha/\beta)^x + (\alpha/\beta)^{-x} = e^{x \log(\alpha/\beta)} + e^{-x \log(\alpha/\beta)} = 2\cosh(x \log(\alpha/\beta)),$$

and

$$(\alpha/\beta)^x - (\beta/\alpha)^x = e^{x \log(\alpha/\beta)} - e^{-x \log(\alpha/\beta)} = 2\sinh(x \log(\alpha/\beta)).$$

If $\Delta < 0$ (so that $\frac{\alpha}{\beta} \notin \mathbb{R}$), take log to be the principal branch of the complex logarithm. Then for all n, we have

$$V_n = \alpha^n + \beta^n = (\alpha\beta)^n \left((\alpha/\beta)^{n/2} + (\beta/\alpha)^{n/2} \right) = 2Q^{n/2} \cosh\left(\frac{n}{2} \log(\alpha/\beta) \right),$$

and

$$U_n = \frac{\alpha^n - \beta^n}{\alpha - \beta} = \frac{(\alpha\beta)^n}{\delta}\left((\alpha/\beta)^{n/2} - (\beta/\alpha)^{n/2}\right) = \frac{2Q^{n/2}}{\delta}\sinh\left(\frac{n}{2}\log(\alpha/\beta)\right).$$

Since (for $z \in \mathbb{C}$)

$$\cos(z/i) = \frac{e^z + e^{-z}}{2} = \cosh(z)$$

and

$$i\sin(z/i) = \frac{e^z - e^{-z}}{2} = \sinh(z),$$

we also have

$$V_n = 2Q^{n/2}\cos\left(\frac{n}{2i}\log(\alpha/\beta)\right),$$
$$U_n = \frac{2iQ^{n/2}}{\delta}\sin\left(\frac{n}{2i}\log(\alpha/\beta)\right).$$

With these last equations in hand, all of our identities between U_n and V_n follow from familiar trigonometric identities. As an example, we show that $2U_{m+n} = V_mU_n + U_mV_n$. With $\theta = \frac{1}{2i}\log(\alpha/\beta)$,

$$\begin{aligned}
2U_{n+m} = \frac{4iQ^{\frac{n+m}{2}}}{\delta}\sin(n\theta + m\theta) &= 2Q^{n/2}\cdot 2Q^{m/2}\cdot\frac{i}{\delta}\left(\cos n\theta\sin m\theta + \sin n\theta\cos m\theta\right)\\
&= 2Q^{n/2}\left(\cos n\theta\left(\frac{2iQ^{m/2}}{\delta}\sin m\theta\right) + \frac{i}{\delta}\sin n\theta\left(2Q^{m/2}\cos m\theta\right)\right)\\
&= \left(2Q^{n/2}\cos n\theta\right)U_m + \left(\frac{2iQ^{n/2}}{\delta}\sin n\theta\right)V_m\\
&= V_nU_m + U_nV_m.
\end{aligned}$$

Number Theoretic Properties of U_n, V_n. In what follows, we assume P and Q to be coprime integers. Our goal in this section is Theorem 5.23, which is the basis for the primality testing in Section 6.

5.6. **Notation.** Recall that for integers n and m, $n|m$ means that n divides m.

5.7. **Example.** Consider again the sequence $\{t_n\} = \{U_n(2,-1)\}$ from Section 4. In Table 1 we list the first 100 terms by their prime factorization.[‡] There are several interesting patterns to notice: when $m|U_n$, it appears that $m|U_{kn}$ for $k = 1, 2, 3, \ldots$; for each odd prime q, $q|U_{q\pm1}$; if k is the highest power of q dividing U_n, then $k+1$ is the highest power of q dividing U_{qn}; $2^k|U_n$ iff $2^k|n$. Curiously, it also appears that $\gcd(U_m, U_n) = U_{\gcd(m,n)}$. We spend the remainder of this section proving several facts similar to these observations. In much of this section we consider the prime divisors of U_n, but not of V_n. We do this because $V_n = \frac{U_{2n}}{U_n}$, so any divisor of V_n is also a divisor of U_{2n} (in [Leh2], Lehmer examines the prime factors of $\frac{U_{kn}}{U_n}$ for a fixed k).

5.8. **Facts.** If $m|n$ then $U_m|U_n$. If $m|n$ and $\frac{n}{m}$ is odd, then $V_m|V_n$.

Proof. Say $n = km$. Then

$$U_n = \frac{\alpha^{km} - \beta^{km}}{\alpha - \beta} = \frac{(\alpha^m)^k - (\beta^m)^k}{\alpha^m - \beta^m}\frac{\alpha^m - \beta^m}{\alpha - \beta}.$$

Now $\frac{(\alpha^m)^k - (\beta^m)^k}{\alpha^m - \beta^m}$ is the kth term in the sequence $\{U_n(V_m, Q^m)\}$ (note that α^m and β^m are the roots of $x^2 - (\alpha^m + \beta^m)x + (\alpha\beta)^m$). We already know that $\frac{\alpha^m - \beta^m}{\alpha - \beta} = U_m(P,Q)$ is an integer; and since V_m and Q^m are integers, so is $U_k(V_m, Q^m) = \frac{(\alpha^m)^k - (\beta^m)^k}{\alpha^m - \beta^m}$.

[‡]See Table 3.1.1 in [Wil2] (page 55) for a similar table with the factorizations of Fibonacci numbers. Interestingly, it was because of studying the factors of Fibonacci numbers that Lucas was led to research his "simply periodic numerical functions."

Now suppose k is odd. Then

$$\begin{aligned} V_n = (\alpha^m)^k + (\beta^m)^k = (\alpha^m)^k - (-\beta^m)^k &= (\alpha^m + \beta^m)\frac{(\alpha^m)^k - (-\beta^m)^k}{\alpha^m - (-\beta^m)} \\ &= V_m(P,Q) \cdot U_k(V_m, -Q^m), \end{aligned}$$

and both $V_m(P,Q)$ and $U_k(V_m, -Q^m)$ are integers. □

5.9. **Definition.** For $m \in \mathbb{N}$, define $\omega(m)$ to be the least integer N (if it exists) such that $m|U_N$. If no such N exists, $\omega(m)$ is undefined. We call $\omega(m)$ (or simply ω) the *rank of apparition* of m.

5.10. **Remark.** For $m > 1$, consider the sequence $\{U_n\}$, reduced modulo m. If $m|Q$, then we have (for $n > 0$)

$$U_{n+1} \equiv_m PU_n \equiv_m P^2U_{n-1} \equiv_m \cdots \equiv_m P^nU_1 = P^n.$$

Now $\gcd(m, P) = 1$ (since $m|Q$ and we assume $\gcd(P, Q) = 1$), so $P^n \not\equiv_m 0$ for any n. So we have:

5.11. **Lemma.** If $m > 1$ and $m|Q$, then $\omega(m)$ is undefined.

We can strengthen this as:

5.12. **Theorem.** If $\gcd(m, Q) \neq 1$, then $\omega(m)$ is undefined.

Proof. If m and Q have a common divisor $d > 1$, then $\omega(d)$ is undefined by the previous lemma. Then m cannot divide any U_n, since otherwise there would be some n for which $d|U_n$ (and therefore $\omega(d) \leq n$). □

Even if $\gcd(m, Q) = 1$, it is not immediately clear whether $\omega(m)$ should be defined. We will see a sufficient condition for the existence of $\omega(m)$ in Theorem 5.17, the proof of which uses the next several lemmas.

5.13. **Lemma.** If Q is odd, then $2|U_2$ or $2|U_3$.

Proof. We consider the sequence $\{U_n\}$ in $\mathbb{Z}_2$. If P is even, then the recurrence is

$$U_{n+1} \equiv_2 U_{n-1},$$

and since U_0 is even, so is U_2. If P is odd, the recurrence is

$$U_{n+1} \equiv_2 U_n + U_{n-1};$$

so

$$U_{n+3} \equiv_2 U_{n+2} + U_{n+1} \equiv_2 (U_{n+1} + U_n) + U_{n+1} \equiv_2 U_n.$$

Now U_0 is even, so U_3 is even. □

Recall the Legendre symbol (a/q).

5.14. **Lemma.** Let q be an odd prime. For $n \in \mathbb{N}$, $U_{qn} \equiv_q (\Delta/q)U_n$ and $V_{qn} \equiv_q V_n$.

Proof. With $m = q$ in (34), and since $\binom{q}{i}$ is divisible by q except when $i = 0$ or $i = q$,

$$2^{q-1}U_{qn} = \sum_{\substack{i=0 \\ i \text{ odd}}}^{q} \binom{q}{i} V_n^{q-i}\Delta^{\frac{i-1}{2}}U_n^i \equiv_q \binom{q}{q} V_n^{q-q}\Delta^{\frac{q-1}{2}}U_n^q \equiv_q \Delta^{\frac{q-1}{2}}U_n^q,$$

$$2^{q-1}V_{qn} = \sum_{\substack{i=0 \\ i \text{ even}}}^{q} \binom{q}{i} V_n^{q-i}\Delta^{\frac{i}{2}}U_n^i \equiv_q \binom{q}{0} V_n^{q-0}\Delta^{\frac{0}{2}}U_n^0 \equiv_q V_n^q.$$

Using Fermat's Little Theorem and Euler's Criterion,[§] we have

$$\begin{aligned} U_{qn} &\equiv_q 2^{q-1}U_{qn} \equiv_q \Delta^{\frac{q-1}{2}}U_n^q \equiv_q (\Delta/q)U_n, \\ V_{qn} &\equiv_q 2^{q-1}V_{qn} \equiv_q V_n^q \equiv_q V_n. \end{aligned}$$

□

5.15. **Lemma.** Let q be an odd prime not dividing Q, and put $\varepsilon = (\Delta/q)$. Then $U_{q-\varepsilon} \equiv_q 0$.

