Faculté des Sciences et Techniques (FAST), Bamako, Mali
École de recherche CIMPA Théorie des Nombres et Algorithmique
Updated: *February 18, 2016*

Pell's equation

Michel Waldschmidt
This text is available on the internet at the address
http://www.math.jussieu.fr/∼miw/articles/pdf/BamakoPell2010.pdf

Contents

[1]

1 On the so–called Pell–Fermat equation

Let D be a positive integer which is not the square of an integer. It follows that $\sqrt{D}$ is an irrational number. The Diophantine equation

$$x^2 - Dy^2 = \pm 1, \tag{1}$$

where the unknowns x and y are in $\mathbf{Z}$, is called *Pell's equation*.

An introduction to the subject has been given in first lecture:
`http://www.math.jussieu.fr/~miw/articles/pdf/PellFermat2010.pdf`
and
`http://www.math.jussieu.fr/~ miw/articles/pdf/PellFermat2010VI.pdf`

Here we supply complete proofs of the results introduced in that lecture.

1.1 Examples of simple continued fractions

The three first examples below are special cases of results initiated by O. Perron [23] and related with real quadratic fields of Richaud-Degert type.

Example 1. Take $D = a^2b^2 + 2b$ where a and b are positive integers. A solution to

$$x^2 - (a^2b^2 + 2b)y^2 = 1$$

is $(x, y) = (a^2b + 1, a)$. As we shall see, this is related with the continued fraction expansion of $\sqrt{D}$ which is

$$\sqrt{a^2b^2 + 2b} = [ab, \overline{a, 2ab}]$$

since

$$t = \sqrt{a^2b^2 + 2b} \Longleftrightarrow t = ab + \cfrac{1}{a + \cfrac{1}{t + ab}}\cdot$$

This includes the examples $D = a^2 + 2$ (take $b = 1$) and $D = b^2 + 2b$ (take $a = 1$). For $a = 1$ and $b = c - 1$, this includes the example $D = c^2 - 1$.

Example 2. Take $D = a^2b^2 + b$ where a and b are positive integers. A solution to

$$x^2 - (a^2b^2 + b)y^2 = 1$$

[1] Les notes sont en anglais, mais les cours seront donnés en français.

is $(x, y) = (2a^2b + 1, 2a)$. The continued fraction expansion of $\sqrt{D}$ is

$$\sqrt{a^2b^2 + b} = [ab, \overline{2a, 2ab}]$$

since

$$t = \sqrt{a^2b^2 + b} \Longleftrightarrow t = ab + \frac{1}{2a + \dfrac{1}{t + ab}}\cdot$$

This includes the example $D = b^2 + b$ (take $a = 1$).

The case $b = 1$, $D = a^2 + 1$ is special: there is an integer solution to

$$x^2 - (a^2 + 1)y^2 = -1,$$

namely $(x, y) = (a, 1)$. The continued fraction expansion of $\sqrt{D}$ is

$$\sqrt{a^2 + 1} = [a, \overline{2a}]$$

since

$$t = \sqrt{a^2 + 1} \Longleftrightarrow t = a + \frac{1}{t + a}\cdot$$

Example 3. Let a and b be two positive integers such that $b^2 + 1$ divides $2ab + 1$. For instance $b = 2$ and $a \equiv 1 \pmod 5$. Write $2ab + 1 = k(b^2 + 1)$ and take $D = a^2 + k$. The continued fraction expansion of $\sqrt{D}$ is

$$[a, \overline{b, b, 2a}]$$

since $t = \sqrt{D}$ satisfies

$$t = a + \frac{1}{b + \dfrac{1}{b + \dfrac{1}{a + t}}} = [a, b, b, a + t].$$

A solution to $x^2 - Dy^2 = -1$ is $x = ab^2 + a + b$, $y = b^2 + 1$.

In the case $a = 1$ and $b = 2$ (so $k = 1$), the continued fraction has period length 1 only:

$$\sqrt{5} = [1, \overline{2}].$$

Example 4. Integers which are *Polygonal numbers* in two ways are given by the solutions to quadratic equations.

Triangular numbers are numbers of the form

$$1 + 2 + 3 + \cdots + n = \frac{n(n + 1)}{2} \quad \text{for } n \geq 1;$$

their sequence starts with

1, 3, 6, 10, 15, 21, 28, 36, 45, 55, 66, 78, 91, 105, 120, 136, 153, 171, ...

`http://www.research.att.com/~njas/sequences/A000217`.

Square numbers are numbers of the form

$$1 + 3 + 5 + \cdots + (2n + 1) = n^2 \quad \text{for } n \geq 1;$$

their sequence starts with

1, 4, 9, 16, 25, 36, 49, 64, 81, 100, 121, 144, 169, 196, 225, 256, 289, ...

`http://www.research.att.com/~njas/sequences/A000290`.

Pentagonal numbers are numbers of the form

$$1 + 4 + 7 + \cdots + (3n + 1) = \frac{n(3n - 1)}{2} \quad \text{for } n \geq 1;$$

their sequence starts with

1, 5, 12, 22, 35, 51, 70, 92, 117, 145, 176, 210, 247, 287, 330, 376, 425, ...

`http://www.research.att.com/~njas/sequences/A000326`.

Hexagonal numbers are numbers of the form

$$1 + 5 + 9 + \cdots + (4n + 1) = n(2n - 1) \quad \text{for } n \geq 1;$$

their sequence starts with

1, 6, 15, 28, 45, 66, 91, 120, 153, 190, 231, 276, 325, 378, 435, 496, 561, ...

`http://www.research.att.com/~njas/sequences/A000384`.

For instance, numbers which are at the same time triangular and squares are the numbers y^2 where (x, y) is a solution to Pell's equation with $D = 8$. Their list starts with

$$0, 1, 36, 1225, 41616, 1413721, 48024900, 1631432881, 55420693056, \ldots$$

See `http://www.research.att.com/~njas/sequences/A001110`.

Example 5. Integer rectangle triangles having sides of the right angle as consecutive integers a and $a + 1$ have an hypothenuse c which satisfies $a^2 + (a + 1)^2 = c^2$. The admissible values for the hypothenuse is the set of positive integer solutions y to Pell's equation $x^2 - 2y^2 = -1$. The list of these hypothenuses starts with

$$1, 5, 29, 169, 985, 5741, 33461, 195025, 1136689, 6625109, 38613965,$$

See `http://www.research.att.com/~njas/sequences/A001653`.

1.2 Existence of integer solutions

Let D be a positive integer which is not a square. We show that Pell's equation (1) has a non–trivial solution $(x, y) \in \mathbf{Z} \times \mathbf{Z}$, that is a solution $\neq (\pm 1, 0)$.

Proposition 2. *Given a positive integer D which is not a square, there exists $(x, y) \in \mathbf{Z}^2$ with $x > 0$ and $y > 0$ such that $x^2 - Dy^2 = 1$.*

Proof. The first step of the proof is to show that there exists a non–zero integer k such that the Diophantine equation $x^2 - Dy^2 = k$ has infinitely many solutions $(x, y) \in \mathbf{Z} \times \mathbf{Z}$. The main idea behind the proof, which will be made explicit in Lemmas 4, 5 and Corollary 6 below, is to relate the integer solutions of such a Diophantine equation with rational approximations x/y of $\sqrt{D}$.

Using the implication (i) $\Rightarrow$ (v) of the irrationality criterion 61 and the fact that $\sqrt{D}$ is irrational, we deduce that there are infinitely many $(x, y) \in \mathbf{Z} \times \mathbf{Z}$ with $y > 0$ (and hence $x > 0$) satisfying

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

For such a (x, y), we have $0 < x < y\sqrt{D} + 1 < y(\sqrt{D} + 1)$, hence

$$0 < |x^2 - Dy^2| = |x - y\sqrt{D}| \cdot |x + y\sqrt{D}| < 2\sqrt{D} + 1.$$

Since there are only finitely integers $k \neq 0$ in the range

$$-(2\sqrt{D} + 1) < k < 2\sqrt{D} + 1,$$

one at least of them is of the form $x^2 - Dy^2$ for infinitely many (x, y).

The second step is to notice that, since the subset of $(x, y) \pmod k$ in $(\mathbf{Z}/k\mathbf{Z})^2$ is finite, there is an infinite subset $E \subset \mathbf{Z} \times \mathbf{Z}$ of these solutions to $x^2 - Dy^2 = k$ having the same $(x \pmod k, y \pmod k)$.

Let (u_1, v_1) and (u_2, v_2) be two distinct elements in E. Define $(x, y) \in \mathbf{Q}^2$ by

$$x + y\sqrt{D} = \frac{u_1 + v_1\sqrt{D}}{u_2 + v_2\sqrt{D}}.$$

From $u_2^2 - Dv_2^2 = k$, one deduces

$$x + y\sqrt{D} = \frac{1}{k}(u_1 + v_1\sqrt{D})(u_2 - v_2\sqrt{D}),$$

hence

$$x = \frac{u_1u_2 - Dv_1v_2}{k}, \quad y = \frac{-u_1v_2 + u_2v_1}{k}.$$

From $u_1 \equiv u_2 \pmod{k}$, $v_1 \equiv v_2 \pmod{k}$ and

$$u_1^2 - Dv_1^2 = k, \quad u_2^2 - Dv_2^2 = k,$$

we deduce

$$u_1u_2 - Dv_1v_2 \equiv u_1^2 - Dv_1^2 \equiv 0 \pmod{k}$$

and

$$-u_1v_2 + u_2v_1 \equiv -u_1v_1 + u_1v_1 \equiv 0 \pmod{k},$$

hence x and y are in $\mathbf{Z}$. Further,

$$\begin{aligned} x^2 - Dy^2 &= (x + y\sqrt{D})(x - y\sqrt{D}) \\ &= \frac{(u_1 + v_1\sqrt{D})(u_1 - v_1\sqrt{D})}{(u_2 + v_2\sqrt{D})(u_2 - v_2\sqrt{D})} \\ &= \frac{u_1^2 - Dv_1^2}{u_2^2 - Dv_2^2} = 1. \end{aligned}$$

It remains to check that $y \neq 0$. If $y = 0$ then $x = \pm 1$, $u_1v_2 = u_2v_1$, $u_1u_2 - Dv_1v_2 = \pm 1$, and

$$ku_1 = \pm u_1(u_1u_2 - Dv_1v_2) = \pm u_2(u_1^2 - Dv_1^2) = \pm ku_2,$$

which implies $(u_1, u_2) = (v_1, v_2)$, a contradiction.

Finally, if $x < 0$ (resp. $y < 0$) we replace x by $-x$ (resp. y by $-y$).

□

Once we have a non–trivial integer solution (x, y) to Pell's equation, we have infinitely many of them, obtained by considering the powers of $x+y\sqrt{D}$.

1.3 All integer solutions

There is a natural order for the positive integer solutions to Pell's equation which can be defined in several ways: we can order them by increasing values of x, or increasing values of y, or increasing values of $x + y\sqrt{D}$ - it is easily checked that the order is the same.

It follows that there is a minimal positive integer solution[2] (x_1, y_1), which is called *the fundamental solution to Pell's equation* $x^2 - Dy^2 = \pm 1$. In the same way, there is a fundamental solution to Pell's equations $x^2 - Dy^2 = 1$.

[2]We use the letter x_1, which should not be confused with the first complete quotient in the section 2.2.2 on continued fractions

Proposition 3. *Denote by (x_1, y_1) the fundamental solution to Pell's equation $x^2 - Dy^2 = \pm 1$. Then the set of all positive integer solutions to this equation is the sequence $(x_n, y_n)_{n\geq 1}$, where x_n and y_n are given by*

$$x_n + y_n\sqrt{D} = (x_1 + y_1\sqrt{D})^n, \quad (n \in \mathbf{Z}, \quad n \geq 1).$$

In other terms, x_n and y_n are defined by the recurrence formulae

$$x_{n+1} = x_n x_1 + D y_n y_1 \qquad \text{and} \qquad y_{n+1} = x_1 y_n + x_n y_1, \qquad (n \geq 1).$$

More explicitly:
• *If $x_1^2 - Dy_1^2 = 1$, then (x_1, y_1) is the fundamental solution to Pell's equation $x^2 - Dy^2 = 1$, and there is no integer solution to Pell's equation $x^2 - Dy^2 = -1$.*
• *If $x_1^2 - Dy_1^2 = -1$, then (x_1, y_1) is the fundamental solution to Pell's equation $x^2 - Dy^2 = -1$, and the fundamental solution to Pell's equation $x^2 - Dy^2 = 1$ is (x_2, y_2). The set of positive integer solutions to Pell's equation $x^2 - Dy^2 = 1$ is $\{(x_n, y_n)\ ;\ n \geq 2 \ \ even\}$, while the set of positive integer solutions to Pell's equation $x^2 - Dy^2 = -1$ is $\{(x_n, y_n)\ ;\ n \geq 1 \ \ odd\}$. The set of all solutions $(x, y) \in \mathbf{Z} \times \mathbf{Z}$ to Pell's equation $x^2 - Dy^2 = \pm 1$ is the set $(\pm x_n, y_n)_{n\in\mathbf{Z}}$, where x_n and y_n are given by the same formula*

$$x_n + y_n\sqrt{D} = (x_1 + y_1\sqrt{D})^n, \quad (n \in \mathbf{Z}).$$

The trivial solution $(1, 0)$ is (x_0, y_0), the solution $(-1, 0)$ is a torsion element of order 2 in the group of units of the ring $\mathbf{Z}[\sqrt{D}]$.

Proof. Let (x, y) be a positive integer solution to Pell's equation $x^2 - Dy^2 = \pm 1$. Denote by $n \geq 0$ the largest integer such that

$$(x_1 + y_1\sqrt{D})^n \leq x + y\sqrt{D}.$$

Hence $x + y\sqrt{D} < (x_1 + y_1\sqrt{D})^{n+1}$. Define $(u, v) \in \mathbf{Z} \times \mathbf{Z}$ by

$$u + v\sqrt{D} = (x + y\sqrt{D})(x_1 - y_1\sqrt{D})^n.$$

From

$$u^2 - Dv^2 = \pm 1 \quad \text{and} \quad 1 \leq u + v\sqrt{D} < x_1 + y_1\sqrt{D},$$

we deduce $u = 1$ and $v = 0$, hence $x = x_n$, $y = y_n$.

□

1.4 On the group of units of $\mathbf{Z}[\sqrt{D}]$

Let D be a positive integer which is not a square. The ring $\mathbf{Z}[\sqrt{D}]$ is the subring of $\mathbf{R}$ generated by $\sqrt{D}$. The map $\sigma : z = x + y\sqrt{D} \longmapsto x - y\sqrt{D}$ is the *Galois automorphism* of this ring. The *norm* $N : \mathbf{Z}[\sqrt{D}] \longrightarrow \mathbf{Z}$ is defined by $N(z) = z\sigma(z)$. Hence

$$N(x + y\sqrt{D}) = x^2 - Dy^2.$$

The restriction of N to the group of unit $\mathbf{Z}[\sqrt{D}]^\times$ of the ring $\mathbf{Z}[\sqrt{D}]$ is a homomorphism from the multiplicative group $\mathbf{Z}[\sqrt{D}]^\times$ to the group of units $\mathbf{Z}^\times$ of $\mathbf{Z}$. Since $\mathbf{Z}^\times = \{\pm 1\}$, it follows that

$$\mathbf{Z}[\sqrt{D}]^\times = \{z \in \mathbf{Z}[\sqrt{D}] \ ; \ N(z) = \pm 1\},$$

hence $\mathbf{Z}[\sqrt{D}]^\times$ is nothing else than the set of $x + y\sqrt{D}$ when (x, y) runs over the set of integer solutions to Pell's equation $x^2 - Dy^2 = \pm 1$.