[§]The former: if $a \in \mathbb{Z}_q^\times$, then $a^{q-1} \equiv_q 1$. The latter: if $a \in \mathbb{Z}_q$, then $a^{\frac{q-1}{2}} \equiv_q (a/q)$.

n	U_n	n	U_n
1	$-$	51	$5^1101^1137^18297^120468307053^1$
2	2^1	52	$2^23^179^1599^122307^133461^166923^1$
3	5^1	53	684804064621612874691^1
4	2^23^1	54	$2^15^17^153^1197^1199^1146449^17761799^1$
5	29^1	55	$29^1109^15741^1183041^1120159269^1$
6	$2^15^17^1$	56	$2^33^113^217^1113^1239^1337^11535466241^1$
7	13^2	57	$5^137^1179057^11311797^153535197^1$
8	$2^33^117^1$	58	$2^144560482149^163018038201^1$
9	5^1197^1	59	$13558774610046711780701^1$
10	$2^129^141^1$	60	$2^23^25^27^111^119^129^131^241^159^1269^1601^12281^1$
11	5741^1	61	$853^119062257^14860136428481^1$
12	$2^23^25^17^111^1$	62	$2^161^11301^1424577^1865087^13272609^1$
13	33461^1	63	$5^113^2197^145697^1198701^1304728101^1$
14	$2^113^2239^1$	64	$2^63^117^1257^1577^11409^1665857^12448769^1$
15	$5^229^1269^1$	65	$29^1521^11949^133461^1804829^13385201^1$
16	$2^43^117^1577^1$	66	$2^15^17^123^1353^15741^137667521^152734529^1$
17	137^18297^1	67	$4289^1364812640489943346522 1^1$
18	$2^15^17^1197^1199^1$	68	$2^23^167^1103^1137^18297^115607^133931^1757793^1$
19	$37^1179057^1$	69	$5^1229^1982789^181042255829535617^1$
20	$2^23^119^129^141^159^1$	70	$2^113^229^141^171^1239^117934071^11800193921^1$
21	$5^113^245697^1$	71	$569^12644939853^1353183656631413^1$
22	$2^123^1353^15741^1$	72	$2^33^35^17^111^117^173^1179^1197^1199^11009^11153^11523089^1$
23	$229^1982789^1$	73	$1873724873^1165338584129036705 7^1$
24	$2^33^25^17^111^117^11153^1$	74	$2^1223^1593^137223^178737^11013411^18761009^1$
25	$29^11549^129201^1$	75	$5^329^1149^1269^11549^129201^11371001^12004001^1$
26	$2^179^1599^133461^1$	76	$2^23^137^1227^1379^119457^134961^1179057^19369319^1$
27	$5^153^1197^1146449^1$	77	$13^25741^123869^141117^1110523055197017^1$
28	$2^23^113^2113^1239^1337^1$	78	$2^15^17^179^1313^1389^1599^133461^14088137^14605197^1$
29	44560482149^1	79	$157^17741^11735683721^1290781543899617^1$
30	$2^15^27^129^131^241^1269^1$	80	$2^43^117^119^129^141^159^1241^1577^15521^1888 01^19393281^1$
31	$61^11301^13272609^1$	81	$5^153^1197^110853^1146449^1244781^1176018980229^1$
32	$2^53^117^1577^1665857^1$	82	$2^12297^1302663^13553471^11746860020068409^1$
33	$5^15741^152734529^1$	83	$1993^1104551336978216677001376175 17^1$
34	$2^1103^1137^18297^115607^1$	84	$2^23^25^17^211^113^283^1113^1239^1251^1337^14663^145697^175937^1$
35	$13^229^11800193921^1$	85	$29^1137^13229^18297^165789^15293801^13276118961^1$
36	$2^23^35^17^111^173^1179^1197^1199^1$	86	$2^111437^119609^1153511^14783321^1890220016097^1$
37	$593^178737^11101341^1$	87	$5^1173^120357^141761^121601752001^144560482149^1$
38	$2^137^1179057^19369319^1$	88	$2^33^117^123^143^189^1353^12113^15741^111483^1967724510017^1$
39	$5^1389^133461^14605197^1$	89	$4125636888562548868221559797461449^1$
40	$2^33^117^119^129^141^159^1241^15521^1$	90	$2^15^27^129^131^241^1197^1199^1269^16481^1238321^11529035201^1$
41	1746860020068409^1	91	$13^3181^133461^1807152590055509778144337^1$
42	$2^15^17^213^2239^14663^145697^1$	92	$2^23^147^1229^1643^1700027^1982789^16771937^1150039121^1$
43	$11437^1890220016097^1$	93	$5^161^11301^15021^13272609^121494743413605524661^1$
44	$2^23^123^143^189^1353^15741^111483^1$	94	$2^13761^191962100830401^1489133282872437279^1$
45	$5^229^1197^1269^16481^1238321^1$	95	$29^137^1179057^1425161653868078671 3636125761^1$
46	$2^147^1229^1982789^16771937^1$	96	$2^53^25^17^111^117^197^1193^1577^11153^113729^1665857^19188923201^1$
47	$3761^191962100830401^1$	97	$476098139432320344529305261222389 3281^1$
48	$2^43^25^17^111^117^197^1577^11153^113729^1$	98	$2^113^2239^1293^11471^18109013290449^140710764977973^1$
49	$13^2293^140710764977973^1$	99	$5^1197^15741^152734529^1930527058020124715920 01^1$
50	$2^129^141^11549^129201^145245801^1$	100	$2^23^119^129^141^159^11549^129201^144465699^145245801^146025899^1$

TABLE 1. The prime factorization of $U_n(2,-1)$ for n from 1 to 100.

Proof. By the previous lemma we have that $U_q \equiv_q \Delta^{\frac{q-1}{2}}$. So if $\varepsilon = 0$ (that is, if $\Delta \equiv_q 0$), then $U_{q-\varepsilon} \equiv_q 0 = \varepsilon$. If $\varepsilon = -1$, then $U_q \equiv_q -1$, and so

$$U_{q-\varepsilon} = U_{q+1} = \frac{V_q U_1 + U_q V_1}{2} = \frac{V_q + PU_q}{2} \equiv_q 2^{-1}(P + (-P)) \equiv_q 0.$$

When $\varepsilon = 1$ (so that $U_q \equiv 1$),

$$U_{q-\varepsilon} = U_{q-1} = \frac{-V_q U_1 + U_q V_1}{2Q} = \frac{-V_q + PU_q}{2Q} \equiv_q (2Q)^{-1}(-P + P) \equiv_q 0. \qquad \square$$

5.16. **Lemma.** Let q be prime and let $k \in \mathbb{N}$. If $q^k | U_n$, then $q^{k+1} | U_{qn}$.

Proof. Suppose first that q is odd. Let $r = \frac{U_n}{q^k}$. Then

$$2^{q-1} U_{qn} = \sum_{\substack{i=0 \\ i \text{ odd}}}^{q} \binom{q}{i} V_n^{q-i} U_n^i \Delta^{\frac{i-1}{2}}.$$

When $i > 1$, $q^{k+1} | U_n^i$, so q^{k+1} divides $\binom{q}{i} V_n^{q-i} U_n^i \Delta^{\frac{i-1}{2}}$. And for $i = 1$,

$$\binom{q}{i} V_n^{q-i} U_n^i \Delta^{\frac{i-1}{2}} = q V_n^{q-1} r q^k = q^{k+1}(r V_n^{q-1}).$$

So $q^{k+1} | 2^{q-1} U_{qn}$. Since q is odd, $q^{k+1} | U_{qn}$. Now suppose $2^k | U_n$. By Lemma (5.22), V_n is even. So $U_{2n} = U_n V_n$ is divisible by 2^{k+1}. $\square$

5.17. **Theorem.** For all $m \in \mathbb{N}$ with $\gcd(m, Q) = 1$ there exists $n \in \mathbb{N}$ such that $m | U_n$.

Proof. Let $m \in \mathbb{N}$. Say $m = q_1^{e_1} q_2^{e_2} \cdots q_r^{e_r}$, where the q_i's are distinct. For each i, put $n_i = q_i^{e_i}(q_i - (\Delta/q_i))$. Then $q_i^{e_i}$ divides U_{n_i}. Put $N = n_1 n_2 \cdots n_r$. Since $n_i | N$ and $q_i^{e_i} | U_{n_i}$, $q_i^{e_i} | U_N$ for each i. So $m | U_N$. $\square$

5.18. **Remark.** Now that we know exactly when $\omega(m)$ is defined, a natural question to ask is whether we can find ω by some formula. It could be a prohibitively time-consuming task to exhaustively search for the first n such that $m | U_n$, so it would be nice to have a simpler method for finding ω. There *are* simpler methods, but finding ω will, in general, still require some work. As Lehmer described the problem (in [Leh2]),

> A definite formula for ω is not to be expected any more [than] a formula for the exponent to which a given number c belongs modulo p.

However, he next states that "In certain special cases ω can be given in advance." With N as in the proof of Theorem 5.17, we saw that $m | U_N$. Using Theorem 5.23, ω is then a divisor of N. We can get a better restriction on the possible values for ω: with $n_1, \ldots, n_r$ as in the proof of Theorem 5.17, and N' equal to the least common multiple of $n_1, \ldots, n_r$ (which is bounded above by $N/2^{r-1}$, as we will see in Section 6), we have that $m | U_{N'}$; so ω is some divisor of N'.

Let $\varepsilon = \pm 1$ and suppose that q is an odd prime such that $\frac{q-\varepsilon}{2}$ is also prime.[¶] If $(\Delta/q) = \varepsilon$, then $\omega(q) | (q - \varepsilon)$. The divisors of $q - \varepsilon$ are 1, 2, $\frac{q-\varepsilon}{2}$, and $q - \varepsilon$. Since $1 \not\equiv_q 0$, $\omega(q) \neq 1$. And unless $P \equiv_q 0$, $\omega(q) \neq 2$. In that case $\omega(q)$ is either $q - \varepsilon$ or $\frac{q-\varepsilon}{2}$. In a certain case we can even determine that $\omega(q)$ is $q - \varepsilon$.

5.19. **Theorem.** Let q be an odd prime. If $(Q/q) = -1$, then $\omega(q)$ is even.

5.20. **Corollary.** Let $\varepsilon = \pm 1$. If q and $\frac{q-\varepsilon}{2}$ are primes, and if $P \not\equiv_q 0$, $(\Delta/q) = \varepsilon$, and $(Q/q) = -1$, then $\omega(q) = q - \varepsilon$.

Proof of Theorem 5.19. We use the identity $V_n^2 - \Delta U_n^2 = 4Q^n$. Since $(Q/q) = -1$, we have $\gcd(q, Q) = 1$. So $\omega = \omega(q)$ is defined, and

$$V_\omega^2 \equiv_q V_\omega^2 - \Delta U_\omega^2 \equiv_q 4Q^\omega.$$

Then $(4Q^\omega/q) = 1$, or $(Q^\omega/q) = 1$. If ω is odd (say $\omega = 2k + 1$), then

$$(Q/q) = ((Q^k)^2/q)(Q/q) = (Q^{2k+1}/q) = 1.$$

So ω must be even. $\square$

For the remainder of this section, we fill in the background necessary to cover the primality tests in Section 6.

[¶] If $\varepsilon = 1$, q is then a so-called "safe prime" and $\frac{q-\varepsilon}{2}$ is called a Sophie Germain prime.

5.21. **Lemma.** Let q be an odd prime not dividing Q, and put $\varepsilon = (\Delta/q)$. If $\varepsilon = \pm 1$, then $V_{q-\varepsilon} \equiv_q 2Q^{\frac{1-\varepsilon}{2}}$.

Proof. If $\varepsilon = -1$, then

$$V_{q-\varepsilon} = V_{q-1} \equiv_q (2Q)^{-1}(V_q V_1 - \Delta U_q U_1) \equiv_q (2Q)^{-1}(P^2 - \Delta) \equiv_q (2Q)^{-1}(4Q) \equiv_q 2.$$

With $\varepsilon = 1$, we get

$$V_{q-\varepsilon} = V_{q+1} \equiv_q 2^{-1}(V_q V_1 + \Delta U_q U_1) \equiv_q 2^{-1}(P^2 - \Delta) \equiv_q 2^{-1}(4Q) \equiv_q 2Q. \qquad \square$$

5.22. **Lemma.** If U_n is even, so is V_n.

Proof. From the identity $V_n^2 - \Delta U_n^2 = 4Q^n$, we have that if U_n is even, then 4 divides both ΔU_n^2 and $4Q^n$, and so 4 divides V_n^2. Then V_n is even. $\square$

5.23. **Theorem.** $m|U_n \Rightarrow \omega(m)|n$.

Proof. Suppose $m|U_n$. Say $n = q\omega + r$, where $0 \le r < \omega$. We want to show that $m|U_r$. By the minimality of ω, we would then have that $r = 0$ (and therefore $\omega|n$). Apply (33), with $m = q\omega$:

$$2Q^{q\omega}U_r = 2Q^{q\omega}U_{n-q\omega} = -V_n U_{q\omega} + U_n V_{q\omega}. \tag{37}$$

By the definition of ω, $m|U_\omega$ (hence $m|U_{q\omega}$). And by hypothesis, $m|U_n$. So $m|2Q^{q\omega}U_r$. Now if m is odd, $m|Q^{q\omega}U_r$. If m is even, then so are U_n and $U_{q\omega}$, as well as (by Lemma 5.22) V_n and $V_{q\omega}$. Then $2m|V_n U_{q\omega}$ and $2m|U_n V_{q\omega}$, and so (because of (37)) $2m|2Q^{q\omega}U_r$, and $m|Q^{q\omega}U_r$. $\square$

Notice that Theorem 5.23 gives us that $m|n$ iff $U_m|U_n$. It also gives us a slightly stronger statement:

5.24. **Theorem.** $\gcd(U_m, U_n) = U_{\gcd(m,n)}$

Proof. Put $d = \gcd(m,n)$, $D = \gcd(U_m, U_n)$, and $\omega = \omega(D)$. From Theorem 5.17 (and since D divides both U_m and U_n) we get $\omega|m$ and $\omega|n$. Then $\omega|d$, and so $U_\omega|U_d$. And since $D|U_\omega$ (by the definition of rank of apparition), we have $D|U_d$. Conversely, $U_d|U_n$ and $U_d|U_m$ (since $d|m$ and $d|n$), so $U_d|D$. $\square$

5.25. **Corollary.** For $n \ge 1$, $\gcd(U_n, U_{n+1}) = 1$.

5.26. **Corollary.** For $n \ge 1$, the greatest common divisor of U_n and V_n is either 1 or 2.

Proof. Say $d = \gcd(U_n, V_n)$. Recall that

$$2U_{n+1} = V_n U_1 + U_n V_1 = V_n + PU_n,$$

so $d|2U_{n+1}$. If d has an odd prime factor q, then $q|U_{n+1}$. But then we would have $\gcd(U_n, U_{n+1}) \ge q > 1$. If $d = 2^k$ for some $k \ge 1$, then $2^{k-1}|U_{n+1}$. Since $\gcd(U_n, U_{n+1}) = 1$, we would have $2^{k-1} = 1$ (or $d = 2$). $\square$

The Sequence $\{(V_n, U_n)\}$ in $\mathbb{Z}_m \times \mathbb{Z}_m$. In this section we consider the sequence of pairs (V_n, U_n) in $\mathbb{Z}_m \times \mathbb{Z}_m$. We examine when this sequence is periodic, and how the period length is related to $\omega(m)$. The results from this section will be used in Section 6 to discuss the security of a certain cryptosystem which employs Lucas sequences.

5.27. **Notation.** For $a, b, a', b' \in \mathbb{Z}_m$, we will use $(a,b) \equiv_m (a',b')$ to mean $a \equiv_m a'$ and $b \equiv_m b'$.

5.28. **Remark.** Suppose that m is odd, and for some $n > 0$, $(V_n, U_n) \equiv_m (2,0)$. Then

$$V_{n+1} \equiv_m \frac{1}{2}(V_n V_1 + \Delta U_n U_1) \equiv_m \frac{1}{2}(PV_n + \Delta U_n) \equiv_m \frac{1}{2}(P \cdot 2 + 0) \equiv_m P,$$

and

$$U_{n+1} \equiv_m \frac{1}{2}(V_n U_1 + U_n V_1) \equiv_m \frac{1}{2}(V_n + PU_n) \equiv_m \frac{1}{2}(2 + 0) \equiv_m 1.$$

By induction on

$$\begin{pmatrix} V_{n+2} \\ U_{n+2} \end{pmatrix} \equiv_m P \begin{pmatrix} V_{n+1} \\ U_{n+1} \end{pmatrix} - Q \begin{pmatrix} V_n \\ U_n \end{pmatrix}$$

we have that $(V_{n+k}, U_{n+k}) \equiv_m (V_k, U_k)$ for all k. That is, the sequence $\{(V_i, U_i)\}$ (in $\mathbb{Z}_m$) is periodic, and its period is at most n (actually, the period will be a divisor of n, as we will see shortly).

5.29. **Definition.** Define $\Omega = \Omega(m)$ to be the least positive n (if it exists) such that $(V_n, U_n) \equiv_m (2,0)$.

5.30. **Remark.** As with ω, it is not clear that $\Omega(m)$ is even defined. Recall that $\omega(m)$ is not defined if $\gcd(m, Q) > 1$. Since $\omega(m)$ is the least n such that $U_n \equiv_m 0$, $\gcd(m,Q) > 1$ also implies $(V_n, U_n) \not\equiv_m (2,0)$ for all n. So as with ω, $\Omega(m)$ is undefined when $\gcd(m,Q) > 1$. Similar to Theorem 5.17 for ω, we have:

5.31. **Theorem.** If $\gcd(m, Q) = 1$, then $\Omega(m)$ exists.

Proof. Since $\gcd(m, Q) = 1$, we can define $\omega = \omega(m)$. For all $k \in \mathbb{N}$, we have $U_{k\omega} \equiv_m 0$, and so

$$V_{k\omega}^2 \equiv_m V_{k\omega}^2 - \Delta U_{k\omega}^2 = 4Q^{k\omega}.$$

Then because $\gcd(Q^\omega, m) = 1$, there is some $k < m$ with $Q^{k\omega} = (Q^\omega)^k \equiv_m 1$. So $V_{k\omega}^2 \equiv_m 4$, and

$$V_{2k\omega} = V_{k\omega}^2 - 2Q^{k\omega} \equiv_m 4 - 2 = 2.$$

That is, $(V_{2k\omega}, U_{2k\omega}) \equiv_m (2, 0)$. □

5.32. **Theorem.** If $(V_n, U_n) \equiv_m (2, 0)$, then $\Omega | n$.

Proof. Say $n = q\Omega + r$, where $0 \le r < \Omega$. We consider (V_r, U_r). Since U_n and U_Ω are each divisible by m, there are integers k and l such that $n = k\omega$, $\Omega = l\omega$. Then $r = (k - ql)\omega$, so $U_r \equiv_m 0$. Now

$$\begin{aligned} Q^{q\Omega} V_r &\equiv_m 2^{-1}(V_n V_{q\Omega} - \Delta U_n U_{q\Omega}) \\ &\equiv_m 2^{-1}(2 \cdot 2 - 0) \\ &\equiv_m 2, \end{aligned}$$

and $Q^\Omega \equiv_m 1$ (combine $V_\Omega^2 - \Delta U_\Omega^2 \equiv_m 4Q^\Omega$ with $V_\Omega^2 \equiv_m 4$ and $U_\Omega \equiv_m 0$; then $4Q^\Omega \equiv_m 4$, or $Q^\Omega \equiv_m 1$), so $V_r \equiv_m 2$. Since $0 \le r < \Omega$ and since Ω is minimal, we get $r = 0$ (or $\Omega | n$). □

5.33. **Notation.** If $\gcd(a, m) = 1$, $\mathrm{ord}_m(a)$ is the least positive n such that $a^n \equiv_m 1$.