Proposition 2 means that $\mathbf{Z}[\sqrt{D}]^\times$ is not reduced to the torsion subgroup ± 1, while Proposition 3 gives the more precise information that this group $\mathbf{Z}[\sqrt{D}]^\times$ is a (multiplicative) abelian group of rank 1: there exists a so–called *fundamental unit* $u \in \mathbf{Z}[\sqrt{D}]^\times$ such that

$$\mathbf{Z}[\sqrt{D}]^\times = \{\pm u^n \ ; \ n \in \mathbf{Z}\}.$$

The fundamental unit $u > 1$ is $x_1 + y_1\sqrt{D}$, where (x_1, y_1) is the fundamental solution to Pell's equation $x^2 - Dy^2 = \pm 1$. Pell's equation $x^2 - Dy^2 = \pm 1$ has integer solutions if and only if the fundamental unit has norm -1.

That the rank of $\mathbf{Z}[\sqrt{D}]^\times$ is at most 1 also follows from the fact that the image of the map

$$\begin{array}{ccc} \mathbf{Z}[\sqrt{D}]^\times & \longrightarrow & \mathbf{R}^2 \\ z & \longmapsto & \big(\log|z|, \ \log|z'|\big) \end{array}$$

is discrete in $\mathbf{R}^2$ and contained in the line $t_1 + t_2 = 0$ of $\mathbf{R}^2$. This proof is not really different from the proof we gave of Proposition 3: the proof that the discrete subgroups of $\mathbf{R}$ have rank ≤ 1 relies on Euclid's division.

Remark. Let d be a non–zero rational integer which is not the square of an integer. Then d is not the square of a rational number, and the field $k = \mathbf{Q}(\sqrt{d})$ is a quadratic extension of $\mathbf{Q}$ (which means a $\mathbf{Q}$–vector space of dimension 2). An element $\alpha \in k$ is an *algebraic integer* if and only if it satisfies the following equivalent conditions:
(i) α is root of a monic polynomial with coefficients in $\mathbf{Z}$.

(*ii*) The irreductible (monic) polynomial of α over $\mathbf{Q}$ has coefficients in $\mathbf{Z}$.
(*iii*) The irreducible polynomial of α over $\mathbf{Z}$ is monic.
(*iv*) The ring $\mathbf{Z}[\alpha]$ is a finitely generated $\mathbf{Z}$–module.
(*v*) The ring $\mathbf{Z}[\alpha]$ is contained in a subring of k which is a finitely generated $\mathbf{Z}$–module.

The set $\mathbf{Z}_k$ of algebraic integers of k is the following ring:

$$\mathbf{Z}_k = \begin{cases} \mathbf{Z} + \mathbf{Z}\sqrt{d} & \text{if } d \equiv 2 \text{ or } 3 \pmod 4 \\ \mathbf{Z} + \mathbf{Z}\dfrac{1+\sqrt{d}}{2} & \text{if } d \equiv 1 \pmod 4. \end{cases}$$

Hence $\mathbf{Z}_k = \mathbf{Z} + \mathbf{Z}\alpha$, where α is any of the two roots of $X^2 - d$ if $d \equiv 2$ or $3 \pmod 4$, and any of the two roots of the polynomial

$$X^2 - X - (d-1)/4 = \frac{1}{4}(2X-1)^2 - d$$

if $d \equiv 1 \pmod 4$.

The *discriminant* D_k of k is the discriminant of the ring of integers of k:

$$D_k = \begin{cases} \det \begin{vmatrix} 2 & 0 \\ 0 & 2d \end{vmatrix} = 4d & \text{if } d \equiv 2 \text{ or } 3 \pmod 4 \\ \\ \det \begin{vmatrix} 2 & 1 \\ 1 & (1+d)/2 \end{vmatrix} = d & \text{if } d \equiv 1 \pmod 4. \end{cases}$$

Hence the discriminant is always congruent to 0 or 1 modulo 4 and the quadratic field is $k = \mathbf{Q}(\sqrt{D_k})$.

The group of units [3] of k is by definition the group of units $\mathbf{Z}_K^\times$ of the ring $\mathbf{Z}_k$. For $d < 0$, it is easy to check that the group of units in k is the following finite group of roots of unity in k:

- $\{1, i, -1, -i\}$ if k has discriminant -4, which means $k = \mathbf{Q}(i)$
- $\{1, \varrho, \varrho^2, -1, -\varrho, -\varrho^2\}$ if k has discriminant -3, where ϱ is a root of $X^2 + X + 1$. The quadratic field with discriminant -3 is $k = \mathbf{Q}(\sqrt{\varrho}) = \mathbf{Q}(\sqrt{-3})$ and ϱ is a primitive cube root of unity.

[3] This is an abuse of language of course, since the units of a field are the non–zero elements of the field; the same applies for *ideals of a number field*, which means ideals of the ring of integers of the number field.

- $\{\pm 1\}$ otherwise.

Assume $d > 0$. Then the roots of unity in k are only ± 1 and the group $\mathbf{Z}_k^\times$ of units of $\mathbf{Z}_k$ is a $\mathbf{Z}$-module of rank 1. Hence it is isomorphic to $\{\pm 1\} \times \mathbf{Z}$. For $d \equiv 2$ and for $d \equiv 3 \pmod 4$, the units $\mathbf{Z}_k^\times$ of k are the elements $x + y\sqrt{D_k} \in k$ such that $(x, y) \in \mathbf{Z} \times \mathbf{Z}$ is a solution of Pell's equation $x^2 - D_k y^2 = \pm 1$. For $d \equiv 1 \pmod 4$, the group of units $\mathbf{Z}_k^\times$ of k is the set of elements $x + y\sqrt{D_k} \in k$ such that $(x, y) \in \mathbf{Z} \times \mathbf{Z}$ is a solution of Pell's equation $x^2 - D_k y^2 = \pm 4$.

1.5 Connection with rational approximation

Lemma 4. *Let D be a positive integer which is not a square. Let x and y be positive rational integers. The following conditions are equivalent:*
(i) $x^2 - Dy^2 = 1$.
(ii) $0 < \dfrac{x}{y} - \sqrt{D} < \dfrac{1}{2y^2\sqrt{D}}$.
(iii) $0 < \dfrac{x}{y} - \sqrt{D} < \dfrac{1}{y^2\sqrt{D} + 1}$.

Proof. We have $\dfrac{1}{2y^2\sqrt{D}} < \dfrac{1}{y^2\sqrt{D} + 1}$, hence (ii) implies (iii).
(i) implies $x^2 > Dy^2$, hence $x > y\sqrt{D}$, and consequently

$$0 < \frac{x}{y} - \sqrt{D} = \frac{1}{y(x + y\sqrt{D})} < \frac{1}{2y^2\sqrt{D}}.$$

(iii) implies

$$x < y\sqrt{D} + \frac{1}{y\sqrt{D}} < y\sqrt{D} + \frac{2}{y},$$

and

$$y(x + y\sqrt{D}) < 2y^2\sqrt{D} + 2,$$

hence

$$0 < x^2 - Dy^2 = y\left(\frac{x}{y} - \sqrt{D}\right)(x + y\sqrt{D}) < 2.$$

Since $x^2 - Dy^2$ is an integer, it is equal to 1. □

The next variant will also be useful.

Lemma 5. *Let D be a positive integer which is not a square. Let x and y be positive rational integers. The following conditions are equivalent:*
(i) $x^2 - Dy^2 = -1$.
(ii) $0 < \sqrt{D} - \dfrac{x}{y} < \dfrac{1}{2y^2\sqrt{D} - 1}$.
(iii) $0 < \sqrt{D} - \dfrac{x}{y} < \dfrac{1}{y^2\sqrt{D}}$.

Proof. We have $\dfrac{1}{2y^2\sqrt{D} - 1} < \dfrac{1}{y^2\sqrt{D}}$, hence (ii) implies (iii).
The condition (i) implies $y\sqrt{D} > x$. We use the trivial estimate

$$2\sqrt{D} > 1 + 1/y^2$$

and write

$$x^2 = Dy^2 - 1 > Dy^2 - 2\sqrt{D} + 1/y^2 = (y\sqrt{D} - 1/y)^2,$$

hence $xy > y^2\sqrt{D} - 1$. From (i) one deduces

$$\begin{aligned} 1 = Dy^2 - x^2 = (y\sqrt{D} - x)(y\sqrt{D} + x) \\ > \left(\sqrt{D} - \frac{x}{y}\right)(y^2\sqrt{D} + xy) \\ > \left(\sqrt{D} - \frac{x}{y}\right)(2y^2\sqrt{D} - 1). \end{aligned}$$

(iii) implies $x < y\sqrt{D}$ and

$$y(y\sqrt{D} + x) < 2y^2\sqrt{D},$$

hence

$$0 < Dy^2 - x^2 = y\left(\sqrt{D} - \frac{x}{y}\right)(y\sqrt{D} + x) < 2.$$

Since $Dy^2 - x^2$ is an integer, it is 1. □

From these two lemmas one deduces:

Corollary 6. *Let D be a positive integer which is not a square. Let x and y be positive rational integers. The following conditions are equivalent:*
(i) $x^2 - Dy^2 = \pm 1$.
(ii) $\left|\sqrt{D} - \dfrac{x}{y}\right| < \dfrac{1}{2y^2\sqrt{D} - 1}$.
(iii) $\left|\sqrt{D} - \dfrac{x}{y}\right| < \dfrac{1}{y^2\sqrt{D} + 1}$.

Proof. If $y > 1$ or $D > 3$ we have $2y^2\sqrt{D} - 1 > y^2\sqrt{D} + 1$, which means that (ii) implies trivially (iii), and we may apply Lemmas 4 and 5.

If $D = 2$ and $y = 1$, then each of the conditions (i), (ii) and (iii) is satisfied if and only if $x = 1$. This follows from

$$2 - \sqrt{2} > \frac{1}{2\sqrt{2} - 1} > \frac{1}{\sqrt{2} + 1} > \sqrt{2} - 1.$$

If $D = 3$ and $y = 1$, then each of the conditions (i), (ii) and (iii) is satisfied if and only if $x = 2$. This follows from

$$3 - \sqrt{3} > \sqrt{3} - 1 > \frac{1}{2\sqrt{3} - 1} > \frac{1}{\sqrt{3} + 1} > 2 - \sqrt{3}.$$

□

It is instructive to compare with Liouville's inequality.

Lemma 7. *Let D be a positive integer which is not a square. Let x and y be positive rational integers. Then*

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

Proof. If $x/y < \sqrt{D}$, then $x \leq y\sqrt{D}$ and from

$$1 \leq Dy^2 - x^2 = (y\sqrt{D} + x)(y\sqrt{D} - x) \leq 2y\sqrt{D}(y\sqrt{D} - x),$$

one deduces

$$\sqrt{D} - \frac{x}{y} > \frac{1}{2y^2\sqrt{D}}.$$

We claim that if $x/y > \sqrt{D}$, then

$$\frac{x}{y} - \sqrt{D} > \frac{1}{2y^2\sqrt{D} + 1}.$$

Indeed, this estimate is true if $x - y\sqrt{D} \geq 1/y$, so we may assume $x - y\sqrt{D} < 1/y$. Our claim then follows from

$$1 \leq x^2 - Dy^2 = (x + y\sqrt{D})(x - y\sqrt{D}) \leq (2y\sqrt{D} + 1/y)(x - y\sqrt{D}).$$

□

This shows that a rational approximation x/y to $\sqrt{D}$, which is only slightly weaker than the limit given by Liouville's inequality, will produce a solution to Pell's equation $x^2 - Dy^2 = \pm 1$. The distance $|\sqrt{D} - x/y|$ cannot be smaller than $1/(2y^2\sqrt{D} + 1)$, but it can be as small as $1/(2y^2\sqrt{D} - 1)$, and for that it suffices that it is less than $1/(y^2\sqrt{D} + 1)$

2 Continued fractions

We first consider generalized continued fractions of the form

$$a_0 + \cfrac{b_1}{a_1 + \cfrac{b_2}{a_2 + \cfrac{b_3}{\ddots}}},$$

which we denote by[4]

$$a_0 + \frac{b_1|}{|a_1} + \frac{b_2|}{|a_2} + \frac{b_3|}{\ddots}.$$

Next we restrict to the special case where $b_1 = b_2 = \cdots = 1$, which yields the simple continued fractions

$$a_0 + \frac{1\,|}{|a_1} + \frac{1\,|}{|a_2} + \cdots = [a_0, a_1, a_2, \dots].$$

2.1 Generalized continued fractions

To start with, $a_0, \dots, a_n, \dots$ and $b_1, \dots, b_n, \dots$ will be independent variables. Later, we shall specialize to positive integers (apart from a_0 which may be negative).

Consider the three rational fractions

$$a_0, \qquad a_0 + \frac{b_1}{a_1} \quad \text{and} \quad a_0 + \cfrac{b_1}{a_1 + \cfrac{b_2}{a_2}}.$$

We write them as

$$\frac{A_0}{B_0}, \qquad \frac{A_1}{B_1} \quad \text{and} \quad \frac{A_2}{B_2}$$

with

$$\begin{array}{lll} A_0 = a_0, & A_1 = a_0a_1 + b_1, & A_2 = a_0a_1a_2 + a_0b_2 + a_2b_1, \\ B_0 = 1, & B_1 = a_1, & B_2 = a_1a_2 + b_2. \end{array}$$

[4] Another notation for $a_0 + \frac{b_1|}{|a_1} + \frac{b_2|}{|a_2} + \cdots + \frac{b_n|}{a_n}$ introduced by Th. Muir and used by Perron in [23] Chap. 1 is

$$K\begin{pmatrix} b_1, \dots, b_n \\ a_0, a_1, \dots, a_n \end{pmatrix}$$

Observe that

$$A_2 = a_2A_1 + b_2A_0, \quad B_2 = a_2B_1 + b_2B_0.$$

Write these relations as

$$\begin{pmatrix} A_2 \\ B_2 \end{pmatrix} = \begin{pmatrix} A_1 & A_0 \\ B_1 & B_0 \end{pmatrix} \begin{pmatrix} a_2 \\ b_2 \end{pmatrix}.$$

In order to iterate the process, it is convenient to work with 2×2 matrices and to write

$$\begin{pmatrix} A_2 & A_1 \\ B_2 & B_1 \end{pmatrix} = \begin{pmatrix} A_1 & A_0 \\ B_1 & B_0 \end{pmatrix} \begin{pmatrix} a_2 & 1 \\ b_2 & 0 \end{pmatrix}.$$

Define inductively two sequences of polynomials with positive rational coefficients A_n and B_n for $n \geq 3$ by

$$\begin{pmatrix} A_n & A_{n-1} \\ B_n & B_{n-1} \end{pmatrix} = \begin{pmatrix} A_{n-1} & A_{n-2} \\ B_{n-1} & B_{n-2} \end{pmatrix} \begin{pmatrix} a_n & 1 \\ b_n & 0 \end{pmatrix}. \tag{8}$$

This means

$$A_n = a_nA_{n-1} + b_nA_{n-2}, \qquad B_n = a_nB_{n-1} + b_nB_{n-2}.$$

This recurrence relation holds for $n \geq 2$. It will also hold for $n = 1$ if we set $A_{-1} = 1$ and $B_{-1} = 0$:

$$\begin{pmatrix} A_1 & A_0 \\ B_1 & B_0 \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ b_1 & 0 \end{pmatrix}$$

and it will hold also for $n = 0$ if we set $b_0 = 1$, $A_{-2} = 0$ and $B_{-2} = 1$:

$$\begin{pmatrix} A_0 & A_{-1} \\ B_0 & B_{-1} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a_0 & 1 \\ b_0 & 0 \end{pmatrix}.$$

Obviously, an equivalent definition is

$$\begin{pmatrix} A_n & A_{n-1} \\ B_n & B_{n-1} \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ b_0 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ b_1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{n-1} & 1 \\ b_{n-1} & 0 \end{pmatrix} \begin{pmatrix} a_n & 1 \\ b_n & 0 \end{pmatrix}. \tag{9}$$

These relations (9) hold for $n \geq -1$, with the empty product (for $n = -1$) being the identity matrix, as always.

Hence $A_n \in \mathbf{Z}[a_0, \ldots, a_n, b_1, \ldots, b_n]$ is a polynomial in $2n + 1$ variables, while $B_n \in \mathbf{Z}[a_1 \ldots, a_n, b_2, \ldots, b_n]$ is a polynomial in $2n - 1$ variables.