5.34. **Remark.** Notice that if $(V_n, U_n) \equiv_m (2, 0)$, we have

$$4Q^n = V_n^2 - \Delta U_n^2 \equiv_m 4,$$

or $Q^n \equiv_m 1$. So if $(V_n, U_n) \equiv_m (2, 0)$, $\mathrm{ord}_m(Q)$ divides n. In particular, $\mathrm{ord}_m(Q)$ divides Ω.

5.35. **Theorem.** If $k = \mathrm{ord}_m(Q^\omega)$, then $\Omega = k\omega$ or $2k\omega$.

Proof. Now $U_n \equiv_m 0$ iff n is a multiple of ω. We want to show that $V_{j\omega} \not\equiv_m 2$ for $j < k$, and that $V_{2k\omega} \equiv_m 2$. Then since $\Omega | 2k\omega$ and $\Omega \not< k\omega$, either $\Omega = k\omega$ or $\Omega = 2k\omega$. Since $(Q^\omega)^k \equiv_m 1$, we have

$$V_{k\omega}^2 = \Delta U_{k\omega}^2 + 4Q^{k\omega} \equiv_m 4,$$

and so

$$V_{2k\omega} = V_{k\omega}^2 - 2Q^{k\omega} \equiv_m 4 - 2 = 2.$$

But $(Q^\omega)^j \not\equiv_m 1$ for $j < k$, so

$$V_{j\omega}^2 \equiv_m 4Q^{j\omega} \not\equiv_m 4$$

(or $V_{j\omega} \not\equiv_m 2$). □

5.36. **Corollary.** If $Q = 1$, then $\Omega = \omega$ or 2ω.

5.37. **Remark.** Suppose that $m = q_1^{e_1} \cdots q_k^{e_k}$. Put $\omega = \omega(m)$. In the proof of Theorem 5.19 we saw that $Q^{\omega(q_i)}$ is a quadratic residue mod q_i for each i. Now ω is a multiple of each $\omega(q_i)$, so Q^ω is also a quadratic residue mod q_i for each i. Then $\mathrm{ord}_{q_i}(Q^\omega) \le \frac{q_i - 1}{2} < \frac{q_i}{2}$, and so $\mathrm{ord}_{q_i^{e_i}}(Q^\omega) \le q_i^{e_i - 1}\frac{q_i}{2} = \frac{q_i^{e_i}}{2}$. So

$$\mathrm{ord}_m(Q^\omega) \le \prod_{i=1}^{k} \mathrm{ord}_{q_i^{e_i}}(Q^\omega) \le \prod_{i=1}^{k} \frac{q_i^{e_i}}{2} = \frac{m}{2^k}.$$

With $\omega \le m + 1$, Theorem 5.35 gives us

$$2\left(\mathrm{ord}_m(Q^\omega)\right)\omega \le 2\frac{m}{2^k}(m + 1) = \frac{m(m+1)}{2^{k-1}}$$

as an upper bound for Ω.

6. Some Applications of Lucas Sequences

Primality Testing. Our notation in this section is the same as from Section 5.

6.1. **Definition.** With $m = q_1^{e_1} \cdots q_k^{e_k}$, put $T(m) = n_1 \cdots n_k$ and $S(m) = \text{lcm}(n_1, \ldots, n_k)$, where for each i $n_i = q_i^{e_i - 1}(q_i - (\Delta/q_i))$.*

6.2. **Lemma.** If a_1, a_2, …, a_k are even integers, then $\text{lcm}(a_1, \ldots, a_k)$ divides $\frac{a_1 \cdots a_k}{2^{k-1}}$.

Proof. Use induction on the fact that if a and b are even, then $\text{lcm}(a, b) | \frac{ab}{2}$, and $\frac{ab}{2}$ is even (because $a|(a \cdot \frac{b}{2})$ and $b|(b \cdot \frac{a}{2})$). □

6.3. **Theorem.** If $m > 1$ is odd and $\omega(m) = m \pm 1$, then m is prime.

Proof. We show that: if m has at least two distinct prime factors, then $\omega(m) \neq m \pm 1$; and if $\omega(q^k) = q^k \pm 1$ for some prime q and some positive integer k, then $k = 1$.

Suppose $m = q_1^{e_1} \cdots q_k^{e_k}$, with $k \geq 2$. Notice that $k \geq 2$ and m odd imply $m \geq 15$. For $i = 1, \ldots, k$ (and with $\varepsilon_i = (\Delta/q_i)$), $q^{e_i - 1}(q_i - \varepsilon_i)$ divides $S(m)$, so $U_{S(m)} \equiv_{q_i} 0$. Then $\omega(m) | S(m)$. We claim that no ε_i is zero. If (for the sake of a contradiction) $\varepsilon_j = 0$ for some j, then $\omega(q_j) | (q_j - \varepsilon_j) = q_j$. Since $q_j \neq 1$, we have $\omega(q_j) > 1$, which gives us that $\omega(q_j) = q_j$. But $q_j | U_{m \pm 1}$, so $\omega(q_j) | (m \pm 1)$. This would imply that $q_j | (\prod q_i^{e_i} \pm 1)$ which is impossible. So $\varepsilon_i = \pm 1$ for all i. Then each $q_i^{e_i - 1}(q_i - \varepsilon_i)$ is even, so we can apply the previous lemma to get $S(m) | \frac{T(m)}{2^{k-1}}$. So then $\omega(m) < \frac{T(m)}{2^{k-1}}$. Now

$$\frac{T(m)}{2^{k-1}} = 2^{1-k} \prod_{i=1}^{k} q_i^{e_i - 1}(q_i - \varepsilon_i) \leq 2^{1-k} \prod_{i=1}^{k} q_i^{e_i} \left(1 - \frac{\varepsilon_i}{q_i}\right) = 2m \prod_{i=1}^{k} \frac{1}{2} \left(1 - \frac{\varepsilon_i}{q_i}\right),$$

which is at most

$$2m \prod_{i=1}^{k} \frac{1}{2} \left(1 + \frac{1}{q_i}\right) \leq 2m \prod_{i=1}^{k} \frac{1}{2} \left(1 + \frac{1}{3}\right) = 2m \left(\frac{2}{3}\right)^k \leq 2m \left(\frac{2}{3}\right)^2 = \frac{8}{9} m,$$

so $\omega(m) \leq \frac{8}{9} m$. But since $m \geq 15$, we have $\frac{8}{9} m < m - 1$, and so $\omega(m) \neq m \pm 1$. Now suppose $m = q^k$, where q is prime and $k \geq 1$. Then $q^k \pm 1$ $(= m \pm 1)$ divides $q^k \pm q^{k-1}$, which is possible only if $q^{k-1} = 1$. □

6.4. **Remark.** Notice that the converse of this theorem is false: for example we see in Table 1 that 13 divides U_7 (that is, $\omega(13) \neq \pm 1$).

This next theorem can be considered Lucas' "Fundamental Theorem."

6.5. **Theorem.** Let $m > 1$ be odd and $\varepsilon = \pm 1$, and say

$$m - \varepsilon = \prod_{i=1}^{k} q_i^{e_i}$$

(where the q_i's are prime and distinct). If $U_{m-\varepsilon} \equiv_m 0$ but $U_{\frac{m-\varepsilon}{q_i}} \not\equiv_m 0$ for each q_i, then m is prime.

Proof. If $q | U_{m-\varepsilon}$, then $\omega(m) | m - \varepsilon$. But $\omega(m) \not| \frac{m-\varepsilon}{q_i}$ for $i = 1, \ldots, k$ (because $q \not| U_{\frac{m-\varepsilon}{q_i}}$), so $\omega(m) = m - \varepsilon$. By the last theorem, m is prime. □

6.6. **Corollary.** (Converse of Fermat's Little Theorem)
Let m be odd and $a \in \mathbb{Z}$. If $a^{m-1} \equiv_m 1$ and $a^{\frac{m-1}{q}} \not\equiv_m 1$ for every prime $q | (m - 1)$, then m is prime.

---

*This definition is slightly different from that in [Leh2] (page 424): our $T(n)$ is $\frac{1}{2}$ of Lehmer's $T(n)$ (but we agree on $S(n)$).

Proof. With $P = a + 1$ and $Q = a$, we get

$$\begin{aligned}\Delta &= P^2 - 4Q = a^2 + 2a + 1 - 4a = (a-1)^2,\\ \delta &= a - 1,\\ \alpha &= \frac{P+\delta}{2} = a,\\ \beta &= \frac{P-\delta}{2} = 1,\\ U_n &= \frac{a^n - 1}{a-1}\end{aligned}$$

(we assume $a \neq 1$, since otherwise $\Delta = 0$). If $a \equiv_m 1$, then $a^{\frac{m-1}{d}} \equiv_m 1$ for all $d|(m-1)$; so we assume $a \not\equiv_m 1$. Then $U_n \equiv_m 0$ iff $a^n \equiv_m 1$, so (with $\varepsilon = 1$) the statement of the last theorem is: "if $a^{m-1} \equiv_m 1$ and $a^{\frac{m-1}{q}} \not\equiv_m 1$ for all prime $q|(m-1)$, then m is prime." □

6.7. **Definition.** The number $M_n = 2^n - 1$ is known as the nth *Mersenne number*. If M_n is prime, it is called a *Mersenne prime*.

6.8. **Remark.** Mersenne numbers lend themselves nicely to primality testing. Since the sole prime factor of $M_k + 1$ is 2, we need only check that $U_{M_k+1} \equiv_{M_k} 0$ and $U_{\frac{M_k+1}{2}} \not\equiv_{M_k} 0$ to show that M_k is prime. The next theorem shows that, in testing whether M_n is prime, it suffices to consider the sequence $\{V_{2^n}\}$. From the duplication formulas, we can compute $V_{2^{n+1}}$ as $(V_{2^n})^2 - 2Q^{2^n}$. When $Q = \pm 1$, we have $V_{2^{n+1}} = (V_{2^n})^2 - 2$. Note however, that $Q = 1$ gives us that $(Q/m) = 1$, in which case we can use the theorem to prove that a number is prime, but not that it is composite.