Exercise 1. *Check, for $n \geq -1$,*

$$B_n(a_1, \ldots, a_n, b_2, \ldots, b_n) = A_{n-1}(a_1, \ldots, a_n, b_2, \ldots, b_n).$$

Lemma 10. *For $n \geq 0$,*

$$a_0 + \frac{b_1|}{|a_1} + \cdots + \frac{b_n|}{|a_n} = \frac{A_n}{B_n}.$$

Proof. By induction. We have checked the result for $n = 0$, $n = 1$ and $n = 2$. Assume the formula holds with $n - 1$ where $n \geq 3$. We write

$$a_0 + \frac{b_1|}{|a_1} + \cdots + \frac{b_{n-1}|}{|a_{n-1}} + \frac{b_n|}{|a_n} = a_0 + \frac{b_1|}{|a_1} + \cdots + \frac{b_{n-1}|}{|x}$$

with

$$x = a_{n-1} + \frac{b_n}{a_n}.$$

We have, by induction hypothesis and by the definition (8),

$$a_0 + \frac{b_1|}{|a_1} + \cdots + \frac{b_{n-1}|}{|a_{n-1}} = \frac{A_{n-1}}{B_{n-1}} = \frac{a_{n-1}A_{n-2} + b_{n-1}A_{n-3}}{a_{n-1}B_{n-2} + b_{n-1}B_{n-3}}.$$

Since A_{n-2}, A_{n-3}, B_{n-2} and B_{n-3} do not depend on the variable a_{n-1}, we deduce

$$a_0 + \frac{b_1|}{|a_1} + \cdots + \frac{b_{n-1}|}{|x} = \frac{xA_{n-2} + b_{n-1}A_{n-3}}{xB_{n-2} + b_{n-1}B_{n-3}}.$$

The product of the numerator by a_n is

$$\begin{aligned}(a_n a_{n-1} + b_n)A_{n-2} + a_n b_{n-1}A_{n-3} &= a_n(a_{n-1}A_{n-2} + b_{n-1}A_{n-3}) + b_n A_{n-2} \\ &= a_n A_{n-1} + b_n A_{n-2} = A_n\end{aligned}$$

and similarly, the product of the denominator by a_n is

$$\begin{aligned}(a_n a_{n-1} + b_n)B_{n-2} + a_n b_{n-1}B_{n-3} &= a_n(a_{n-1}B_{n-2} + b_{n-1}B_{n-3}) + b_n B_{n-2} \\ &= a_n B_{n-1} + b_n B_{n-2} = B_n.\end{aligned}$$

□

From (9), taking the determinant, we deduce, for $n \geq -1$,

$$A_n B_{n-1} - A_{n-1}B_n = (-1)^{n+1} b_0 \cdots b_n. \tag{11}$$

which can be written, for $n \geq 1$,

$$\frac{A_n}{B_n} - \frac{A_{n-1}}{B_{n-1}} = \frac{(-1)^{n+1} b_0 \cdots b_n}{B_{n-1} B_n}. \tag{12}$$

Adding the telescoping sum, we get, for $n \geq 0$,

$$\frac{A_n}{B_n} = A_0 + \sum_{k=1}^{n} \frac{(-1)^{k+1} b_0 \cdots b_k}{B_{k-1} B_k}. \tag{13}$$

We now substitute for $a_0, a_1, \ldots$ and $b_1, b_2, \ldots$ rational integers, all of which are ≥ 1, apart from a_0 which may be ≤ 0. We denote by p_n (resp. q_n) the value of A_n (resp. B_n) for these special values. Hence p_n and q_n are rational integers, with $q_n > 0$ for $n \geq 0$. A consequence of Lemma 10 is

$$\frac{p_n}{q_n} = a_0 + \frac{b_1|}{|a_1} + \cdots + \frac{b_n|}{|a_n} \quad \text{for} \quad n \geq 0.$$

We deduce from (8),

$$p_n = a_n p_{n-1} + b_n p_{n-2}, \qquad q_n = a_n q_{n-1} + b_n q_{n-2} \quad \text{for} \quad n \geq 0,$$

and from (11),

$$p_n q_{n-1} - p_{n-1} q_n = (-1)^{n+1} b_0 \cdots b_n \quad \text{for} \quad n \geq -1,$$

which can be written, for $n \geq 1$,

$$\frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}} = \frac{(-1)^{n+1} b_0 \cdots b_n}{q_{n-1} q_n}. \tag{14}$$

Adding the telescoping sum (or using (13)), we get the alternating sum

$$\frac{p_n}{q_n} = a_0 + \sum_{k=1}^{n} \frac{(-1)^{k+1} b_0 \cdots b_k}{q_{k-1} q_k}. \tag{15}$$

Recall that for real numbers a, b, c, d, with b and d positive, we have

$$\frac{a}{b} < \frac{c}{d} \implies \frac{a}{b} < \frac{a+c}{b+d} < \frac{c}{d}. \tag{16}$$

Since a_n and b_n are positive for $n \geq 0$, we deduce that for $n \geq 2$, the rational number

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

lies between p_{n-1}/q_{n-1} and p_{n-2}/q_{n-2}. Therefore we have

$$\frac{p_2}{q_2} < \frac{p_4}{q_4} < \cdots < \frac{p_{2n}}{q_{2n}} < \cdots < \frac{p_{2m+1}}{q_{2m+1}} < \cdots < \frac{p_3}{q_3} < \frac{p_1}{q_1}. \tag{17}$$

From (14), we deduce, for $n \geq 3$, $q_{n-1} > q_{n-2}$, hence $q_n > (a_n + b_n)q_{n-2}$.

The previous discussion was valid without any restriction, now we assume $a_n \geq b_n$ for all sufficiently large n, say $n \geq n_0$. Then for $n > n_0$, using $q_n > 2b_n q_{n-2}$, we get

$$\left|\frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}}\right| = \frac{b_0 \cdots b_n}{q_{n-1}q_n} < \frac{b_n \cdots b_0}{2^{n-n_0} b_n b_{n-1} \cdots b_{n_0+1} q_{n_0} q_{n_0-1}} = \frac{b_{n_0} \cdots b_0}{2^{n-n_0} q_{n_0} q_{n_0-1}}$$

and the right hand side tends to 0 as n tends to infinity. Hence the sequence $(p_n/q_n)_{n\geq 0}$ has a limit, which we denote by

$$x = a_0 + \frac{b_1|}{|a_1} + \cdots + \frac{b_{n-1}|}{|a_{n-1}} + \frac{b_n|}{|a_n} + \cdots$$

From (15), it follows that x is also given by an alternating series

$$x = a_0 + \sum_{k=1}^{\infty} \frac{(-1)^{k+1} b_0 \cdots b_k}{q_{k-1}q_k}.$$

We now prove that x is irrational. Define, for $n \geq 0$,

$$x_n = a_n + \frac{b_{n+1}|}{|a_{n+1}} + \cdots$$

so that $x = x_0$ and, for all $n \geq 0$,

$$x_n = a_n + \frac{b_{n+1}}{x_{n+1}}, \qquad x_{n+1} = \frac{b_{n+1}}{x_n - a_n}$$

and $a_n < x_n < a_n + 1$. Hence for $n \geq 0$, x_n is rational if and only if x_{n+1} is rational, and therefore, if x is rational, then all x_n for $n \geq 0$ are also rational. Assume x is rational. Consider the rational numbers x_n with $n \geq n_0$ and select a value of n for which the denominator v of x_n is minimal, say $x_n = u/v$. From

$$x_{n+1} = \frac{b_{n+1}}{x_n - a_n} = \frac{b_{n+1}v}{u - a_n v} \quad \text{with} \quad 0 < u - a_n v < v,$$

it follows that x_{n+1} has a denominator strictly less than v, which is a contradiction. Hence x is irrational.

Conversely, given an irrational number x and a sequence $b_1, b_2, \dots$ of positive integers, there is a unique integer a_0 and a unique sequence $a_1, \dots, a_n, \dots$ of positive integers satisfying $a_n \geq b_n$ for all $n \geq 1$, such that

$$x = a_0 + \frac{b_1|}{|a_1} + \cdots + \frac{b_{n-1}|}{|a_{n-1}} + \frac{b_n|}{|a_n} + \cdots$$

Indeed, the unique solution is given inductively as follows: $a_0 = \lfloor x \rfloor$, $x_1 = b_1/\{x\}$, and once $a_0, \dots, a_{n-1}$ and $x_1, \dots, x_n$ are known, then a_n and x_{n+1} are given by

$$a_n = \lfloor x_n \rfloor, \quad x_{n+1} = b_{n+1}/\{x_n\},$$

so that for $n \geq 1$ we have $0 < x_n - a_n < 1$ and

$$x = a_0 + \frac{b_1|}{|a_1} + \cdots + \frac{b_{n-1}|}{|a_{n-1}} + \frac{b_n|}{|x_n}.$$

Here is what we have proved.

Proposition 18. *Given a rational integer a_0 and two sequences $a_0, a_1, \dots$ and $b_1, b_2, \dots$ of positive rational integers with $a_n \geq b_n$ for all sufficiently large n, the infinite continued fraction*

$$a_0 + \frac{b_1|}{|a_1} + \cdots + \frac{b_{n-1}|}{|a_{n-1}} + \frac{b_n|}{|a_n} + \cdots$$

exists and is an irrational number.
Conversely, given an irrational number x and a sequence $b_1, b_2, \dots$ of positive integers, there is a unique $a_0 \in \mathbf{Z}$ and a unique sequence $a_1, \dots, a_n, \dots$ of positive integers satisfying $a_n \geq b_n$ for all $n \geq 1$ such that

$$x = a_0 + \frac{b_1|}{|a_1} + \cdots + \frac{b_{n-1}|}{|a_{n-1}} + \frac{b_n|}{|a_n} + \cdots$$

These results are useful for proving the irrationality of π and e^r when r is a non–zero rational number, following the proof by Lambert. See for instance Chapter 7 (Lambert's Irrationality Proofs) of David Angell's course on Irrationality and Transcendence([5]) at the University of New South Wales:

`http://www.maths.unsw.edu.au/~angell/5535/`

The following example is related with Lambert's proof [16]:

$$\tanh z = \frac{z|}{|1} + \frac{z^2|}{|\,3} + \frac{z^2|}{|\,5} + \cdots + \frac{z^2\ \ |}{|2n+1} + \cdots$$

[5] I found this reference from the website of John Cosgrave
`http://staff.spd.dcu.ie/johnbcos/transcendental_numbers.htm`.

Here, z is a complex number and the right hand side is a complex valued function. Here are other examples (see Sloane's Encyclopaedia of Integer Sequences([6]))

$$\frac{1}{\sqrt{e}-1} = 1 + \frac{2|}{|3} + \frac{4|}{|5} + \frac{6|}{|7} + \frac{8|}{|9} + \cdots = 1.541\,494\,082\ldots \qquad \text{(A113011)}$$

$$\frac{1}{e-1} = \frac{1|}{|1} + \frac{2|}{|2} + \frac{3|}{|3} + \frac{4|}{|4} + \cdots = 0.581\,976\,706\ldots \qquad \text{(A073333)}$$

Remark. *A variant of the algorithm of simple continued fractions is the following. Given two sequences $(a_n)_{n\geq 0}$ and $(b_n)_{n\geq 0}$ of elements in a field K and an element x in K, one defines a sequence (possibly finite) $(x_n)_{n\geq 1}$ of elements in K as follows. If $x = a_0$, the sequence is empty. Otherwise x_1 is defined by $x = a_0 + (b_1/x_1)$. Inductively, once $x_1, \ldots, x_n$ are defined, there are two cases:*

- *If $x_n = a_n$, the algorithm stops.*
- *Otherwise, x_{n+1} is defined by*

$$x_{n+1} = \frac{b_{n+1}}{x_n - a_n}, \quad \text{so that} \quad x_n = a_n + \frac{b_{n+1}}{x_{n+1}}.$$

If the algorithm does not stop, then for any $n \geq 1$, one has

$$x = a_0 + \frac{b_1|}{|a_1} + \cdots + \frac{b_{n-1}|}{|a_{n-1}} + \frac{b_n|}{|x_n}.$$

In the special case where $a_0 = a_1 = \cdots = b_1 = b_2 = \cdots = 1$, the set of x such that the algorithm stops after finitely many steps is the set $(F_{n+1}/F_n)_{n\geq 1}$ of quotients of consecutive Fibonacci numbers. In this special case, the limit of

$$a_0 + \frac{b_1|}{|a_1} + \cdots + \frac{b_{n-1}|}{|a_{n-1}} + \frac{b_n|}{|a_n}$$

is the Golden ratio, which is independent of x, of course!

2.2 Simple continued fractions

We restrict now the discussion of § 2.1 to the case where $b_1 = b_2 = \cdots = b_n = \cdots = 1$. We keep the notations A_n and B_n which are now polynomials in $\mathbf{Z}[a_0, a_1, \ldots, a_n]$ and $\mathbf{Z}[a_1, \ldots, a_n]$ respectively, and when we specialize to

[6] `http://www.research.att.com/~njas/sequences/`

integers $a_0, a_1, \ldots, a_n \ldots$ with $a_n \geq 1$ for $n \geq 1$ we use the notations p_n and q_n for the values of A_n and B_n.

The recurrence relations (8) are now, for $n \geq 0$,

$$\begin{pmatrix} A_n & A_{n-1} \\ B_n & B_{n-1} \end{pmatrix} = \begin{pmatrix} A_{n-1} & A_{n-2} \\ B_{n-1} & B_{n-2} \end{pmatrix} \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix}, \tag{19}$$

while (9) becomes, for $n \geq -1$,

$$\begin{pmatrix} A_n & A_{n-1} \\ B_n & B_{n-1} \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{n-1} & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix}. \tag{20}$$

From Lemma 10 one deduces, for $n \geq 0$,

$$[a_0, \ldots, a_n] = \frac{A_n}{B_n}.$$

Taking the determinant in (20), we deduce the following special case of (11)

$$A_n B_{n-1} - A_{n-1} B_n = (-1)^{n+1}.$$

The specialization of these relations to integral values of $a_0, a_1, a_2 \ldots$ yields

$$\begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix} = \begin{pmatrix} p_{n-1} & p_{n-2} \\ q_{n-1} & q_{n-2} \end{pmatrix} \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix} \qquad \text{for } n \geq 0, \tag{21}$$

$$\begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{n-1} & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix} \qquad \text{for } n \geq -1, \tag{22}$$

$$[a_0, \ldots, a_n] = \frac{p_n}{q_n} \qquad \text{for } n \geq 0$$

and

$$p_n q_{n-1} - p_{n-1} q_n = (-1)^{n+1} \qquad \text{for } n \geq -1. \tag{23}$$

From (23), it follows that for $n \geq 0$, the fraction p_n/q_n is in lowest terms: $\gcd(p_n, q_n) = 1$.

Transposing (22) yields, for $n \geq -1$,

$$\begin{pmatrix} p_n & q_n \\ p_{n-1} & q_{n-1} \end{pmatrix} = \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_{n-1} & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix}$$

from which we deduce, for $n \geq 1$,

$$[a_n, \ldots, a_0] = \frac{p_n}{p_{n-1}} \quad \text{and} \quad [a_n, \ldots, a_1] = \frac{q_n}{q_{n-1}}$$

Lemma 24. *For $n \geq 0$,*

$$p_n q_{n-2} - p_{n-2} q_n = (-1)^n a_n.$$

Proof. We multiply both sides of (21) on the left by the inverse of the matrix

$$\begin{pmatrix} p_{n-1} & p_{n-2} \\ q_{n-1} & q_{n-2} \end{pmatrix} \quad \text{which is} \quad (-1)^n \begin{pmatrix} q_{n-2} & -p_{n-2} \\ -q_{n-1} & p_{n-1} \end{pmatrix}.$$

We get

$$(-1)^n \begin{pmatrix} p_n q_{n-2} - p_{n-2} q_n & p_{n-1} q_{n-2} - p_{n-2} q_{n-1} \\ -p_n q_{n-1} + p_{n-1} q_n & 0 \end{pmatrix} = \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix}$$

□

2.2.1 Finite simple continued fraction of a rational number

Let u_0 and u_1 be two integers with u_1 positive. The first step in Euclid's algorithm to find the gcd of u_0 and u_1 consists in dividing u_0 by u_1:

$$u_0 = a_0 u_1 + u_2$$

with $a_0 \in \mathbf{Z}$ and $0 \leq u_2 < u_1$. This means

$$\frac{u_0}{u_1} = a_0 + \frac{u_2}{u_1},$$

which amonts to dividing the rational number $x_0 = u_0/u_1$ by 1 with quotient a_0 and remainder $u_2/u_1 < 1$. This algorithms continues with

$$u_m = a_m u_{m+1} + u_{m+2},$$

where a_m is the integral part of $x_m = u_m/u_{m+1}$ and $0 \leq u_{m+2} < u_{m+1}$, until some $u_{\ell+2}$ is 0, in which case the algorithms stops with

$$u_\ell = a_\ell u_{\ell+1}.$$

Since the gcd of u_m and u_{m+1} is the same as the gcd of u_{m+1} and u_{m+2}, it follows that the gcd of u_0 and u_1 is $u_{\ell+1}$. This is how one gets the regular continued fraction expansion $x_0 = [a_0, a_1, \ldots, a_\ell]$, where $\ell = 0$ in case x_0 is a rational integer, while $a_\ell \geq 2$ if x_0 is a rational number which is not an integer.

Exercise 2. *Compare with the geometrical construction of the continued fraction given in the beamer presentation.*
Give a variant of this geometrical construction where rectangles are replaced by segments.

Proposition 25. *Any finite regular continued fraction*

$$[a_0, a_1, \dots, a_n],$$

where $a_0, a_1, \dots, a_n$ are rational numbers with $a_i \geq 2$ for $1 \leq i \leq n$ and $n \geq 0$, represents a rational number. Conversely, any rational number x has two representations as a continued fraction, the first one, given by Euclid's algorithm, is

$$x = [a_0, a_1, \dots, a_n]$$

and the second one is

$$x = [a_0, a_1, \dots, a_{n-1}, a_n - 1, 1].$$

If $x \in \mathbf{Z}$, then $n = 0$ and the two simple continued fractions representations of x are $[x]$ and $[x-1, 1]$, while if x is not an integer, then $n \geq 1$ and $a_n \geq 2$. For instance the two continued fractions of 1 are $[1]$ and $[0, 1]$, they both end with 1. The two continued fractions of 0 are $[0]$ and $[-1, 1]$, the first of which is the unique continued fraction which ends with 0.

We shall use later (in the proof of Lemma 30 in § 3.2) the fact that any rational number has one simple continued fraction expansion with an odd number of terms and one with an even number of terms.

2.2.2 Infinite simple continued fraction of an irrational number

Given a rational integer a_0 and an infinite sequence of positive integers $a_1, a_2, \dots$, the continued fraction

$$[a_0, a_1, \dots, a_n, \dots]$$

represents an irrational number. Conversely, given an irrational number x, there is a unique representation of x as an infinite simple continued fraction

$$x = [a_0, a_1, \dots, a_n, \dots]$$

Definitions The numbers a_n are the *partial quotients*, the rational numbers

$$\frac{p_n}{q_n} = [a_0, a_1, \dots, a_n]$$

are the *convergents* (in French *réduites*), and the numbers

$$x_n = [a_n, a_{n+1}, \ldots]$$

are the *complete quotients.*

From these definitions we deduce, for $n \geq 0$,

$$x = [a_0, a_1, \ldots, a_n, x_{n+1}] = \frac{x_{n+1}p_n + p_{n-1}}{x_{n+1}q_n + q_{n-1}}. \tag{26}$$

Lemma 27. *For $n \geq 0$,*

$$q_n x - p_n = \frac{(-1)^n}{x_{n+1}q_n + q_{n-1}}.$$

Proof. From (26) one deduces

$$x - \frac{p_n}{q_n} = \frac{x_{n+1}p_n + p_{n-1}}{x_{n+1}q_n + q_{n-1}} - \frac{p_n}{q_n} = \frac{(-1)^n}{(x_{n+1}q_n + q_{n-1})q_n}.$$

□

Corollary 28. *For $n \geq 0$,*

$$\frac{1}{q_{n+1} + q_n} < |q_n x - p_n| < \frac{1}{q_{n+1}}.$$

Proof. Since a_{n+1} is the integral part of x_{n+1}, we have

$$a_{n+1} < x_{n+1} < a_{n+1} + 1.$$

Using the recurrence relation $q_{n+1} = a_{n+1}q_n + q_{n-1}$, we deduce

$$q_{n+1} < x_{n+1}q_n + q_{n-1} < a_{n+1}q_n + q_{n-1} + q_n = q_{n+1} + q_n.$$

□

In particular, since $x_{n+1} > a_{n+1}$ and $q_{n-1} > 0$, one deduces from Lemma 27

$$\frac{1}{(a_{n+1} + 2)q_n^2} < \left| x - \frac{p_n}{q_n} \right| < \frac{1}{a_{n+1}q_n^2}. \tag{29}$$

Therefore any convergent p/q of x satisfies $|x - p/q| < 1/q^2$ (compare with (i) $\Rightarrow$ (v) in Proposition 61). Moreover, if a_{n+1} is large, then the approximation p_n/q_n is sharp. Hence, large partial quotients yield good rational approximations by truncating the continued fraction expansion just before the given partial quotient.

3 Continued fractions and Pell's Equation

3.1 The main lemma

The theory which follows is well–known (a classical reference is the book [23] by O. Perron), but the point of view which we develop here is slightly different from most classical texts on the subject. We follow [3, 4, 32]. An important role in our presentation of the subject is the following result (Lemma 4.1 in [26]).

Lemma 30. *Let $\epsilon = \pm 1$ and let a, b, c, d be rational integers satisfying*

$$ad - bc = \epsilon$$

and $d \geq 1$. Then there is a unique finite sequence of rational integers $a_0, \ldots, a_s$ with $s \geq 1$ and $a_1, \ldots, a_{s-1}$ positive, such that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_s & 1 \\ 1 & 0 \end{pmatrix} \tag{31}$$

These integers are also characterized by

$$\frac{b}{d} = [a_0, a_1, \ldots, a_{s-1}], \quad \frac{c}{d} = [a_s, \ldots, a_1], \quad (-1)^{s+1} = \epsilon. \tag{32}$$

For instance, when $d = 1$, for b and c rational integers,

$$\begin{pmatrix} bc+1 & b \\ c & 1 \end{pmatrix} = \begin{pmatrix} b & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} c & 1 \\ 1 & 0 \end{pmatrix}$$

and

$$\begin{pmatrix} bc-1 & b \\ c & 1 \end{pmatrix} = \begin{pmatrix} b-1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} c-1 & 1 \\ 1 & 0 \end{pmatrix}.$$

Proof. We start with unicity. If $a_0, \ldots, a_s$ satisfy the conclusion of Lemma 30, then by using (31), we find $b/d = [a_0, a_1, \ldots, a_{s-1}]$. Taking the transpose, we also find $c/d = [a_s, \ldots, a_1]$. Next, taking the determinant, we obtain $(-1)^{s+1} = \epsilon$. The last equality fixes the parity of s, and each of the rational numbers b/d, c/d has a unique continued fraction expansion whose length has a given parity (cf. Proposition 25). This proves the unicity of the factorisation when it exists.

For the existence, we consider the simple continued fraction expansion of c/d with length of parity given by the last condition in (32), say $c/d =$

$[a_s, \dots, a_1]$. Let a_0 be a rational integer such that the distance between b/d and $[a_0, a_1, \dots, a_{s-1}]$ is $\le 1/2$. Define a', b', c', d' by

$$\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_s & 1 \\ 1 & 0 \end{pmatrix}.$$

We have

$$d' > 0, \quad a'd' - b'c' = \epsilon, \quad \frac{c'}{d'} = [a_s, \dots, a_1] = \frac{c}{d}$$

and

$$\frac{b'}{d'} = [a_0, a_1, \dots, a_{s-1}], \quad \left| \frac{b'}{d'} - \frac{b}{d} \right| \le \frac{1}{2}.$$

From $\gcd(c, d) = \gcd(c', d') = 1$, $c/d = c'/d'$ and $d > 0$, $d' > 0$ we deduce $c' = c$, $d' = d$. From the equality between the determinants we deduce $a' = a + kc$, $b' = b + kd$ for some $k \in \mathbf{Z}$, and from

$$\frac{b'}{d'} - \frac{b}{d} = k$$

we conclude $k = 0$, $(a', b', c', d') = (a, b, c, d)$. Hence (31) follows.

□

Corollary 33. *Assume the hypotheses of Lemma 30 are satisfied.*
(a) If $c > d$, then $a_s \ge 1$ and

$$\frac{a}{c} = [a_0, a_1, \dots, a_s].$$

(b) If $b > d$, then $a_0 \ge 1$ and

$$\frac{a}{b} = [a_s, \dots, a_1, a_0].$$

The following examples show that the hypotheses of the corollary are not superfluous:

$$\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} b & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

$$\begin{pmatrix} b-1 & b \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} b-1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

and

$$\begin{pmatrix} c-1 & 1 \\ c & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} c-1 & 1 \\ 1 & 0 \end{pmatrix}.$$

Proof of Corollary 33. Any rational number $u/v > 1$ has two continued fractions. One of them starts with 0 only if $u/v = 1$ and the continued fraction is $[0, 1]$. Hence the assumption $c > d$ implies $a_s > 0$. This proves part (a), and part (b) follows by transposition (or repeating the proof). □

Another consequence of Lemma 30 is the following classical result (Satz 13 p. 47 of [23]).

Corollary 34. *Let a, b, c, d be rational integers with $ad - bc = \pm 1$ and $c > d > 0$. Let x and y be two irrational numbers satisfying $y > 1$ and*

$$x = \frac{ay + b}{cy + d}.$$

Let $x = [a_0, a_1, \ldots]$ be the simple continued fraction expansion of x. Then there exists $s \geq 1$ such that

$$a = p_s, \quad b = p_{s-1}, \quad c = q_s, \quad r = q_{s-1}, \quad y = x_{s+1}.$$

Proof. Using lemma 30, we write

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a'_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a'_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a'_s & 1 \\ 1 & 0 \end{pmatrix}$$

with $a'_1, \ldots, a'_{s-1}$ positive and

$$\frac{b}{d} = [a'_0, a'_1, \ldots, a'_{s-1}], \quad \frac{c}{d} = [a'_s, \ldots, a'_1].$$

From $c > d$ and corollary 33, we deduce $a'_s > 0$ and

$$\frac{a}{c} = [a'_0, a'_1, \ldots, a'_s] = \frac{p'_s}{q'_s}, \quad x = \frac{p'_s y + p'_{s-1}}{q'_s y + q'_{s-1}} = [a'_0, a'_1, \ldots, a'_s, y].$$

Since $y > 1$, it follows that $a'_i = a_i$, $p'_i = q'_i$ for $0 \leq i \leq s$ and $y = x_{s+1}$.

□

Remark.
In [12], § 4, there is a variant of the matrix formula (21) for the simple continued fraction of a real number.

Given integers $a_0, a_1, \ldots$ with $a_i > 0$ for $i \geq 1$ and writing, for $n \geq 0$, as usual, $p_n/q_n = [a_0, a_1, \ldots, a_n]$, one checks, by induction on n, the two formulae

$$\left.\begin{array}{ll} \begin{pmatrix} 1 & a_0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ a_1 & 1 \end{pmatrix} \cdots \begin{pmatrix} 1 & a_n \\ 0 & 1 \end{pmatrix} & = \begin{pmatrix} p_{n-1} & p_n \\ q_{n-1} & q_n \end{pmatrix} \quad \text{if } n \text{ is even} \\ \begin{pmatrix} 1 & a_0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ a_1 & 1 \end{pmatrix} \cdots \begin{pmatrix} 1 & 0 \\ a_n & 1 \end{pmatrix} & = \begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix} \quad \text{if } n \text{ is odd} \end{array}\right\} \tag{35}$$

Define two matrices U (up) and L (low) in $\mathrm{GL}_2(\mathbf{R})$ of determinant $+1$ by

$$U = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad L = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

For p and q in $\mathbf{Z}$, we have

$$U^p = \begin{pmatrix} 1 & p \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad L^q = \begin{pmatrix} 1 & 0 \\ q & 1 \end{pmatrix},$$

so that these formulae (35) are

$$U^{a_0} L^{a_1} \cdots U^{a_n} = \begin{pmatrix} p_{n-1} & p_n \\ q_{n-1} & q_n \end{pmatrix} \quad \text{if } n \text{ is even}$$

and

$$U^{a_0} L^{a_1} \cdots L^{a_n} = \begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix} \quad \text{if } n \text{ is odd.}$$

The connexion with Euclid's algorithm is

$$U^{-p} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a - pc & b - pd \\ c & d \end{pmatrix} \quad \text{and} \quad L^{-q} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c - qa & d - qb \end{pmatrix}.$$

The corresponding variant of Lemma 30 is also given in [12], § 4: *If a, b, c, d are rational integers satisfying $b > a > 0$, $d > c \geq 0$ and $ad - bc = 1$, then there exist rational integers $a_0, \ldots, a_n$ with n even and $a_1, \ldots, a_n$ positive, such that*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & a_0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ a_1 & 1 \end{pmatrix} \cdots \begin{pmatrix} 1 & a_n \\ 0 & 1 \end{pmatrix}$$

These integers are uniquely determined by $b/d = [a_0, \ldots, a_n]$ with n even.

3.2 Simple Continued fraction of $\sqrt{D}$

An infinite sequence $(a_n)_{n\geq 1}$ is *periodic* if there exists a positive integer s such that

$$a_{n+s} = a_n \quad \text{for all } n \geq 1. \tag{36}$$

In this case, the finite sequence $(a_1, \ldots, a_s)$ is called a *period* of the original sequence. For the sake of notation, we write

$$(a_1, a_2, \ldots) = (\overline{a_1, \ldots, a_s}).$$

If s_0 is the smallest positive integer satisfying (36), then the set of s satisfying (36) is the set of positive multiples of s_0. In this case $(a_1, \ldots, a_{s_0})$ is called *the fundamental period* of the original sequence.

Theorem 37. *Let D be a positive integer which is not a square. Write the simple continued fraction of $\sqrt{D}$ as $[a_0, a_1, \ldots]$ with $a_0 = \lfloor \sqrt{D} \rfloor$.*
(a) The sequence $(a_1, a_2, \ldots)$ is periodic.
(b) Let (x, y) be a positive integer solution to Pell's equation $x^2 - Dy^2 = \pm 1$. Then there exists $s \geq 1$ such that $x/y = [a_0, \ldots, a_{s-1}]$ and

$$(a_1, a_2, \ldots, a_{s-1}, 2a_0)$$

is a period of the sequence $(a_1, a_2, \ldots)$. Further, $a_{s-i} = a_i$ for $1 \leq i \leq s-1$. One says that the word $a_1, \ldots, a_{s-1}$ is a palindrome. [7]
(c) Let $(a_1, a_2, \ldots, a_{s-1}, 2a_0)$ be a period of the sequence $(a_1, a_2, \ldots)$. Set $x/y = [a_0, \ldots, a_{s-1}]$. Then $x^2 - Dy^2 = (-1)^s$.
(d) Let s_0 be the length of the fundamental period. Then for $i \geq 0$ not multiple of s_0, we have $a_i \leq a_0$.