6.9. **Theorem.** Suppose $m = 2^k + \varepsilon$, where $k > 1$ and $\varepsilon = \pm 1$. Then $V_{2^{k-1}} = V_{\frac{m-\varepsilon}{2}} \equiv_m 0$ iff m is prime and $(Q/m) = -1$.

Proof. Suppose m is prime and $(Q/m) = -1$. We know that $U_{m-\varepsilon} \equiv_m 0$, and (since $m - \varepsilon$ is even)

$$U_{m-\varepsilon} = (U_{\frac{m-\varepsilon}{2}})(V_{\frac{m-\varepsilon}{2}}).$$

We also have

$$V_{m-\varepsilon} = (V_{\frac{m-\varepsilon}{2}})^2 - 2Q^{\frac{m-\varepsilon}{2}},$$

or

$$(V_{\frac{m-\varepsilon}{2}})^2 = V_{m-\varepsilon} + 2Q^{\frac{m-\varepsilon}{2}}.$$

From (5.21), $V_{m-\varepsilon} \equiv_m 2Q^{\frac{1-\varepsilon}{2}}$, so

$$(V_{\frac{m-\varepsilon}{2}})^2 \equiv_m 2Q^{\frac{1-\varepsilon}{2}} + 2Q^{\frac{m-\varepsilon}{2}} \equiv_m 2Q^{\frac{1-\varepsilon}{2}}(1 + 2Q^{\frac{m-1}{2}}) \equiv_m 2Q^{\frac{1-\varepsilon}{2}}(1 + (Q/m)).$$

This is zero, since $(Q/m) = -1$. Conversely, suppose $V_{\frac{m-\varepsilon}{2}} \equiv_m 0$. Since the only prime factor of $m - \varepsilon$ is 2, we have that m is prime if $U_{m-\varepsilon} \equiv_m 0$ and $U_{\frac{m-\varepsilon}{2}} \not\equiv_m 0$. Note that $V_{\frac{m-\varepsilon}{2}}$ divides $U_{m-\varepsilon}$, so $U_{m-\varepsilon} \equiv_m 0$. And $\gcd(U_{\frac{m-\varepsilon}{2}}, V_{\frac{m-\varepsilon}{2}}) \leq 2 < m$, so $U_{\frac{m-\varepsilon}{2}} \not\equiv_m 0$. □

As a particular case, we have the next theorem, which is commonly known as the Lucas-Lehmer test for Mersenne primes.

6.10. **Theorem.** Put $S_1 = 4$, and $S_{n+1} = S_n^2 - 2$ for $n > 1$. Then for odd k, M_k is prime iff it divides S_{k-1}.

Proof. Say $k = 2j + 1$. Consider the sequence $\{V_{2^n}\}$ with $P = 2^{j+1}$ and $Q = -1$. We have

$$V_2 = P^2 - 2Q = 2(2^{j+1}) + 2,$$

or (since $2^{2j+1} \equiv_m 1$) $V_{2^1} \equiv_m 4 = S_1$. Also

$$V_{2^{n+1}} = V_{2^n}^2 - 2(-1)^{2^n} = V_{2^n}^2 - 2,$$

so by induction, $V_{2^n} \equiv_m S_n$ for all n. If m divides S_{k-1}, then $V_{2^{k-1}} \equiv_m 0$; so by the previous theorem, m is prime. Conversely, suppose m is prime. Put $\varepsilon = (\Delta/m)$. Now

$$\Delta = P^2 - 4Q = (P^2 - 2Q) - 2Q \equiv_m 4 + 2 = 6,$$

so[†]

$$(\Delta/m) = (6/m) = (2/m)(3/m) = -(2/m)(m/3)$$

(we have $(3/m) = -(m/3)$ because $m \equiv_4 3$). And

$$m = 2^{2j+1} - 1 = 2 \cdot 4^j - 1 \equiv_3 2 - 1 = 1,$$

so $(m/3) = 1$. Since $m \equiv_8 7$, $(2/m) = 1$. Then $\varepsilon = -1$. Because $m \equiv_4 3$, we have $(Q/m) = -1$. So by the previous theorem,

$$S_{k-1} \equiv_m V_{2^{k-1}} \equiv_m V_{\frac{m-\varepsilon}{2}} \equiv_m 0. \qquad \square$$

6.11. **Remark.** An alternate proof can be found in [Leh3], in which the trick is to put $\sigma_n = 2^{2^{n-1}} S_n$. Then $\sigma_{n+1} = \sigma_n^2 - 2 \cdot 2^{2^n}$, much like $V_{2^{n+1}} = V_{2^n}^2 - 2 \cdot Q^{2^n}$. There are infinitely many values of S_1 for which the sequence $\{S_n\}$ (with $S_{n+1} = S_n^2 - 2$) can be used to detect primes of the form $2^k - 1$. Lehmer (in [Leh2], pages 444 and 445) describes how to find some of these (using linear recurrences!), and lists the known values less than 10^9:

$$S_1 = 4,\ 10,\ 52,\ 724,\ 970,\ 10084,\ 95050,\ 140452,\ 1956244,\ 9313930,\ 27246964.$$

Cryptosystems Based on Lucas Sequences. One of the main ideas behind the RSA public key cryptosystem is that for any finite group G, there are $e, d \in \mathbb{Z}$ such that $g^{ed} = g$ for all $g \in G$ (so $ed = k|G| + 1$ for some $k \in \mathbb{Z}$). The integer e is made public, along with certain parameters that specify the group being used (this information is the "public key"). Anyone wishing to encrypt a message M with this public key computes $M' = M^e$, and decryption requires possession of d (the "private key"):

$$(M')^d = (M^e)^d = M^{ed} = M.$$

The other main idea behind RSA is that the security of this system ultimately depends on the difficulty of finding $|G|$: if $|G|$ is known, d can easily be computed (for any e) by solving $xe \equiv_{|G|} 1$ for x. In practice, $G = \mathbb{Z}_N^\times$, where $N = p \cdot q$, and p and q are odd primes. Then $|G| = \phi(N) = (p-1)(q-1)$. An integer $e < N$ is selected with $\gcd(e, \phi(N)) = 1$, and then $d < N$ is found such that $ed \equiv_{\phi(N)} 1$ (the pair (N, e) is the public key, and d is the private key). Given only N and e, finding d requires either good luck or knowledge of the factorization of N. We can choose p and q so that the expected time to factor N, even with the best currently-known algorithms, is arbitrarily long.

In Section 2 we defined (for $r, D \in \mathbb{Z}_m$ nonzero)

$$\mathcal{L}_{\mathbb{Z}_m} = \{(a, b) \in \mathbb{Z}_m \times \mathbb{Z}_m : a^2 - Db^2 \equiv_m r^2\}.$$

Recall that $\mathcal{L}_{\mathbb{Z}_{pq}}$ is an abelian group of order $\psi = (p - (D/p))(q - (D/q))$. Then if $e, d \in \mathbb{Z}$ are such that $ed \equiv_\psi 1$, we have

$$((a, b)^e)^d = (a, b)^{ed} = (a, b).$$

So we can design an RSA-type cryptosystem around the group $\mathcal{L}_{\mathbb{Z}_{pq}}$. Whereas with RSA we can encode a message $M < pq$ directly into $\mathbb{Z}_{pq}$ by inclusion, some care would need to be taken to encode the message M into $\mathcal{L}_{\mathbb{Z}_{pq}}$: for an arbitrary D, $\mathcal{L}_{\mathbb{Z}_{pq}}$ might not contain a point of the form (M, b) or (a, M). Notice that if we put $D = M^2 - r^2$, then

$$M^2 - D \cdot 1^2 = M^2 - (M^2 - r^2) = r^2,$$

so $(M, 1) \in \mathcal{L}_{\mathbb{Z}_{pq}}$. Then we could proceed as before: the message $(M, 1)$ is encrypted as $(X, Y) = (M, 1)^e$, and the message is decrypted as $(X, Y)^d$. However, d depends on D, which depends on the message M. But note that with

$$\psi' = (p-1)(p+1)(q-1)(q+1)$$

(or ψ' equal to any common multiple of those four factors), $|\mathcal{L}_{\mathbb{Z}_{pq}}| = (p - (D/p))(q - (D/q))$ will divide ψ' regardless of D. Even with a proper d in hand, computing $(X, Y)^d$ would still require knowledge of D. If both X and Y are known, we could recover D by computing $Y^{-2}(X^2 - r^2)$ (the computations being performed in $\mathbb{Z}_{pq}$).[‡]

[†]Here we use some standard facts about the Legendre symbol (for p, q odd primes): $(ab/q) = (a/q)(b/q)$; $(2/q) = 1$ if $q \equiv_8 \pm 1$; $(q/p) = (p/q)$ if $q \equiv_4 1$ or $p \equiv_4 1$, and $(q/p) = -(p/q)$ if $q \equiv_4 p \equiv_4 3$. This last fact is known as the Law of Quadratic Reciprocity.

[‡]If e is chosen properly, we can be sure that Y^{-2} is defined.

It might happen that the the original message M is less than $\sqrt{pq}$. In that case, knowing both X and Y is enough to recover M: from X and Y, we can find D; then since $D = M^2 - r^2$, $M = \sqrt{D - r^2}$ (this is the square root in $\mathbb{R}$, not just the modular square root in $\mathbb{Z}_{pq}$). A cryptosystem in which the pair (X, Y) is sent would therefore have an obvious weakness: any message $M < \sqrt{pq}$ could not be securely encrypted. One way around this is to make sure that only messages greater than $\sqrt{pq}$ are encrypted. Another way is to send just X or just Y. By putting certain restrictions on $\mathcal{L}_{\mathbb{Z}_{pq}}$, we can use Lucas sequences to easily recover M from just X, without requiring D.