If $(a_1, a_2, \ldots, a_{s-1}, 2a_0)$ is a period of the sequence $(a_1, a_2, \ldots)$, then

$$\sqrt{D} = [a_0, \overline{a_1, \ldots, a_{s-1}, 2a_0}] = [a_0, a_1, \ldots, a_{s-1}, a_0 + \sqrt{D}].$$

[7] Note (2016). As kindly pointed out to me by Yoishi Motohashi, the fact that the word $a_1, \ldots, a_{s-1}$ is a palindrom is proved in 'Essai sur la théorie des nombres' by Legendre (1798).
In his first paper published at the age of 17 by Evariste Galois, it is proved that if the expansion of a quadratic irrational α is purely periodic, then the same is true for the conjugate α' of α, and the continued fraction of α' is obtained by reversing the order of the continued fraction of α. Besides, this continued fraction is a palindrom if and only if $\alpha\alpha' = -1$.
É. Galois, *Démonstration d'un théorème sur les fractions continues périodiques.*
Annales de Mathématiques Pures et Appliquées, **19** (1828-1829), p. 294-301.
`http://archive.numdam.org/article/AMPA_1828-1829__19__294_0.pdf`
For more information on these contributions by Galois, see
`https://www.bibnum.education.fr/mathematiques/algebre/demonstration-d-un-theoreme-sur-les-fractions-continues-periodiques`

Consider the fundamental period $(a_1, a_2, \ldots, a_{s_0-1}, a_{s_0})$ of the sequence $(a_1, a_2, \ldots)$. By part (b) of Theorem 37 we have $a_{s_0} = 2a_0$, and by part (d), it follows that s_0 is the smallest index i such that $a_i > a_0$.

From (b) and (c) in Theorem 37, it follows that the fundamental solution (x_1, y_1) to Pell's equation $x^2 - Dy^2 = \pm 1$ is given by $x_1/y_1 = [a_0, \ldots, a_{s_0-1}]$, and that $x_1^2 - Dy_1^2 = (-1)^{s_0}$. Therefore, if s_0 is even, then there is no solution to the Pell's equation $x^2 - Dy^2 = -1$. If s_0 is odd, then (x_1, y_1) is the fundamental solution to Pell's equation $x^2 - Dy^2 = -1$, while the fundamental solution (x_2, y_2) to Pell's equation $x^2 - Dy^2 = 1$ is given by $x_2/y_2 = [a_0, \ldots, a_{2s-1}]$.

It follows also from Theorem 37 that the $(ns_0 - 1)$-th convergent

$$x_n/y_n = [a_0, \ldots, a_{ns_0-1}]$$

satisfies

$$x_n + y_n\sqrt{D} = (x_1 + y_1\sqrt{D})^n. \tag{38}$$

We shall check this relation directly (Lemma 42).

Proof. Start with a positive solution (x, y) to Pell's equation $x^2 - Dy^2 = \pm 1$, which exists according to Proposition 2. Since $Dy \geq x$ and $x > y$, we may use lemma 30 and corollary 33 with

$$a = Dy, \quad b = c = x, \quad d = y$$

and write

$$\begin{pmatrix} Dy & x \\ x & y \end{pmatrix} = \begin{pmatrix} a'_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a'_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a'_s & 1 \\ 1 & 0 \end{pmatrix} \tag{39}$$

with positive integers $a'_0, \ldots, a'_s$ and with $a'_0 = \lfloor \sqrt{D} \rfloor$. Then the continued fraction expansion of Dy/x is $[a'_0, \ldots, a'_s]$ and the continued fraction expansion of x/y is $[a'_0, \ldots, a'_{s-1}]$.

Since the matrix on the left hand side of (39) is symmetric, the word $a'_0, \ldots, a'_s$ is a palindrome. In particular $a'_s = a'_0$.

Consider the periodic continued fraction

$$\delta = [a'_0, \overline{a'_1, \ldots, a'_{s-1}, 2a'_0}].$$

This number δ satisfies

$$\delta = [a'_0, a'_1, \ldots, a'_{s-1}, a'_0 + \delta].$$

Using the inverse of the matrix

$$\begin{pmatrix} a_0' & 1 \\ 1 & 0 \end{pmatrix} \quad \text{which is} \quad \begin{pmatrix} 0 & 1 \\ 1 & -a_0' \end{pmatrix},$$

we write

$$\begin{pmatrix} a_0' + \delta & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} a_0' & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \delta & 1 \end{pmatrix}$$

Hence the product of matrices associated with the continued fraction of δ

$$\begin{pmatrix} a_0' & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1' & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{s-1}' & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_0' + \delta & 1 \\ 1 & 0 \end{pmatrix}$$

is

$$\begin{pmatrix} Dy & x \\ x & y \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \delta & 1 \end{pmatrix} = \begin{pmatrix} Dy + \delta x & x \\ x + \delta y & y \end{pmatrix}.$$

It follows that

$$\delta = \frac{Dy + \delta x}{x + \delta y},$$

hence $\delta^2 = D$. As a consequence, $a_i' = a_i$ for $0 \leq i \leq s-1$ while $a_s' = a_0$, $a_s = 2a_0$.

This proves that if (x, y) is a non–trivial solution to Pell's equation $x^2 - Dy^2 = \pm 1$, then the continued fraction expansion of $\sqrt{D}$ is of the form

$$\sqrt{D} = [a_0, \overline{a_1, \ldots, a_{s-1}, 2a_0}] \tag{40}$$

with $a_1, \ldots, a_{s-1}$ a palindrome, and x/y is given by the convergent

$$x/y = [a_0, a_1, \ldots, a_{s-1}]. \tag{41}$$

Consider a convergent $p_n/q_n = [a_0, a_1, \ldots, a_n]$. If $a_{n+1} = 2a_0$, then (29) with $x = \sqrt{D}$ implies the upper bound

$$\left| \sqrt{D} - \frac{p_n}{q_n} \right| \leq \frac{1}{2a_0 q_n^2},$$

and it follows from Corollary 6 that (p_n, q_n) is a solution to Pell's equation $p_n^2 - Dq_n^2 = \pm 1$. This already shows that $a_i < 2a_0$ when $i+1$ is not the length of a period. We refine this estimate to $a_i \leq a_0$.

Assume $a_{n+1} \geq a_0 + 1$. Since the sequence $(a_m)_{m \geq 1}$ is periodic of period length s_0, for any m congruent to n modulo s_0, we have $a_{m+1} > a_0$. For these m we have

$$\left| \sqrt{D} - \frac{p_m}{q_m} \right| \leq \frac{1}{(a_0 + 1) q_m^2}.$$

For sufficiently large m congruent to n modulo s we have

$$(a_0+1)q_m^2 > q_m^2\sqrt{D}+1.$$

Corollary 6 implies that (p_m, q_m) is a solution to Pell's equation $p_m^2 - Dq_m^2 = \pm 1$. Finally, Theorem 37 implies that $m+1$ is a multiple of s_0, hence $n+1$ also.

□

3.3 Connection between the two formulae for the n-th positive solution to Pell's equation

Lemma 42. *Let D be a positive integer which is not a square. Consider the simple continued fraction expansion $\sqrt{D} = [a_0, \overline{a_1, \ldots, a_{s_0-1}, 2a_0}]$ where s_0 is the length of the fundamental period. Then the fundamental solution (x_1, y_1) to Pell's equation $x^2 - Dy^2 = \pm 1$ is given by the continued fraction expansion $x_1/y_1 = [a_0, a_1, \ldots, a_{s_0-1}]$. Let $n \geq 1$ be a positive integer. Define (x_n, y_n) by $x_n/y_n = [a_0, a_1, \ldots, a_{ns_0-1}]$. Then $x_n + y_n\sqrt{D} = (x_1 + y_1\sqrt{D})^n$.*

This result is a consequence of the two formulae we gave for the n-th solution (x_n, y_n) to Pell's equation $x^2 - Dy^2 = \pm 1$. We check this result directly.

Proof. From Lemma 30 and relation (39), one deduces

$$\begin{pmatrix} Dy_n & x_n \\ x_n & y_n \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{ns_0-1} & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Since

$$\begin{pmatrix} Dy_n & x_n \\ x_n & y_n \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -a_0 \end{pmatrix} = \begin{pmatrix} x_n & Dy_n - a_0x_n \\ y_n & x_n - a_0y_n \end{pmatrix},$$

we obtain

$$\begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{ns_0-1} & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} x_n & Dy_n - a_0x_n \\ y_n & x_n - a_0y_n \end{pmatrix}. \qquad (43)$$

Notice that the determinant is $(-1)^{ns_0} = x_n^2 - Dy_n^2$. Formula (43) for $n+1$ and the periodicity of the sequence $(a_1, \ldots, a_n, \ldots)$ with $a_{s_0} = 2a_0$ give :

$$\begin{pmatrix} x_{n+1} & Dy_{n+1} - a_0x_{n+1} \\ y_{n+1} & x_{n+1} - a_0y_{n+1} \end{pmatrix} = \begin{pmatrix} x_n & Dy_n - a_0x_n \\ y_n & x_n - a_0y_n \end{pmatrix} \begin{pmatrix} 2a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{s_0-1} & 1 \\ 1 & 0 \end{pmatrix}.$$

Take first $n = 1$ in (43) and multiply on the left by

$$\begin{pmatrix} 2a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -a_0 \end{pmatrix} = \begin{pmatrix} 1 & a_0 \\ 0 & 1 \end{pmatrix}.$$

Since

$$\begin{pmatrix} 1 & a_0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 & Dy_1 - a_0x_1 \\ y_1 & x_1 - a_0y_1 \end{pmatrix} = \begin{pmatrix} x_1 + a_0y_1 & (D - a_0^2)y_1 \\ y_1 & x_1 - a_0y_1 \end{pmatrix}.$$

we deduce

$$\begin{pmatrix} 2a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{s_0-1} & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} x_1 + a_0y_1 & (D - a_0^2)y_1 \\ y_1 & x_1 - a_0y_1 \end{pmatrix}.$$

Therefore

$$\begin{pmatrix} x_{n+1} & Dy_{n+1} - a_0x_{n+1} \\ y_{n+1} & x_{n+1} - a_0y_{n+1} \end{pmatrix} = \begin{pmatrix} x_n & Dy_n - a_0x_n \\ y_n & x_n - a_0y_n \end{pmatrix} \begin{pmatrix} x_1 + a_0y_1 & (D - a_0^2)y_1 \\ y_1 & x_1 - a_0y_1 \end{pmatrix}.$$

The first column gives

$$x_{n+1} = x_nx_1 + Dy_ny_1 \qquad \text{and} \qquad y_{n+1} = x_1y_n + x_ny_1,$$

which was to be proved.

□

3.4 Records

For large D, Pell's equation may obviously have small integer solutions. Examples are

for $D = m^2 - 1$ with $m \geq 2$, the numbers $x = m$, $y = 1$ satisfy $x^2 - Dy^2 = 1$,

for $D = m^2 + 1$ with $m \geq 1$, the numbers $x = m$, $y = 1$ satisfy $x^2 - Dy^2 = -1$,

for $D = m^2 \pm m$ with $m \geq 2$, the numbers $x = 2m \pm 1$, $y = 2$ satisfy $x^2 - Dy^2 = 1$,

for $D = t^2m^2 + 2m$ with $m \geq 1$ and $t \geq 1$, the numbers $x = t^2m + 1$, $y = t$ satisfy $x^2 - Dy^2 = 1$.

On the other hand, relatively small values of D may lead to large fundamental solutions. Tables are available on the internet[8].

[8] For instance:
Tomás Oliveira e Silva: Record-Holder Solutions of Pell's Equation
`http://www.ieeta.pt/~tos/pell.html`.

For D a positive integer which is not a square, denote by $S(D)$ the base 10 logarithm of x_1, when (x_1, y_1) is the fundamental solution to $x^2 - Dy^2 = 1$. The number of decimal digits of the fundamental solution x_1 is the integral part of $S(D)$ plus 1. For instance, when $D = 61$, the fundamental solution (x_1, y_1) is

$$x_1 = 1\,766\,319\,049, \quad y_1 = 226\,153\,980$$

and $S(61) = \log_{10} x_1 = 9.247\,069\ldots$

An integer D is a *record holder* for S if $S(D') < S(D)$ for all $D' < D$.

Here are the record holders up to 1021:

D	2	5	10	13	29	46	53	61	109
$S(D)$	0.477	0.954	1.278	2.812	3.991	4.386	4.821	9.247	14.198

D	181	277	397	409	421	541	661	1021
$S(D)$	18.392	20.201	20.923	22.398	33.588	36.569	37.215	47.298

Some further records with number of digits successive powers of 10:

D	3061	169789	12765349	1021948981	85489307341
$S(D)$	104.051	1001.282	10191.729	100681.340	1003270.151

3.5 Periodic continued fractions

An infinite sequence $(a_n)_{n\geq 0}$ is said to be *ultimately periodic* if there exists $n_0 \geq 0$ and $s \geq 1$ such that

$$a_{n+s} = a_n \quad \text{for all } n \geq n_0. \tag{44}$$

The set of s satisfying this property (3.5) is the set of positive multiples of an integer s_0, and $(a_{n_0}, a_{n_0+1}, \ldots, a_{n_0+s_0-1})$ is called *the fundamental period*.

A continued fraction with a sequence of partial quotients satisfying (44) will be written

$$[a_0, a_1, \ldots, a_{n_0-1}, \overline{a_{n_0}, \ldots, a_{n_0+s-1}}].$$

Example. For D a positive integer which is not a square, setting $a_0 = \lfloor \sqrt{D} \rfloor$, we have by Theorem 37

$$a_0 + \sqrt{D} = [\overline{2a_0, a_1, \ldots, a_{s-1}}] \quad \text{and} \quad \frac{1}{\sqrt{D} - a_0} = [\overline{a_1, \ldots, a_{s-1}, 2a_0}].$$

Lemma 45 (Euler 1737)**.** *If an infinite continued fraction*

$$x = [a_0, a_1, \ldots, a_n, \ldots]$$

is ultimately periodic, then x is a quadratic irrational number.

Proof. Since the continued fraction of x is infinite, x is irrational. Assume first that the continued fraction is periodic, namely that (44) holds with $n_0 = 0$:

$$x = [\overline{a_0, \ldots, a_{s-1}}].$$

This can be written

$$x = [a_0, \ldots, a_{s-1}, x].$$

Hence

$$x = \frac{p_{s-1}x + p_{s-2}}{q_{s-1}x + q_{s-2}}.$$

It follows that

$$q_{s-1}X^2 + (q_{s-2} - p_{s-1})X - p_{s-2}$$

is a non–zero quadratic polynomial with integer coefficients having x as a root. Since x is irrational, this polynomial is irreducible and x is quadratic.

In the general case where (44) holds with $n_0 > 0$, we write

$$x = [a_0, a_1, \ldots, a_{n_0-1}, \overline{a_{n_0}, \ldots, a_{n_0+s-1}}] = [a_0, a_1, \ldots, a_{n_0-1}, y],$$

where $y = [\overline{a_{n_0}, \ldots, a_{n_0+s-1}}]$ is a periodic continued fraction, hence is quadratic. But

$$x = \frac{p_{n_0-1}y + p_{n_0-2}}{q_{n_0-1}y + q_{n_0-2}},$$

hence $x \in \mathbf{Q}(y)$ is also quadratic irrational.

□

Lemma 46 (Lagrange, 1770)**.** *If x is a quadratic irrational number, then its continued fraction*

$$x = [a_0, a_1, \ldots, a_n, \ldots]$$

is ultimately periodic.

Proof. For $n \geq 0$, define $d_n = q_n x - p_n$. According to Corollary 28, we have $|d_n| < 1/q_{n+1}$.

Let $AX^2 + BX + C$ with $A > 0$ be an irreducible quadratic polynomial having x as a root. For each $n \geq 2$, we deduce from (26) that the convergent x_n is a root of a quadratic polynomial $A_nX^2 + B_nX + C_n$, with

$$\begin{aligned} A_n &= Ap_{n-1}^2 + Bp_{n-1}q_{n-1} + Cq_{n-1}^2, \\ B_n &= 2Ap_{n-1}p_{n-2} + B(p_{n-1}q_{n-2} + p_{n-2}q_{n-1}) + 2Cq_{n-1}q_{n-2}, \\ C_n &= A_{n-1}. \end{aligned}$$

Using $Ax^2 + Bx + C = 0$, we deduce

$$\begin{aligned} A_n &= (2Ax + B)d_{n-1}q_{n-1} + Ad_{n-1}^2, \\ B_n &= (2Ax + B)(d_{n-1}q_{n-2} + d_{n-2}q_{n-1}) + 2Ad_{n-1}d_{n-2}. \end{aligned}$$

There are similar formulae expressing A, B, C as homogeneous linear combinations of A_n, B_n, C_n, and since $(A, B, C) \neq (0, 0, 0)$, it follows that $(A_n, B_n, C_n) \neq (0, 0, 0)$. Since x_n is irrational, one deduces $A_n \neq 0$.

From the inequalities

$$q_{n-1}|d_{n-2}| < 1, \quad q_{n-2}|d_{n-1}| < 1, \quad q_{n-1} < q_n, \quad |d_{n-1}d_{n-2}| < 1,$$

one deduces

$$\max\{|A_n|,\ |B_n|/2, |C_n|\} < A + |2Ax + B|.$$

This shows that $|A_n|$, $|B_n|$ and $|C_n|$ are bounded independently of n. Therefore there exists $n_0 \geq 0$ and $s > 0$ such that $x_{n_0} = x_{n_0+s}$. From this we deduce that the continued fraction of x_{n_0} is purely periodic, hence the continued fraction of x is ultimately periodic. □

A *reduced quadratic irrational number* is an irrational number $x > 1$ which is a root of a degree 2 polynomial $ax^2 + bx + c$ with rational integer coefficients, such that the other root x' of this polynomial, which is the *Galois conjugate of* x, satisfies $-1 < x' < 0$. If x is reduced, then so is $-1/x'$.

Lemma 47. *A continued fraction*

$$x = [a_0, a_1, \ldots, a_n \ldots]$$

is purely periodic if and only if x is a reduced quadratic irrational number. In this case, if $x = [\overline{a_0, a_1, \ldots, a_{s-1}}]$ and if x' is the Galois conjugate of x, then

$$-1/x' = [\overline{a_{s-1}, \ldots, a_1, a_0}]$$

Proof. Assume first that the continued fraction of x is purely periodic:

$$x = [\overline{a_0, a_1, \dots, a_{s-1}}].$$

From $a_s = a_0$ we deduce $a_0 > 0$, hence $x > 1$. From $x = [a_0, a_1, \dots, a_{s-1}, x]$ and the unicity of the continued fraction expansion, we deduce

$$x = \frac{p_{s-1}x + p_{s-2}}{q_{s-1}x + q_{s-2}} \quad \text{and} \quad x = x_s.$$

Therefore x is a root of the quadratic polynomial

$$P_s(X) = q_{s-1}X^2 + (q_{s-2} - p_{s-1})X - p_{s-2}.$$

This polynomial P_s has a positive root, namely $x > 1$, and a negative root x', with the product $xx' = -p_{s-2}/q_{s-1}$. We transpose the relation

$$\begin{pmatrix} p_{s-1} & p_{s-2} \\ q_{s-1} & q_{s-2} \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{s-1} & 1 \\ 1 & 0 \end{pmatrix}$$

and obtain

$$\begin{pmatrix} p_{s-1} & q_{s-1} \\ p_{s-2} & q_{s-2} \end{pmatrix} = \begin{pmatrix} a_{s-1} & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Define

$$y = [\overline{a_{s-1}, \dots, a_1, a_0}],$$

so that $y > 1$,

$$y = [a_{s-1}, \dots, a_1, a_0, y] = \frac{p_{s-1}y + q_{s-1}}{p_{s-2}y + q_{s-2}}$$

and y is the positive root of the polynomial

$$Q_s(X) = p_{s-2}X^2 + (q_{s-2} - p_{s-1})X - q_{s-1}.$$

The polynomials P_s and Q_s are related by $Q_s(X) = -X^2 P_s(-1/X)$. Hence $y = -1/x'$.

For the converse, assume $x > 1$ and $-1 < x' < 0$. Let $(x_n)_{n\geq 1}$ be the sequence of complete quotients of x. For $n \geq 1$, define x'_n as the Galois conjugate of x_n. One deduces by induction that $x'_n = a_n + 1/x'_{n+1}$, that $-1 < x'_n < 0$ (hence x_n is reduced), and that a_n is the integral part of $-1/x'_{n+1}$.

If the continued fraction expansion of x were not purely periodic, we would have

$$x = [a_0, \ldots, a_{h-1}, \overline{a_h, \ldots, a_{h+s-1}}]$$

with $a_{h-1} \neq a_{h+s-1}$. By periodicity we have $x_h = [a_h, \ldots, a_{h+s-1}, x_h]$, hence $x_h = x_{h+s}$, $x'_h = x'_{h+s}$. From $x'_h = x'_{h+s}$, taking integral parts, we deduce $a_{h-1} = a_{h+s-1}$, a contradiction. □

Corollary 48. *If $r > 1$ is a rational number which is not a square, then the continued fraction expansion of $\sqrt{r}$ is of the form*

$$\sqrt{r} = [a_0, \overline{a_1, \ldots, a_{s-1}, 2a_0}]$$

with $a_1, \ldots, a_{s-1}$ a palindrome and $a_0 = \lfloor \sqrt{r} \rfloor$.
Conversely, if the continued fraction expansion of an irrational number $t > 1$ is of the form

$$t = [a_0, \overline{a_1, \ldots, a_{s-1}, 2a_0}]$$

with $a_1, \ldots, a_{s-1}$ a palindrome, then t^2 is a rational number.

Proof. If $t^2 = r$ is rational > 1, then for and $a_0 = \lfloor \sqrt{t} \rfloor$ the number $x = t + a_0$ is reduced. Since $t' + t = 0$, we have

$$-\frac{1}{x'} = \frac{1}{x - 2a_0}.$$

Hence

$$x = [\overline{2a_0, a_1, \ldots, a_{s-1}}], \quad -\frac{1}{x'} = [\overline{a_{s-1}, \ldots, a_1, 2a_0}]$$

and $a_1, \ldots, a_{s-1}$ a palindrome.

Conversely, if $t = [a_0, \overline{a_1, \ldots, a_{s-1}, 2a_0}]$ with $a_1, \ldots, a_{s-1}$ a palindrome, then $x = t + a_0$ is periodic, hence reduced, and its Galois conjugate x' satisfies

$$-\frac{1}{x'} = [\overline{a_1, \ldots, a_{s-1}, 2a_0}] = \frac{1}{x - 2a_0},$$

which means $t + t' = 0$, hence $t^2 \in \mathbf{Q}$. □

Lemma 49 (Serret, 1878)**.** *Let x and y be two irrational numbers with continued fractions*

$$x = [a_0, a_1, \ldots, a_n \ldots] \quad \text{and} \quad y = [b_0, b_1, \ldots, b_m \ldots]$$

respectively. Then the two following properties are equivalent.

(i) *There exists a matrix* $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *with rational integer coefficients and determinant* ± 1 *such that*

$$y = \frac{ax+b}{cx+d}.$$

(ii) *There exists* $n_0 \geq 0$ *and* $m_0 \geq 0$ *such that* $a_{n_0+k} = b_{m_0+k}$ *for all* $k \geq 0$.

Condition (i) means that x *and* y *are equivalent modulo the action of* $\mathrm{GL}_2(\mathbf{Z})$ *by homographies.*

Condition (ii) means that there exists integers n_0, m_0 and a real number $t > 1$ such that

$$x = [a_0, a_1, \ldots, a_{n_0-1}, t] \quad \text{and} \quad y = [b_0, b_1, \ldots, b_{m_0-1}, t].$$

Example.

$$\text{If } x = [a_0, a_1, x_2], \text{ then } \quad -x = \begin{cases} [-a_0 - 1, 1, a_1 - 1, x_2] & \text{if } a_1 \geq 2, \\ [-a_0 - 1, 1 + x_2] & \text{if } a_1 = 1. \end{cases} \tag{50}$$

Proof. We already know by (26) that if x_n is a complete quotient of x, then x and x_n are equivalent modulo $\mathrm{GL}_2(\mathbf{Z})$. Condition (ii) means that there is a partial quotient of x and a partial quotient of y which are equal. By transitivity of the $\mathrm{GL}_2(\mathbf{Z})$ equivalence, (ii) implies (i).

Conversely, assume (i):

$$y = \frac{ax+b}{cx+d}.$$

Let n be a sufficiently large number. From

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix} = \begin{pmatrix} u_n & u_{n-1} \\ v_n & v_{n-1} \end{pmatrix}$$

with

$$\begin{aligned} u_n &= ap_n + bq_n, \quad u_{n-1} = ap_{n-1} + bq_{n-1}, \\ v_n &= cp_n + dq_n, \quad v_{n-1} = cp_{n-1} + dq_{n-1}, \end{aligned}$$

we deduce

$$y = \frac{u_n x_{n+1} + u_{n-1}}{v_n x_{n+1} + v_{n-1}}.$$

We have $v_n = (cx+d)q_n + c\delta_n$ with $\delta_n = p_n - q_n x$. We have $q_n \to \infty$, $q_n \geq q_{n-1} + 1$ and $\delta_n \to 0$ as $n \to \infty$. Hence, for sufficiently large n, we have $v_n > v_{n-1} > 0$. From part 1 of Corollary 33, we deduce

$$\begin{pmatrix} u_n & u_{n-1} \\ v_n & v_{n-1} \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_s & 1 \\ 1 & 0 \end{pmatrix}$$

with $a_0, \dots, a_s$ in $\mathbf{Z}$ and $a_1, \dots, a_s$ positive. Hence

$$y = [a_0, a_1, \dots, a_s, x_{n+1}].$$

□

A computational proof of (i) ⇒ (ii). Another proof is given by Bombieri [3] (Theorem A.1 p. 209). He uses the fact that $\mathrm{GL}_2(\mathbf{Z})$ is generated by the two matrices

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

The associated fractional linear transformations are K and J defined by

$$K(x) = x + 1 \quad \text{and} \quad J(x) = 1/x.$$

We have $J^2 = 1$ and

$$K([a_0, t]) = [a_0 + 1, t], \quad K^{-1}([a_0, t]) = [a_0 - 1, t].$$

Also $J([a_0, t]) = [0, a_0, t]$ if $a_0 > 0$ and $J([0, t]) = \lfloor t \rfloor$. According to (50), the continued fractions of x and $-x$ differ only by the first terms. This completes the proof. [9]

□

3.6 Diophantine approximation and simple continued fractions

Lemma 51 (Lagrange, 1770)**.** *The sequence* $(|q_n x - p_n|)_{n \geq 0}$ *is strictly decreasing: for* $n \geq 1$ *we have*

$$|q_n x - p_n| < |q_{n-1} x - p_{n-1}|.$$

Proof. We use Lemma 27 twice: on the one hand

$$|q_n x - p_n| = \frac{1}{x_{n+1} q_n + q_{n-1}} < \frac{1}{q_n + q_{n-1}}$$

because $x_{n+1} > 1$, on the other hand

$$|q_{n-1} x - p_{n-1}| = \frac{1}{x_n q_{n-1} + q_{n-2}} > \frac{1}{(a_n + 1) q_{n-1} + q_{n-2}} = \frac{1}{q_n + q_{n-1}}$$

because $x_n < a_n + 1$. □

[9] Bombieri in [3] gives formulae for $J([a_0, t])$ when $a_0 \leq -1$. He distinguishes eight cases, namely four cases when $a_0 = -1$ ($a_1 > 2$, $a_1 = 2$, $a_1 = 1$ and $a_3 > 1$, $a_1 = a_3 = 1$), two cases when $a_0 = -2$ ($a_1 > 1$, $a_1 = 1$) and two cases when $a_0 \leq -3$ ($a_1 > 1$, $a_1 = 1$). Here, (50) enables us to simplify his proof by reducing to the case $a_0 \geq 0$.

Corollary 52. *The sequence* $(|x - p_n/q_n|)_{n\geq 0}$ *is strictly decreasing: for* $n \geq 1$ *we have*

$$\left|x - \frac{p_n}{q_n}\right| < \left|x - \frac{p_{n-1}}{q_{n-1}}\right|.$$

Proof. For $n \geq 1$, since $q_{n-1} < q_n$, we have

$$\left|x - \frac{p_n}{q_n}\right| = \frac{1}{q_n}|q_n x - p_n| < \frac{1}{q_n}|q_{n-1}x - p_{n-1}| = \frac{q_{n-1}}{q_n}\left|x - \frac{p_{n-1}}{q_{n-1}}\right| < \left|x - \frac{p_{n-1}}{q_{n-1}}\right|.$$

□

Here is the *law of best approximation* of the simple continued fraction.

Lemma 53. *Let* $n \geq 0$ *and* $(p, q) \in \mathbf{Z} \times \mathbf{Z}$ *with* $q > 0$ *satisfy*

$$|qx - p| < |q_n x - p_n|.$$

Then $q \geq q_{n+1}$.

Proof. The system of two linear equations in two unknowns u, v

$$\begin{cases} p_n u + p_{n+1} v &= p \\ q_n u + q_{n+1} v &= q \end{cases} \tag{54}$$

has determinant ± 1, hence there is a solution $(u, v) \in \mathbf{Z} \times \mathbf{Z}$.

Since $p/q \neq p_n/q_n$, we have $v \neq 0$.

If $u = 0$, then $v = q/q_{n+1} > 0$, hence $v \geq 1$ and $q \geq q_{n+1}$.

We now assume $uv \neq 0$.

Since q, q_n and q_{n+1} are > 0, it is not possible for u and v to be both negative. In case u and v are positive, the desired result follows from the second relation of (54). Hence one may suppose u and v of opposite signs. Since $q_n x - p_n$ and $q_{n+1}x - p_{n+1}$ also have opposite signs, the numbers $u(q_n x - p_n)$ and $v(q_{n+1}x - p_{n+1})$ have same sign, and therefore

$$|q_n x - p_n| \leq |u(q_n x - p_n)| + |v(q_{n+1}x - p_{n+1})| = |qx - p| < |q_n x - p_n|,$$

which is a contradiction.

□

A consequence of Lemma 53 is that the sequence of p_n/q_n produces the best rational approximations to x in the following sense: any rational number p/q with denominator $q < q_n$ has $|qx - p| > |q_n x - p_n|$. This is sometimes referred to as *best rational approximations of type* 0.

Corollary 55. *The sequence $(q_n)_{n\geq 0}$ of denominators of the convergents of a real irrational number x is the increasing sequence of positive integers for which*

$$\|q_n x\| < \|qx\| \quad \textit{for} \quad 1 \leq q < q_n.$$

As a consequence,

$$\|q_n x\| = \min_{1\leq q\leq q_n} \|qx\|.$$

The theory of continued fractions is developed starting from Corollary 55 as a definition of the sequence $(q_n)_{n\geq 0}$ in Cassels's book [7].

Corollary 56. *Let $n \geq 0$ and $p/q \in \mathbf{Q}$ with $q > 0$ satisfy*

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

Then $q > q_n$.

Proof. For $q \leq q_n$ we have

$$\left| x - \frac{p}{q} \right| = \frac{1}{q}|qx - p| > \frac{1}{q}|q_n x - p_n|\frac{q_n}{q} \left| x - \frac{p_n}{q_n} \right| \geq \left| x - \frac{p_n}{q_n} \right|.$$

□

Corollary 56 shows that the denominators q_n of the convergents are also among the *best rational approximations of type* 1 in the sense that

$$\left| x - \frac{p}{q} \right| > \left| x - \frac{p_n}{q_n} \right| \quad \text{for} \quad 1 \leq q < q_n,$$

but they do not produce the full list of them: to get the complete set, one needs to consider also some of the rational fractions of the form

$$\frac{p_{n-1} + ap_n}{q_{n-1} + aq_n}$$

with $0 \leq a \leq a_{n+1}$ (*semi–convergents*) – see for instance [23], Chap. II, § 16.