6.12. **Notation.** For $D, m \in \mathbb{Z}$, let $L(D, m)$ denote the group $\{(a, b) \in \mathbb{Z}_m \times \mathbb{Z}_m : a^2 - Db^2 \equiv_m 4\}$. When the value of D is important we will use $L(D, m)$. We will still let $\mathcal{L}_{\mathbb{Z}_m}$ denote an arbitrary $L(D, m)$.

Recall that the Lucas sequences satisfy $V_n^2 - \Delta U_n^2 = 4Q^n$ (where $\Delta = P^2 - 4Q$). When $Q = 1$, this is

$$V_n^2 - (P^2 - 4)U_n^2 = 4.$$

Now $(V_1, U_1) = (P, 1)$ is the fundamental solution of

$$x^2 - (P^2 - 4)y^2 = 4,$$

and so $(V_n, U_n) = (P, 1)^n$ for all n (here we are exponentiating in the group $\mathcal{L}_{\mathbb{Z}}$, where $r = 2$ and $D = P^2 - 4$). Then also the sequence $\{(V_n, U_n)\}$ (reduced mod pq) coincides with the sequence $\{(P, 1)^n\}$ in $L(\Delta, pq)$. Since $(P, 1)^n \equiv_{pq} (V_n, U_n)$, we can use the duplication formulas to efficiently encrypt and decrypt.
We return to considering the problem of recovering the message M from X. If, instead of $\mathcal{L}_{\mathbb{Z}_{pq}}$, we take our group to be $L(\Delta, pq)$, we can express X using a Lucas sequence: since

$$(X, Y) = (M, 1)^e \equiv_{pq} (V_e(M, 1), U_e(M, 1)),$$

we have $X \equiv_{pq} V_e(M, 1)$. Now recall Equation (36) from Section 5: for integers m and n,

$$V_{mn}(P, Q) = V_m(V_n(P, Q), Q^n);$$

in particular,

$$V_{mn}(P, 1) = V_m(V_n(P, 1), 1).$$

Also recall that

$$(V_{ed}(P, 1), U_{ed}(P, 1)) \equiv_{pq} (P, 1)^{ed} = (P, 1).$$

So

$$M \equiv_{pq} V_{ed}(M, 1) = V_d(V_e(M, 1), 1) \equiv_{pq} V_d(X, 1).$$

The LUC cryptosystem. The LUC cryptosystem is an analog of RSA which uses the groups $\mathcal{L}_{\mathbb{Z}_{pq}}$ (with $r = 2$, and where D depends on the message M). More specifically, it uses the subgroups generated by points of the form $(P, 1)$. Each of these subgroups coincides with the set of terms in the sequence $\{(V_n(P, 1), U_n(P, 1))\}$ (where these V_n and U_n are reduced mod pq). A sketch of LUC can be found in the 1989 book [Bre] (on page 193). LUC was described in more detail in the 1993 paper [Smi-Len]. To make explicit the workings of this cryptosystem:

(a) Choose distinct odd primes p and q. Put $N = pq$ and $\psi' = (p - 1)(p + 1)(q - 1)(q + 1)$. Pick $e < N$ to be coprime with ψ', and find $d < N$ such that $ed \equiv_{\psi'} 1$.
(b) To encrypt the message $M < N$, compute

$$X \equiv_N V_e(M, 1).$$

(c) To decrypt, compute

$$V_d(X, 1) = V_d(V_e(M, 1), 1) = V_{ed}(M, 1) \equiv_N M.$$

6.13. **Remark.** Consider again the abstract version of RSA, and the problem of finding the private key d. One method to search for the private key d is to compute $X = M^e$, where M is some dummy message, and then exhaustively search for a solution to $X^k = M$. Unless M has "small" order, this exhaustive search method is hopeless in practice. So it is nice to know something about the order of arbitrary elements of G.

With LUC, we want to prevent there being points of the form $(a, b)^e \in \mathcal{L}_{\mathbb{Z}_{pq}}$ with "small" order. By picking e to be coprime with ψ', we ensure that any $(a, b)^e$ has the same order as (a, b). We claim that without having the factorization of N, and unless p and q are of a specific form, we cannot hope to have (for all $(a, b) \in \mathcal{L}_{\mathbb{Z}_{pq}}$) an exact formula for the order of (a, b). To see this, first suppose $(a, b) \in \mathcal{L}_{\mathbb{Z}_{pq}}$. Put $P = a$ and $Q = 1$. Then

$$\Delta = P^2 - 4Q = a^2 - 4 = Db^2,$$

and so the map $(x, y) \mapsto (x, \frac{y}{b})$ is an isomorphism between $L(D, pq)$ and $L(\Delta, pq)$. So the order of the point $(a, b) \in L(D, pq)$ is the same as the order of the point $(a, 1) \in L(\Delta, pq)$, which is the same as $\Omega(pq)$ (recall that this is the period of the sequence $\{V_n(a,1), U_n(a,1)\}$ in $\mathbb{Z}_{pq} \times \mathbb{Z}_{pq}$). Theorem 5.35 (with $Q = 1$) states that $\Omega(pq) = \omega(pq)$ or $2\omega(pq)$, so finding Ω is no easier than finding $\omega(pq)$ (which is no easier than finding the order of $a \in \mathbb{Z}_{pq}^{\times}$).

Appendix A: Continued Fractions

Notes. The references used for this appendix were [Khi] and [Har-Wri]. Due to the sparsity of literature regarding nonsimple continued fractions, much of the material herein was developed independently.

Let $a_0 \in \mathbb{R}$, and for $n \in \mathbb{N}$ let a_n and b_n be positive reals. By $\{a_n\}$ and $\{b_n\}$ we mean the sequences $\{a_n\}_{n=0}^{\infty}$ and $\{b_n\}_{n=1}^{\infty}$. For any k, $a_{k-1} + \frac{b_k}{a_k}$ is defined, and so

$$a_{k-2} + \cfrac{b_{k-2}}{a_{k-1} + \cfrac{b_k}{a_k}}$$

is defined, and so

$$a_{k-3} + \cfrac{b_{k-2}}{a_{k-2} + \cfrac{b_{k-2}}{a_{k-1} + \cfrac{b_k}{a_k}}}$$

is defined, and so on. Thus the expression

$$a_0 + \cfrac{b_1}{a_1 + \cfrac{b_2}{a_2 + \dots \cfrac{b_n}{a_n}}} \tag{38}$$

(which we will call a *finite continued fraction*) is defined for all n. We also define an *infinite continued fraction* to be an expression of the form

$$a_0 + \cfrac{b_1}{a_1 + \cfrac{b_2}{a_2 + \cfrac{b_3}{a_3 + \dots}}} \tag{39}$$

(without regard to questions of convergence).

A.1. **Notation.** Denote (38) by

$$a_0 + \frac{b_1}{a_1+}\frac{b_2}{a_2+}\cdots\frac{b_n}{a_n}$$

and let

$$a_0 + \frac{b_1}{a_1+}\frac{b_2}{a_2+}\cdots$$

denote (39).

A.2. **Definition.** For $n = 0, 1, \dots$, put

$$c_n = a_0 + \frac{b_1}{a_1+}\frac{b_2}{a_2+}\cdots\frac{b_n}{a_n}.$$

We call $\{c_n\}$ the sequence of *convergents* (though this sequence need not converge) of the continued fraction.

A.3. **Definition.** Define the sequences $\{p_n\}$ and $\{q_n\}$ by: $p_0 = a_0$, $p_1 = a_0a_1 + b_1$, $q_0 = 1$, $q_1 = a_1$, and

$$\begin{pmatrix} p_n \\ q_n \end{pmatrix} = a_n \begin{pmatrix} p_{n-1} \\ q_{n-1} \end{pmatrix} + b_n \begin{pmatrix} p_{n-2} \\ q_{n-2} \end{pmatrix}$$

for $n > 1$. Because of Claim A.4, we call $\{p_n\}$ the sequence of *numerators* of the convergents, and we call $\{q_n\}$ the sequence of *denominators* of the convergents.

Notice that q_n is positive (since it is a polynomial in $a_1, \ldots, a_n$ with positive coefficients). Also notice that $\{q_n\}$ is eventually strictly increasing if $a_n \geq 1$ for all but finitely many n or if $b_n \geq 1$ for all but finitely many n.

A.4. **Claim.** For all sequences $\{a_n\}$ and $\{b_n\}$, form the sequences $\{c_n\}$, $\{p_n\}$, and $\{q_n\}$ as above. Then $c_n = \frac{p_n}{q_n}$ for all n.