Lemma 57 (Vahlen, 1895)**.** *Among two consecutive convergents p_n/q_n and p_{n+1}/q_{n+1}, one at least satisfies $|x - p/q| < 1/2q^2$.*

Proof. Since $x - p_n/q_n$ and $x - p_{n-1}/q_{n-1}$ have opposite signs,

$$\left|x - \frac{p_n}{q_n}\right| + \left|x - \frac{p_{n-1}}{q_{n-1}}\right| = \left|\frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}}\right| = \frac{1}{q_n q_{n-1}} < \frac{1}{2q_n^2} + \frac{1}{2q_{n-1}^2}\cdot$$

The last inequality is $ab < (a^2 + b^2)/2$ for $a \neq b$ with $a = 1/q_n$ and $b = 1/q_{n-1}$. Therefore,

$$\text{either} \quad \left|x - \frac{p_n}{q_n}\right| < \frac{1}{2q_n^2} \quad \text{or} \quad \left|x - \frac{p_{n-1}}{q_{n-1}}\right| < \frac{1}{2q_{n-1}^2}\cdot$$

□

Lemma 58 (É. Borel, 1903)**.** *Among three consecutive convergents p_{n-1}/q_{n-1}, p_n/q_n and p_{n+1}/q_{n+1}, one at least satisfies $|x - p/q| < 1/\sqrt{5}q^2$.*

Compare with the implication (i) ⇒ (vi) in the irrationality criterion below (Proposition 61 in § 4.1).

That the constant $\sqrt{5}$ cannot be replaced by a larger one is proved in Lemma 66. This is true for any number with a continued fraction expansion having all but finitely many partial quotients equal to 1 (which means the Golden number Φ and all rational numbers which are equivalent to Φ modulo $\mathrm{GL}_2(\mathbf{Z})$).

Proof. Recall Lemma 27: for $n \geq 0$,

$$q_n x - p_n = \frac{(-1)^n}{x_{n+1}q_n + q_{n-1}}\cdot$$

Therefore $|q_n x - p_n| < 1/\sqrt{5}q_n$ if and only if $|x_{n+1}q_n + q_{n-1}| > \sqrt{5}q_n$. Define $r_n = q_{n-1}/q_n$. Then this condition is equivalent to $|x_{n+1} + r_n| > \sqrt{5}$.

Recall the inductive definition of the convergents:

$$x_{n+1} = a_{n+1} + \frac{1}{x_{n+2}}\cdot$$

Also, using the definitions of r_n, r_{n+1}, and the inductive relation $q_{n+1} = a_{n+1}q_n + q_{n-1}$, we can write

$$\frac{1}{r_{n+1}} = a_{n+1} + r_n.$$

Eliminate a_{n+1}:

$$\frac{1}{x_{n+2}} + \frac{1}{r_{n+1}} = x_{n+1} + r_n.$$

Assume now

$$|x_{n+1} + r_n| \leq \sqrt{5} \quad \text{and} \quad |x_{n+2} + r_{n+1}| \leq \sqrt{5}.$$

We deduce

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

which yields

$$r_{n+1}^2 - \sqrt{5} r_{n+1} + 1 \leq 0.$$

The roots of the polynomial $X^2 - \sqrt{5}X + 1$ are $\Phi = (1 + \sqrt{5})/2$ and $\Phi^{-1} = (\sqrt{5} - 1)/2$. Hence $r_{n+1} > \Phi^{-1}$ (the strict inequality is a consequence of the irrationality of the Golden ratio). .

This estimate follows from the hypotheses $|q_n x - p_n| < 1/\sqrt{5}q_n$ and $|q_{n+1}x - p_{n+1}| < 1/\sqrt{5}q_{n+1}$. If we also had $|q_{n+2}x - p_{n+2}| < 1/\sqrt{5}q_{n+2}$, we would deduce in the same way $r_{n+2} > \Phi^{-1}$. This would give

$$1 = (a_{n+2} + r_{n+1})r_{n+2} > (1 + \Phi^{-1})\Phi^{-1} = 1,$$

which is impossible.

□

Lemma 59 (Legendre, 1798)**.** *If $p/q \in \mathbf{Q}$ satisfies $|x - p/q| \leq 1/2q^2$, then p/q is a convergent of x.*

Proof. Let r and s in $\mathbf{Z}$ satisfy $1 \leq s < q$. From

$$1 \leq |qr - ps| = |s(qx - p) - q(sx - r)| \leq s|qx - p| + q|sx - r| \leq \frac{s}{2q} + q|sx - r|$$

one deduces

$$q|sx - r| \geq 1 - \frac{s}{2q} > \frac{1}{2} \geq q|qx - p|.$$

Hence $|sx - r| > |qx - p|$ and therefore Lemma 53 implies that p/q is a convergent of x.

□

3.7 A criterion for the existence of a solution to the negative Pell equation

Here is a recent result on the existence of a solution to Pell's equation $x^2 - Dy^2 = -1$

Proposition 60 (R.A. Mollin, A. Srinivasan[10])**.** *Let d be a positive integer which is not a square. Let (x_0, y_0) be the fundamental solution to Pell's equation $x^2 - dy^2 = 1$. Then the equation $x^2 - dy^2 = -1$ has a solution if and only if $x_0 \equiv -1 \pmod{2d}$.*

Proof. If $a^2 - db^2 = -1$ is the fundamental solution to $x^2 - dy^2 = -1$, then $x_0 + y_0\sqrt{d} = (a + b\sqrt{d})^2$, hence

$$x_0 = a^2 + db^2 = 2db^2 - 1 \equiv -1 \pmod{2d}.$$

Conversely, if $x_0 = 2dk - 1$, then $x_0^2 = 4d^2k^2 - 4dk + 1 = dy_0^2 + 1$, hence $4dk^2 - 4k = y_0^2$. Therefore y_0 is even, $y_0 = 2z$, and $k(dk - 1) = z^2$. Since k and $dk - 1$ are relatively prime, both are squares, $k = b^2$ and $dk - 1 = a^2$, which gives $a^2 - db^2 = -1$. □

3.8 Arithmetic varieties

Let D be a positive integer which is not a square. Define $\mathcal{G} = \{(x, y) \in \mathbf{R}^2 \ ; \ x^2 - Dy^2 = 1\}$.

The map

$$\begin{array}{ccc} \mathcal{G} & \longrightarrow & \mathbf{R}^\times \\ (x, y) & \longmapsto & t = x + y\sqrt{D} \end{array}$$

is bijective: the inverse of that map is obtained by writing $u = 1/t$, $2x = t + u$, $2y\sqrt{D} = t - u$, so that $t = x + y\sqrt{D}$ and $u = x - y\sqrt{D}$. By transfer of structure, this endows $\mathcal{G}$ with a multiplicative group structure, which is isomorphic to $\mathbf{R}^\times$, for which

$$\begin{array}{ccc} \mathcal{G} & \longrightarrow & \mathrm{GL}_2(\mathbf{R}) \\ (x, y) & \longmapsto & \begin{pmatrix} x & Dy \\ y & x \end{pmatrix}. \end{array}$$

is an injective group homomorphism. Let $G(\mathbf{R})$ be its image, which is therefore isomorphic to $\mathbf{R}^\times$.

A matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ respects the quadratic form $x^2 - Dy^2$ if and only if

$$(ax + by)^2 - D(cx + dy)^2 = x^2 - Dy^2,$$

which can be written

$$a^2 - Dc^2 = 1, \quad b^2 - Dd^2 = D, \quad ab = cdD.$$

[10] Pell equation: non-principal Lagrange criteria and central norms; Canadian Math. Bull., to appear

Hence the group of matrices of determinant 1 with coefficients in $\mathbf{Z}$ which respect the quadratic form $x^2 - Dy^2$ is the group

$$G(\mathbf{Z}) = \left\{ \begin{pmatrix} a & Dc \\ c & a \end{pmatrix} \in \mathrm{GL}_2(\mathbf{Z}) \right\}.$$

According to the work of Siegel, Harish–Chandra, Borel and Godement, the quotient of $G(\mathbf{R})$ by $G(\mathbf{Z})$ is compact. Hence $G(\mathbf{Z})$ is infinite (of rank 1 over $\mathbf{Z}$), which means that there are infinitely many solutions to the equation $a^2 - Dc^2 = 1$.

This is not a new proof of Proposition 2, but an interpretation and a generalization. Such results are valid for *arithmetic varieties*[11].

4 More on Diophantine Approximation

4.1 Irrationality Criterion

Proposition 61. *Let ϑ be a real number. The following conditions are equivalent:*
(i) *ϑ is irrational.*
(ii) *For any $\epsilon > 0$, there exists $(p, q) \in \mathbf{Z}^2$ such that $q > 0$ and*

$$0 < |q\vartheta - p| < \epsilon.$$

(iii) *For any $\epsilon > 0$, there exist two linearly independent linear forms in two variables*

$$L_0(X_0, X_1) = a_0X_0 + b_0X_1 \quad \textit{and} \quad L_1(X_0, X_1) = a_1X_0 + b_1X_1,$$

with rational integer coefficients, such that

$$\max\left\{|L_0(1, \vartheta)|\ ,\ |L_1(1, \vartheta)|\right\} < \epsilon.$$

(iv) *For any real number $Q > 1$, there exists an integer q in the range $1 \leq q < Q$ and a rational integer p such that*

$$0 < |q\vartheta - p| < \frac{1}{Q}.$$

[11] See for instance Nicolas Bergeron, "Sur la forme de certains espaces provenant de constructions arithmétiques", Images des Mathématiques, (2004).
`http://www.math.jussieu.fr/~bergeron/Recherche_files/Images.pdf`.

(v) *There exist infinitely many* $p/q \in \mathbf{Q}$ *such that*

$$\left|\vartheta - \frac{p}{q}\right| < \frac{1}{q^2}.$$

(vi) *There exist infinitely many* $p/q \in \mathbf{Q}$ *such that*

$$\left|\vartheta - \frac{p}{q}\right| < \frac{1}{\sqrt{5}q^2}.$$

The implication (vi) $\Rightarrow$ (v) is trivial. We shall prove (i) $\Rightarrow$ (vi) later (in the section on continued fractions). We now prove the equivalence between the other conditions of Proposition 61 as follows:

$$\text{(iv)} \Rightarrow \text{(ii)} \Rightarrow \text{(iii)} \Rightarrow \text{(i)} \Rightarrow \text{(iv)} \Rightarrow \text{(v) and (v)} \Rightarrow \text{(ii)}.$$

Notice that given a positive integer q, there is at most one value of p such that $|q\vartheta - p| < 1/2$, namely the nearest integer to $q\vartheta$. Hence, when we approximate ϑ by a rational number p/q, we have only one free parameter in $\mathbf{Z}_{>0}$, namely q.

In condition (v), there is no need to assume that the left hand side is not 0: if one $p/q \in \mathbf{Q}$ produces 0, then all other ones do not, and there are again infinitely many of them.

Proof of (iv) $\Rightarrow$ (ii). Using (iv) with Q satisfying $Q > 1$ and $Q \geq 1/\epsilon$, we get (ii). □

Proof of (v) $\Rightarrow$ (ii). According to (v), there is an infinite sequence of distinct rational numbers $(p_i/q_i)_{i\geq 0}$ with $q_i > 0$ such that

$$\left|\vartheta - \frac{p_i}{q_i}\right| < \frac{1}{\sqrt{5}q_i^2}.$$

For each q_i, there is a single value for the numerator p_i for which this inequality is satisfied. Hence the set of q_i is unbounded. Taking $q_i \geq 1/\epsilon$ yields (ii). □

Proof of (ii) $\Rightarrow$ (iii). Let $\epsilon > 0$. From (ii) we deduce the existence of $(p, q) \in \mathbf{Z} \times \mathbf{Z}$ with $q > 0$ and $\gcd(p, q) = 1$ such that

$$0 < |q\vartheta - p| < \epsilon.$$

We use (ii) once more with ϵ replaced by $|q\vartheta - p|$. There exists $(p', q') \in \mathbf{Z} \times \mathbf{Z}$ with $q' > 0$ such that

$$0 < |q'\vartheta - p'| < |q\vartheta - p|. \tag{62}$$

Define $L_0(X_0, X_1) = pX_0 - qX_1$ and $L_1(X_0, X_1) = p'X_0 - q'X_1$. It only remains to check that $L_0(X_0, X_1)$ and $L_1(X_0, X_1)$ are linearly independent. Otherwise, there exists $(s, t) \in \mathbf{Z}^2 \setminus (0, 0)$ such that $sL_0 = tL_1$. Hence $sp = tp'$, $sq = tq'$, and $p/q = p'/q'$. Since $\gcd(p, q) = 1$, we deduce $t = 1$, $p' = sp$, $q' = sq$ and $q'\vartheta - p' = s(q\vartheta - p)$. This is not compatible with (62). □

Proof of (iii) ⇒ (i). Assume $\vartheta \in \mathbf{Q}$, say $\vartheta = a/b$ with $\gcd(a, b) = 1$ and $b > 0$. For any non–zero linear form $L \in \mathbf{Z}X_0 + \mathbf{Z}X_1$, the condition $L(1, \vartheta) \neq 0$ implies $|L(1, \vartheta)| \geq 1/b$, hence for $\epsilon = 1/b$ condition (iii) does not hold. □

Proof of (i) ⇒ (iv) *using Dirichlet's box principle.* Let $Q > 1$ be a given real number. Define $N = \lceil Q \rceil$: this means that N is the integer such that $N - 1 < Q \leq N$. Since $Q > 1$, we have $N \geq 2$.

Let $\vartheta \in \mathbf{R} \setminus \mathbf{Q}$. Consider the subset E of the unit interval $[0, 1]$ which consists of the $N + 1$ elements

$$0,\ \{\vartheta\},\ \{2\vartheta\},\ \{3\vartheta\},\ \ldots, \{(N-1)\vartheta\},\ 1.$$

Since ϑ is irrational, these $N + 1$ elements are pairwise distinct. Split the interval $[0, 1]$ into N intervals

$$I_j = \left[\frac{j}{N}\,,\, \frac{j+1}{N}\right] \quad (0 \leq j \leq N - 1).$$

One at least of these N intervals, say I_{j_0}, contains at least two elements of E. Apart from 0 and 1, all elements $\{q\vartheta\}$ in E with $1 \leq q \leq N - 1$ are irrational, hence belong to the union of the *open* intervals $(j/N,\ (j+1)/N)$ with $0 \leq j \leq N - 1$.

If $j_0 = N - 1$, then the interval

$$I_{j_0} = I_{N-1} = \left[1 - \frac{1}{N}\ ,\ 1\right]$$

contains 1 as well as another element of E of the form $\{q\vartheta\}$ with $1 \leq q \leq N - 1$. Set $p = \lfloor q\vartheta \rfloor + 1$. Then we have $1 \leq q \leq N - 1 < Q$ and

$$p - q\vartheta = \lfloor q\vartheta \rfloor + 1 - \lfloor q\vartheta \rfloor - \{q\vartheta\} = 1 - \{q\vartheta\}, \quad \text{hence} \quad 0 < p - q\vartheta < \frac{1}{N} \leq \frac{1}{Q}.$$

Otherwise we have $0 \le j_0 \le N-2$ and I_{j_0} contains two elements $\{q_1\vartheta\}$ and $\{q_2\vartheta\}$ with $0 \le q_1 < q_2 \le N-1$. Set

$$q = q_2 - q_1, \quad p = \lfloor q_2\vartheta \rfloor - \lfloor q_1\vartheta \rfloor.$$

Then we have $0 < q = q_2 - q_1 \le N-1 < Q$ and

$$|q\vartheta - p| = |\{q_2\vartheta\} - \{q_1\vartheta\}| < 1/N \le 1/Q.$$

□

Remark. Theorem 1.A in Chap. II of [28] states that for any real number ϑ, for any real number $Q > 1$, there exists an integer q in the range $1 \le q < Q$ and a rational integer p such that

$$\left|\vartheta - \frac{p}{q}\right| \le \frac{1}{qQ}.$$

The proof given there yields strict inequality $|q\vartheta - p| < 1/Q$ in case Q is not an integer. In the case where Q is an integer and ϑ is rational, the result does not hold with a strict inequality in general. For instance, if $\vartheta = a/b$ with $\gcd(a,b) = 1$ and $b \ge 2$, there is a solution p/q to this problem with strict inequality for $Q = b+1$, but not for $Q = b$.