Proof. Let $\{a_n\}$ and $\{b_n\}$ be given. By definition,

$$c_0 = a_0 = \frac{a_0}{1} = \frac{p_0}{q_0},$$
$$c_1 = a_0 + \frac{b_1}{a_1} = \frac{a_0 a_1 + b_1}{a_1} = \frac{p_1}{q_1}.$$

Suppose $k > 1$ and the claim is true for all $j \leq k$. Consider two sequences $\{a'_n\}$ and $\{b'_n\}$ such that $b'_n = b_n$ for $n \leq k$, $a'_n = a_n$ for $n < k$, and $a'_k = a_k + \frac{b_{k+1}}{a_{k+1}}$. Then

$$c_{k+1} = a_0 + \frac{b_1}{a_1+}\frac{b_2}{a_2+}\cdots\frac{b_k}{a_k+}\frac{b_{k+1}}{a_{k+1}} = a_0 + \frac{b_1}{a_1+}\frac{b_2}{a_2+}\cdots\frac{b_k}{a_k + \frac{b_{k+1}}{a_{k+1}}} = c'_k = \frac{p'_k}{q'_k} = \frac{a'_k p'_{k-1} + b'_k p'_{k-2}}{a'_k q'_{k-1} + b'_k q'_{k-2}},$$

which (since $p'_n = p_n$ and $q'_n = q_n$ for $n < k$) is equal to

$$\frac{a'_k p_{k-1} + b_k p_{k-2}}{a'_k q_{k-1} + b_k q_{k-2}} = \frac{\left(a_k + \frac{b_{k+1}}{a_{k+1}}\right) p_{k-1} + b_k p_{k-2}}{\left(a_k + \frac{b_{k+1}}{a_{k+1}}\right) q_{k-1} + b_k q_{k-2}} = \frac{(a_k p_{k-1} + b_k p_{k-2}) + \frac{b_{k+1}}{a_{k+1}} p_{k-1}}{(a_k q_{k-1} + b_k q_{k-2}) + \frac{b_{k+1}}{a_{k+1}} q_{k-1}}$$
$$= \frac{p_k + \frac{b_{k+1}}{a_{k+1}} p_{k-1}}{q_k + \frac{b_{k+1}}{a_{k+1}} q_{k-1}}$$
$$= \frac{a_{k+1} p_k + b_{k+1} p_{k-1}}{a_{k+1} q_k + b_{k+1} q_{k-1}}$$
$$= \frac{p_{k+1}}{q_{k+1}}.$$

By induction, $c_n = \frac{p_n}{q_n}$ for all n. □

The Convergence/Divergence of $\{c_n\}$. Notice that c_{n+2} is a weighted average of c_n and c_{n+1}: with

$$\lambda = \frac{a_{n+2} q_{n+1}}{q_{n+2}} = \frac{a_{n+2} q_{n+1}}{a_{n+2} q_{n+1} + b_{n+2} q_n}$$

(so that $0 < \lambda < 1$),

$$\lambda c_{n+1} + (1-\lambda) c_n = \frac{a_{n+2} q_{n+1}}{q_{n+2}} \cdot \frac{p_{n+1}}{q_{n+1}} + \frac{b_{n+2} q_n}{q_{n+2}} \cdot \frac{p_n}{q_n} = \frac{a_{n+2} p_{n+1} + b_{n+2} p_n}{q_{n+2}} = \frac{p_{n+2}}{q_{n+2}} = c_{n+2}.$$

Now $c_1 = a_0 + \frac{b_1}{a_1} > a_0 = c_0$, and

$$c_0 < c_2 < c_1,$$
$$c_0 < c_2 < c_3 < c_1,$$
$$c_0 < c_2 < c_4 < c_3 < c_1,$$

and inductively we see that $\{c_{2n}\}$ and $\{c_{2n+1}\}$ are (respectively) strictly increasing and strictly decreasing, and are both bounded above by c_1 and bounded below by c_0. So both sequences are convergent. To explore when $\{c_n\}$ converges (that is, when $|c_{n+1} - c_n| \to 0$), we use the following fact:

$$p_{n+1} q_n - p_n q_{n+1} = (-1)^n b_1 b_2 \cdots b_{n+1}. \tag{40}$$

Proof. Notice that for all $n > 0$,

$$\begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix} \begin{pmatrix} a_n & 1 \\ b_n & 0 \end{pmatrix} = \begin{pmatrix} p_{n+1} & p_n \\ q_{n+1} & q_n \end{pmatrix}.$$

For $n > 0$ put $M_n = \begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix}$ and $T_n = \begin{pmatrix} a_n & 1 \\ b_n & 0 \end{pmatrix}$; so

$$M_{n+1} = M_n T_{n+1} = (M_{n-1} T_n) T_{n+1} = \cdots = M_1 T_2 T_3 \cdots T_{n+1},$$

and

$$\begin{aligned} p_{n+1}q_n - p_nq_{n+1} = \det M_{n+1} &= (\det M_1)(\det T_2)(\det T_3)\cdots(\det T_{n+1}) \\ &= (p_1q_0 - p_0q_1)(-b_2)(-b_3)\cdots(-b_{n+1}) \\ &= ((a_0a_1 + b_1) - (a_0a_1))(-b_2)(-b_3)\cdots(-b_{n+1}) \\ &= (-1)^n b_1b_2\cdots b_{n+1}. \end{aligned}$$

□

From this fact we have

$$|c_{n+1} - c_n| = \left|\frac{p_{n+1}}{q_{n+1}} - \frac{p_n}{q_n}\right| = \frac{|p_{n+1}q_n - p_nq_{n+1}|}{q_nq_{n+1}} = \frac{b_1b_2\cdots b_{n+1}}{q_nq_{n+1}}.$$

Put $B_n = b_1b_2\cdots b_n$. Then

$$\begin{aligned} |c_{n+1} - c_n| = \frac{B_{n+1}}{q_{n+1}q_n} = \left(\frac{b_{n+1}}{q_{n+1}}q_{n-1}\right)\left(\frac{B_n}{q_nq_{n-1}}\right) &= \left(\frac{b_{n+1}}{q_{n+1}}q_{n-1}\right)\left(\frac{b_n}{q_n}q_{n-2}\right)\left(\frac{B_{n-1}}{q_{n-1}q_{n-2}}\right) \\ &\vdots \\ &= \left(\frac{b_{n+1}}{q_{n+1}}q_{n-1}\right)\cdots\left(\frac{b_2}{q_2}q_0\right)\left(\frac{B_1}{q_1q_0}\right) \\ &= \frac{B_1}{q_1q_0}\prod_1^n \frac{b_{k+1}}{q_{k+1}}q_{k-1} \\ &= \frac{b_1}{q_1}\prod_1^n b_{k+1}\left(\frac{q_{k+1}}{q_{k-1}}\right)^{-1} \\ &= \frac{b_1}{q_1}\prod_1^n b_{k+1}\left(a_{k+1}\frac{q_k}{q_{k-1}} + b_{k+1}\right)^{-1} \\ &= \frac{b_1}{q_1}\prod_1^n \left(1 + \frac{a_{k+1}}{b_{k+1}}\frac{q_k}{q_{k-1}}\right)^{-1}. \end{aligned}$$

So $\{c_n\}$ converges iff $\prod(1 + \frac{a_{k+1}}{b_{k+1}}\frac{q_k}{q_{k-1}}) \to \infty$, iff $\sum \frac{a_{k+1}}{b_{k+1}}\frac{q_k}{q_{k-1}} \to \infty$.* In particular, we have the following:

A.5. **Theorem.** When $\{a_n\}, \{b_n\} \subset \mathbb{N}$, $\{c_n\}$ converges if the sequence $\{\frac{b_n}{a_n}\}$ is bounded.

Simple Continued Fractions. When each a_n is an integer (and $a_n > 0$ for $n > 0$), an expression of the form

$$a_0 + \frac{1}{a_1+}\frac{1}{a_2+}\cdots\frac{1}{a_n}$$

is a *(finite) simple continued fraction*, and the expression

$$a_0 + \frac{1}{a_1+}\frac{1}{a_2+}\cdots$$

is an *infinite simple continued fraction*. Since $\frac{1}{a_n}$ is bounded for each n, the sequence $\{c_n\}$ converges.

A.6. **Fact.** Every irrational $x \in \mathbb{R}$ has an expression as an infinite simple continued fraction.

Proof. Let $x \in \mathbb{R}\backslash\mathbb{Q}$. Put $a_0 = \lfloor x \rfloor$ and $r_1 = x - a_0$, and $a_1 = \lfloor \frac{1}{r_1} \rfloor$. Having found $a_0, a_1, \ldots, a_{n-1}$ and $r_1, \ldots, r_{n-1}$, put $r_n = \frac{1}{r_{n-1}} - a_{n-1}$ and $a_n = \lfloor \frac{1}{r_n} \rfloor$. Each c_n is a rational function of $a_0, a_1, \ldots, a_n$; so since $x \notin \mathbb{Q}$, each r_n is nonzero (and so the sequence $\{a_n\}$ is infinite). Notice that $c_0 < x < c_1$. If $c_n \nrightarrow x$, then for some n either $c_{2n} < x < c_{2n+2}$ (so $x < c_{2n+2} < c_{2n+1}$) or $c_{2n+1} < x < c_{2n-1}$ (so $c_{2n} < c_{2n+1} < x$). But

$$\begin{aligned} x = a_0 + r_1 = a_0 + \frac{1}{a_1 + r_2} = a_0 + \frac{1}{a_1 + \dfrac{1}{a_2 + r_3}} = \cdots &= a_0 + \frac{1}{a_1+}\frac{1}{a_2+}\cdots\frac{1}{a_n + r_{n+1}} \\ &= a_0 + \frac{1}{a_1+}\frac{1}{a_2+}\cdots\frac{1}{a_n + \frac{1}{1/r_{n+1}}}, \end{aligned}$$

*Recall that when $\{\alpha_n\}$ is a sequence of positive reals, each less than 1, the product $\prod(1 + \alpha_n)$ converges/diverges iff the sum $\sum \alpha_n$ converges/diverges.

and so for all n,

$$x = \frac{r_{n+1}p_n + p_{n-1}}{r_{n+1}q_n + q_{n-1}} = \frac{r_{n+1}q_n}{r_{n+1}q_n + q_{n-1}}\frac{p_n}{q_n} + \frac{q_{n-1}}{r_{n+1}q_n + q_{n-1}}\frac{p_{n-1}}{q_{n-1}}$$

(that is, x is strictly between c_{n-1} and c_n). So $c_n \to x$.

A.7. **Fact.** Let $x \in \mathbb{R}$. For each convergent $\frac{p_n}{q_n}$,

$$\left|\frac{p_n}{q_n} - x\right| < \frac{1}{q_n^2}.$$

Proof. Since x lies strictly between c_n and c_{n+1},

$$|c_n - x| < |c_n - c_{n+1}| = \frac{|p_{n+1}q_n - p_nq_{n+1}|}{|p_np_{n+1}|} = \frac{1}{p_np_{n+1}} < \frac{1}{p_n^2}.$$

Putting Facts (A.6) and (A.7) together, we have:

A.8. **Corollary.** For every irrational x, there are infinitely many pairs (p, q) of coprime integers such that

$$\left|\frac{p}{q} - x\right| < \frac{1}{q}.$$

Appendix B: Finite Fields

Notes. Our treatment of finite fields follows that of [Ire-Ros].