However, when Q is an integer and ϑ is irrational, the number $|q\vartheta - p|$ is irrational (recall that $q > 0$), hence not equal to $1/Q$.

Proof of (iv) ⇒ (v). Assume (iv). We already know that (iv) ⇒ (i), hence ϑ is irrational.

Let $\{q_1, \ldots, q_N\}$ be a finite set of positive integers. We are going to show that there exists a positive integer $q \notin \{q_1, \ldots, q_N\}$ satisfying the condition (v). Denote by $\|\cdot\|$ the distance to the nearest integer: for $x \in \mathbf{R}$,

$$\|x\| = \min_{a \in \mathbf{Z}} |x - a|.$$

Since ϑ is irrational, it follows that for $1 \le j \le N$, the number $\|q_j\vartheta\|$ is non–zero. Let $Q > 1$ satisfy

$$Q > \left(\min_{1 \le j \le N} \|q_j\vartheta\|\right)^{-1}.$$

From (iv) we deduce that there exists an integer q in the range $1 \le q < Q$ such that

$$0 < \|q\vartheta_i\| \le \frac{1}{Q}.$$

The right hand side is $< 1/q$, and the choice of Q implies $q \notin \{q_1, \ldots, q_N\}$.
□

4.2 Liouville's inequality

The main Diophantine tool for proving transcendence results is Liouville's inequality.

Recall that the ring $\mathbf{Z}[X]$ is factorial, its irreducible elements of positive degree are the non-constant polynomials with integer coefficients which are irreducible in $\mathbf{Q}[X]$ (i.e., not a product of two non-constant polynomials in $\mathbf{Q}[X]$) and have content 1. The *content* of a polynomial in $\mathbf{Z}[X]$ is the greatest common divisor of its coefficients.

The *minimal polynomial* of an algebraic number α is the unique irreducible polynomial $P \in \mathbf{Z}[X]$ which vanishes at α and has a positive leading coefficient.

The next lemma is one of many variants of Liouville's inequality (see, for instance, [28]), which is close to the original one of 1844.

Lemma 63. *Let α be an algebraic number of degree $d \geq 2$ and minimal polynomial $P \in \mathbf{Z}[X]$. Define $c = |P'(\alpha)|$. Let $\epsilon > 0$. Then there exists an integer q_0 such that, for any $p/q \in \mathbf{Q}$ with $q \geq q_0$,*

$$\left|\alpha - \frac{p}{q}\right| \geq \frac{1}{(c+\epsilon)q^d}.$$

Proof. The result is trivial if α is not real: an admissible value for q_0 is

$$q_0 = (c|\Im m(\alpha)|)^{-1/d}.$$

Assume now α is real. Let q be a sufficiently large positive integer and let p be the nearest integer to $q\alpha$. In particular,

$$\left|\alpha - \frac{p}{q}\right| \leq \frac{1}{2q}.$$

Denote by a_0 the leading coefficient of P and by $\alpha_1, \ldots, \alpha_d$ the roots with $\alpha_1 = \alpha$. Hence

$$P(X) = a_0(X-\alpha_1)(X-\alpha_2)\cdots(X-\alpha_d)$$

and

$$q^d P(p/q) = a_0 q^d \prod_{i=1}^{d} \left(\frac{p}{q} - \alpha_i\right). \tag{64}$$

Also

$$P'(\alpha) = a_0 \prod_{i=2}^{d} (\alpha - \alpha_i).$$

The left hand side of (64) is a rational integer. It is not zero because P is irreducible of degree ≥ 2. For $i \geq 2$ we use the estimate

$$\left|\alpha_i - \frac{p}{q}\right| \leq |\alpha_i - \alpha| + \frac{1}{2q}\cdot$$

We deduce

$$1 \leq q^d a_0 \left|\alpha - \frac{p}{q}\right| \prod_{i=2}^{d} \big(|\alpha_i - \alpha| + \frac{1}{2q}\big).$$

For sufficiently large q the right hand side is bounded from above by

$$q^d \left|\alpha - \frac{p}{q}\right| (|P'(\alpha)| + \epsilon).$$

□

4.1.2 Liouville's inequality for quadratic numbers

Consider Lemma 63 in the special case $d = 2$ where α is a quadratic algebraic number. Write its minimal polynomial $f(X) = aX^2 + bX + c$ and let $\Delta := b^2 - 4ac$ be its discriminant. Since we are interested in the approximation of α by rational numbers, we assume $\Delta > 0$. If $\alpha = (-b \pm \sqrt{\Delta})/2a$, then the other root is $\alpha' = (-b \mp \sqrt{\Delta})/2a$ and

$$f'(\alpha) = a(\alpha - \alpha') = \pm\sqrt{\Delta}.$$

Lemma 65. *Let α be an algebraic number of degree 2 and minimal polynomial $P \in \mathbf{Z}[X]$. Define $c = |P'(\alpha)|$. Let $\epsilon > 0$. Then there exists an integer q_0 such that, for any $p/q \in \mathbf{Q}$ with $q \geq q_0$,*

$$\left|\alpha - \frac{p}{q}\right| \geq \frac{1}{(\sqrt{\Delta} + \epsilon)q^2}\cdot$$

The smallest positive discriminant of an irreducible quadratic polynomial with coefficients in $\mathbf{Z}$ is 5, which is the value of the discriminant of $X^2 - X - 1$, with roots Φ and $-\Phi^{-1}$ where $\Phi = 1.6180339887499\ldots$ denotes the Golden ratio.

The next result deals with the Fibonacci sequence $(F_n)_{n\geq 0}$:

$$F_0 = 0,\ F_1 = 1,\ F_n = F_{n-1} + F_{n-2} \quad (n \geq 2).$$

Lemma 66. *For any $q \geq 1$ and any $p \in \mathbf{Z}$,*

$$\left|\Phi - \frac{p}{q}\right| > \frac{1}{\sqrt{5}q^2 + (q/2)}.$$

On the other hand

$$\lim_{n\to\infty} F_{n-1}^2 \left|\Phi - \frac{F_n}{F_{n-1}}\right| = \frac{1}{\sqrt{5}}.$$

Proof. It suffices to prove the lower bound when p is the nearest integer to $q\Phi$. From $X^2 - X - 1 = (X - \Phi)(X + \Phi^{-1})$ we deduce

$$p^2 - pq - q^2 = q^2 \left(\frac{p}{q} - \Phi\right)\left(\frac{p}{q} + \Phi^{-1}\right).$$

The left hand side is a non-zero rational integer, hence has absolute value at least 1. We now bound the absolute value of the right hand side from above. Since $p < q\Phi + (1/2)$ and $\Phi + \Phi^{-1} = \sqrt{5}$ we have

$$\frac{p}{q} + \Phi^{-1} < \sqrt{5} + \frac{1}{2q}.$$

Hence

$$1 < q^2 \left|\frac{p}{q} - \Phi\right| \left(\sqrt{5} + \frac{1}{2q}\right)$$

The first part of Lemma 66 follows.

The real vector space of sequences $(v_n)_{n\geq 0}$ satisfying $v_n = v_{n-1} + v_{n-2}$ has dimension 2, a basis is given by the two sequences $(\Phi^n)_{n\geq 0}$ and $((-\Phi^{-1})^n)_{n\geq 0}$. From this one easily deduces the formula

$$F_n = \frac{1}{\sqrt{5}}(\Phi^n - (-1)^n \Phi^{-n})$$

due to A. De Moivre (1730), L. Euler (1765) and J.P.M. Binet (1843). It follows that F_n is the nearest integer to

$$\frac{1}{\sqrt{5}}\Phi^n,$$

hence the sequence $(u_n)_{n\geq 2}$ of quotients of Fibonacci numbers

$$u_n = F_n/F_{n-1}$$

satisfies $\lim_{n\to\infty} u_n = \Phi$.

By induction one easily checks

$$F_n^2 - F_n F_{n-1} - F_{n-1}^2 = (-1)^{n-1}$$

for $n \geq 1$. The left hand side is $F_{n-1}^2(u_n - \Phi)(u_n + \Phi^{-1})$, as we already saw. Hence

$$F_{n-1}^2 |\Phi - u_n| = \frac{1}{\Phi^{-1} + u_n},$$

and the limit of the right hand side is $1/(\Phi + \Phi^{-1}) = 1/\sqrt{5}$. The result follows.

□

Remark. *The sequence $u_n = F_n/F_{n-1}$ is also defined by*

$$u_2 = 2,\ u_n = 1 + \frac{1}{u_{n-1}},\quad (n \geq 3).$$

Hence

$$u_n = 1 + \cfrac{1}{1 + \cfrac{1}{u_{n-2}}} = 1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{u_{n-3}}}} = \cdots$$

Remark. *It is known (see for instance [28] p. 25) that if k is a positive integer, if an irrational real number ϑ has a continued fraction expansion $[a_0; a_1, a_2, \dots]$ with $a_n \geq k$ for infinitely many n, then*

$$\liminf_{q \to \infty} q^2 \left| \vartheta - \frac{p}{q} \right| \leq \frac{1}{\sqrt{4 + k^2}}.$$

References

[1] E. J. Barbeau, *Pell's equation*, Problem Books in Mathematics, Springer-Verlag, New York, 2003.

[2] N. Bergeron, *Sur la forme de certains espaces provenant de constructions arithmétiques*, Images des Mathématiques, (2004).
http://people.math.jussieu.fr/∼bergeron/Recherche_files/Images.pdf.

[3] E. Bombieri, *Continued fractions and the Markoff tree*, Expo. Math., 25 (2007), pp. 187–213.

[4] E. Bombieri and A. J. van der Poorten, *Continued fractions of algebraic numbers*, in Computational algebra and number theory (Sydney, 1992), vol. 325 of Math. Appl., Kluwer Acad. Publ., Dordrecht, 1995, pp. 137–152.

[5] J.-P. Borel and F. Laubie, *Quelques mots sur la droite projective réelle*, J. Théor. Nombres Bordeaux, 5 (1993), pp. 23–51.

[6] Z. I. Borevitch and I. R. Chafarevitch, *Théorie des nombres*, Les Grands Classiques Gauthier-Villars. [Gauthier-Villars Great Classics], Éditions Jacques Gabay, Sceaux, 1993. Translated from the Russian by Myriam Verley and Jean-Luc Verley, Reprint of the 1967 French translation.

[7] J. W. S. Cassels, *An introduction to Diophantine approximation*, Hafner Publishing Co., New York, 1972. Facsimile reprint of the 1957 edition, Cambridge Tracts in Mathematics and Mathematical Physics, No. 45.

[8] H. Cohen, *Computational aspects of number theory*, in Mathematics unlimited—2001 and beyond, Springer, Berlin, 2001, pp. 301–330.

[9] H. Davenport, *The higher arithmetic*, Cambridge University Press, Cambridge, eighth ed., 2008. An introduction to the theory of numbers, With editing and additional material by James H. Davenport.

[10] D. Duverney, *Théorie des nombres – Cours et exercices corrigés*, Dunod, 1998.

[11] A. Faisant, *L'équation diophantienne du second degré*, Hermann, 1991.

[12] P. Flajolet, B. Vallée, and I. Vardi, *Continued fractions from euclid to the present day*. 44p.
`http://www.lix.polytechnique.fr/Labo/Ilan.Vardi/continued_fractions.ps`.

[13] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, Oxford University Press, Oxford, sixth ed., 2008. Revised by D. R. Heath-Brown and J. H. Silverman.

[14] M. Hindry, *Arithmétique*, Calvage & Mounet, 2008.

[15] M. J. Jacobson, Jr. and H. C. Williams, *Solving the Pell equation*, CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, Springer, New York, 2009.

[16] H. Lambert, *Mémoire sur quelques propriétés remarquables des quantités transcendantes circulaires et logarithmiques*, Mémoires de l'Académie des Sciences de Berlin, 17 (1768), pp. 265–322. Math. Werke, t. II
http://www.bibnum.education.fr/mathematiques/theorie-des-nombres/lambert-et-l-irrationalite-de-n-1761.

[17] H. W. Lenstra, Jr., *Solving the Pell equation*, Notices Amer. Math. Soc., 49 (2002), pp. 182–192.
http://www.ams.org/notices/200202/fea-lenstra.pdf.

[18] W. J. LeVeque, *Topics in number theory. Vol. I, II*, Dover Publications Inc., Mineola, NY, 2002. Reprint of the 1956 original [Addison-Wesley Publishing Co., Inc., Reading, Mass.; MR0080682 (18,283d)], with separate errata list for this edition by the author.

[19] N. H. McCoy, *The theory of numbers*, The Macmillan Co., New York, 1965.

[20] R. A. Mollin, *Quadratics*, CRC Press Series on Discrete Mathematics and its Applications, CRC Press, Boca Raton, FL, 1996.

[21] L. J. Mordell, *Diophantine equations*, Pure and Applied Mathematics, Vol. 30, Academic Press, London, 1969.

[22] I. Niven, H. S. Zuckerman, and H. L. Montgomery, *An introduction to the theory of numbers*, John Wiley & Sons Inc., New York, fifth ed., 1991.

[23] O. Perron, *Die Lehre von den Kettenbrüchen. Dritte, verbesserte und erweiterte Aufl. Bd. II. Analytisch-funktionentheoretische Kettenbrüche*, B. G. Teubner Verlagsgesellschaft, Stuttgart, 1957.

[24] G. N. Raney, *On continued fractions and finite automata*, Math. Ann., 206 (1973), pp. 265–283.

[25] K. H. Rosen, *Elementary number theory and its applications*, Addison-Wesley, Reading, MA, fourth ed., 2000.

[26] D. Roy, *On the continued fraction expansion of a class of numbers*, in Diophantine approximation, vol. 16 of Dev. Math., SpringerWien-NewYork, Vienna, 2008, pp. 347–361.
http://arxiv.org/abs/math/0409233.

[27] W. M. Schmidt, *Diophantine approximation*, vol. 785 of Lecture Notes in Mathematics, Springer, Berlin, 1980.

[28] ——, *Diophantine approximation*, vol. **785**, Lecture Notes in Mathematics. Berlin-Heidelberg-New York: Springer-Verlag, 1980, new ed. 2001.

[29] M. R. Schroeder, *Number theory in science and communication*, vol. 7 of Springer Series in Information Sciences, Springer-Verlag, Berlin, fourth ed., 2006. With applications in cryptography, physics, digital information, computing, and self-similarity.

[30] W. Sierpiński, *Elementary theory of numbers*, vol. 31 of North-Holland Mathematical Library, North-Holland Publishing Co., Amsterdam, second ed., 1988. Edited and with a preface by Andrzej Schinzel.

[31] H. M. Stark, *An introduction to number theory*, MIT Press, Cambridge, Mass., 1978.

[32] A. J. van der Poorten, *An introduction to continued fractions*, in Diophantine analysis (Kensington, 1985), vol. 109 of London Math. Soc. Lecture Note Ser., Cambridge Univ. Press, Cambridge, 1986, pp. 99–138.

[33] I. Vardi, *Archimedes' cattle problem*, Amer. Math. Monthly, 105 (1998), pp. 305–319.
`http://www.lix.polytechnique.fr/Labo/Ilan.Vardi/cattle.tex`.

[34] A. Weil, *Number theory*, Modern Birkhäuser Classics, Birkhäuser Boston Inc., Boston, MA, 2007. An approach through history from Hammurapi to Legendre, Reprint of the 1984 edition.

[35] D. Zagier, *On the number of Markoff numbers below a given bound*, Math. Comp., 39 (1982), pp. 709–723.

Michel WALDSCHMIDT
Université P. et M. Curie (Paris VI)
Institut Mathématique de Jussieu
Problèmes Diophantiens, Case 247
4, Place Jussieu
75252 Paris CEDEX 05, France
`miw@math.jussieu.fr` `http://www.math.jussieu.fr/~miw/`

This text is available on the internet at the address
http://www.math.jussieu.fr/∼miw/articles/pdf/BamakoPell2010.pdf