The goal of this section is Theorem B.3, the statement of which is that every finite field is cyclic. We begin with some standard facts before we discuss the construction of finite fields.

Fact. If $\mathbb{F}$ is a finite field, then for some prime q, $qa = 0$ for every $a \in \mathbb{F}$.

Fact. If $\mathbb{F}$ is a finite field, then for some prime q and some $k \in \mathbb{N}$, $|\mathbb{F}| = q^k$.

Fact. If $\mathbb{F}$ and $\mathbb{F}'$ are finite fields and $|\mathbb{F}| = |\mathbb{F}'|$, then $\mathbb{F} \cong \mathbb{F}'$.

So if $\mathbb{F}$ is a field of q^k elements, we can consider $\mathbb{F}$ to be "the" field of q^k elements, which we denote as $\mathbb{F}_{q^k}$. In particular, since $\mathbb{Z}_q$ is a field of q elements, we sometimes use $\mathbb{Z}_q$ and $\mathbb{F}_q$ interchangeably.

Constructing $\mathbb{F}_{q^k}$. We assume $k \geq 2$. Let $f(x)$ be a k-degree irreducible polynomial in $\mathbb{F}_q[x]$, say

$$f(x) = x^{k-1} + c_{k-2}x^{k-2} + \cdots + c_2x^2 + c_1x + c_0.$$

Define the symbol α to be a root of $f(x)$. Then $\mathbb{F}_{q^k}$ is the set of polynomials (over $\mathbb{F}_q$) in α and of degree less than k: each element of $\mathbb{F}_{q^k}$ is of the form

$$b_0 + b_1\alpha + b_2\alpha^2 + \cdots + b_{k-1}\alpha^{k-1},$$

where each $b_i \in \mathbb{F}_q$. Addition and multiplication are as with polynomials over $\mathbb{F}_q$, and then reduced according to the rule

$$\alpha^{k-1} = -c_{k-2}\alpha^{k-2} - \cdots - c_2\alpha^2 - c_1\alpha - c_0.$$

Another way of describing this construction is that we consider $\mathbb{F}_{q^k}$ to be the quotient ring $\mathbb{F}_q[x]/(f(x))$.

Example. In $\mathbb{F}_3$, $x^2 - 2$ is irreducible (that is, the Legendre symbol $(2/3)$ equals -1). Define α to be a square root of 2 in $\mathbb{F}_3$. The elements of (our version of) $\mathbb{F}_9$ are:

$$0, 1, 2, \alpha, 1 + \alpha, 2 + \alpha, 2\alpha, 1 + 2\alpha, 2 + 2\alpha.$$

We add and multiply these elements as with polynomials over $\mathbb{F}_3$, and then reduce by the rule $\alpha^2 = 2$; thus we have for example

$$(1 + \alpha)(2 + \alpha) = 2 + 3\alpha + \alpha^2 = 2 + \alpha^2 = 2 + 2 = 1.$$

Definition. Define $\phi : \mathbb{N} \to \mathbb{N}$ as

$$\phi(n) = |\{a \in \mathbb{N} : a \leq n \text{ and } \gcd(a, n) = 1\}|.$$

This is known as *Euler's ϕ function*, or as the *totient function*.

For example, we have $\phi(1) = 1$, $\phi(2) = 1$, $\phi(3) = 2$, $\phi(4) = 2$, $\phi(5) = 4$, $\phi(6) = 2$. It is well-known that a formula for ϕ is

$$\phi(n) = \prod p_i^{e_i - 1}(p_i - 1),$$

where $\prod p_i^{e_i}$ is the prime factorization of n. In particular, $\phi(n) \geq 2$ when $n \geq 3$.

B.1. **Fact.** $\sum_{d|n} \phi(d) = n$.

B.2. **Lemma.** If $d|(q^k - 1)$, then $x^d - 1$ has d distinct roots in $\mathbb{F}_{q^k}$.

Proof. Say $N = q^k - 1$. For every $a \in \mathbb{F}_{q^k}$, $a^{q^k} = a$ ($a^N = 1$ for $a \in \mathbb{F}_{q^k}^\times$, and $0^{q^k} = 0$). So there are $q^k = |\mathbb{F}_{q^k}|$ distinct roots of $x^{q^k} - x$, and so $x^N - 1$ has N distinct roots. If $d|N$ (say $N = kd$), then

$$(x^d - 1)(x^{d(k-1)} + x^{d(k-2)} + \cdots + x^d + 1) = x^{kd} - 1 = x^N - 1.$$

Let i and j be the number of distinct roots of (respectively) $x^d - 1$ and

$$x^{d(k-1)} + x^{d(k-2)} + \cdots + x^d + 1.$$

Now i is at most d and j is at most $d(k-1)$. Since $x^N - 1$ has exactly $kd = d + d(k-1)$ distinct roots, we must have $i = d$ (as well as $j = d(k-1)$). □

Notation. For $a \in \mathbb{F}^\times$, let $\mathrm{ord}(a)$ denote the least positive n such that $a^n = 1$.

B.3. **Theorem.** $\mathbb{F}_{q^k}^\times$ is cyclic.

Proof. For all $n \in \mathbb{N}$ put $C_n = \{a \in \mathbb{F}_{q^k}^\times : a^n = 1\}$. Put $N = q^k - 1$. For each n put

$$\psi(n) = |\{a \in \mathbb{F}_{q^k}^\times : \mathrm{ord}(a) = n\}|.$$

Suppose $n|N$. By the lemma, $x^n - 1$ has n distinct roots (so $|C_n| = n$). Also,

$$|C_n| = \sum_{d|n} \psi(d).$$

So for all $n|N$,

$$n = \sum_{d|n} \psi(d).$$

We claim that $\psi(n) = \phi(n)$ for all $n|N$, which we show by induction. Obviously $\psi(1) = 1 = \phi(1)$. Now suppose $n|N$, $n \geq 1$, and $\psi(d) = \phi(d)$ for all $d < n$ which divide n. Then

$$\psi(n) = n - \sum_{\substack{d|n \\ d<n}} \psi(d) = n - \sum_{\substack{d|n \\ d<n}} \phi(d) = n - (n - \phi(n)) = \phi(n).$$

In particular, $\psi(n) > 0$; so there is some $g \in \mathbb{F}_{q^k}^\times$ with $\mathrm{ord}(g) = N$. □

References

[Bar] E. Barbeau, *Pell's Equation*, Springer-Verlag, New York, 2003.

[Bel] E.T. Bell, *A Revision of the Algebra of Lucas Functions*, Ann. of Math. (3), 36 (1935), 733-742.

[Bre] D. Bressoud, *Factorization and Primality Testing*, Springer-Verlag, New York, 1989.

[Bur] D. Burton, *Elementary Number Theory, Fourth Edition*, McGraw-Hill, New York, 1998.

[Car] R.D. Carmichael, *On the Numerical Factors of the Arithmetic Forms* $\alpha^2 \pm \beta^n$. Ann. of Math. (2), Vol. 15, No. 1 (1913), 30-70.

[Dic] L.E. Dickson, *History of the Theory of Numbers* (3 vols.), 1919, 1920, and 1923. Reprinted by Chelsea Pub. Co., New York, 1952.

[Edw] H.M. Edwards, *Fermat's Last Theorem: A Genetic Introduction to Algebraic Number Theory*, Springer-Verlag, New York, 1977.

[Gau] C.F. Gauss, *Disquisitiones Arithmeticae*, 1801. Reprinted in English translation by Yale University Press, New Haven, CT, 1966.

[Har-Wri] G.H. Hardy and E.M. Wright, *An Introduction to the Theory of Numbers, Fifth Edition*, Oxford, 1979.

[Hun] T. Hungerford, *Algebra*, Springer-Verlag, New York, 1974.

[Ire-Ros] K. Ireland and M. Rosen, *A Classical Introduction to Modern Number Theory, Second Edition*, Springer-Verlag, New York, 1990.

[Kap] I. Kaplansky, *Lucas' Tests for Mersenne Numbers*, Amer. Math. Monthly 52, (1945), 188-190.

[Khi] A.Ya. Khinchin, *Continued Fractions*, 1964. Reprinted by Dover, New York, 1997.
[Kno] K. Knopp, *Infinite Sequences and Series*, Dover, New York, 1956.
[Leh1] D.H. Lehmer, *Tests for Primality by the Converse of Fermat's Theorem*, Bull. AMS, 33 (1927), 327-340.
[Leh2] D.H. Lehmer, *An Extended Theory of Lucas' Functions*, Ann. of Math. (2), 31 (1930), 419-448.
[Leh3] D.H. Lehmer, *On Lucas' Test for the Primality of Mersenne's Numbers*, J. Lon. Math. Soc., 10, (1935), 162-165.
[Luc] E. Lucas, *Théorie des Fonctions Numériques Simplement Périodiques*, Amer. J. Math., 1 (1878), 184-240, 289-321.
[Rez] B. Reznick, *One Introduction to Mathematical Research*, Retrieved from http://www.math.uiuc.edu/ reznick/mhori.pdf.
[RSA] Rivest, Shamir, Adleman, *A Method for Obtaining Digital Signatures and Public Key Cryptosystems*, Comm. ACM, Vol. 21, (1978) 120-126.
[Smi-Len] P. Smith and M.J.J. Lennon, *LUC: A New Public Key System*, In Ninth IFIP Symposium on Computer Security, Elsevier Science Publishers (1993), 103-117.
[Wil1] H.C. Williams, *Édouard Lucas and Primality Testing*, Vol. 22 of *Canadian Mathematical Society Series of Monographs and Advanced Texts*, Wiley, New York, 1998.
[Wil2] H.C. Williams, *A $p+1$ Method of Factoring*, Math. Comp., 39, No. 159 (1982), 225-234.