DOI: 10.15439/2017F314
ISSN 2300-5963 ACSIS, Vol. 11

Formalization of Pell's Equation in the Mizar System

Marcin Acewicz
University of Białystok,
Ciolkowskiego 1M, 15-245 Bialystok, Poland
Email: acewiczmarcin@gmail.com

Karol Pąk
University of Białystok,
Ciolkowskiego 1M, 15-245 Bialystok, Poland
Email: pakkarol@uwb.edu.pl

***Abstract*—We present a case study on a formalization of a textbook theorem that is listed as #39 at Freek Wiedijk's list of "Top 100 mathematical theorems". We focus on the formalization of the theorem that Pell's equation $x^2 - Dy^2 = 1$ has infinitely many solutions in positive integers for a given non square natural number D. We present also a formalization of the theorem that based on the least fundamental solution of the equation we can simply calculate algebraically each remaining solution.**

I. Introduction

THE work under the rigorous control of proof-assistants on a high formal level of trust eliminates all gaps which sometimes occur in informal proofs, especially in large ones. Any attempt to analyze the details of such formal certification is difficult in principle, however there are some exceptions. There are proof scripts whose authors put an extra effort to improve their readability [1], [2].

A. Paper Content and Contributions

We present our experience with the formalization of theorems related to the solvability of Pell's equation in the Mizar system[3], where we tried to obtain readable formalization. We focus on the approach represented in the textbook [4]. We show that the effort associated with this formalization is nontrivial, since we have to add to informal proofs all technical details that have been originally omitted.

Note that each fragment of the Mizar proof scripts contained in this paper comes from [5] available in the Mizar distribution.

II. Pell's Equation

Pell's equation (called alternatively the Fermat equation) is a special case of the quadratic Diophantine equation having the form $x^2 - Dy^2 = 1$, with D be a nonzero integer number. Generally, it is assumed that D is not a square since otherwise the equation can be solved using the difference of squares $x^2 - Dy^2 = (x + dy)(x - dy) = 1$. However in the context of Pell's equation, only non zero pairs of integers are being considered as solutions, excluding the trivial cases $x = 1$, $y = 0$ and $x = -1$, $y = 0$.

The solution of Pell's equation has been applied in many branches of mathematics. As the most basic we indicate here that based on solutions for a given non square natural D, we obtain a rational approximation for $\sqrt{D}$. There is also a correspondence between the solvability of Pell's equation and a special case of Dirichlet's unit theorem. It is also important to note that the Stormer's theorem applies Pell's equation to find pairs of consecutive smooth numbers.

From our point of view, the most significant application of Pell's equation was done by Yuri Matiyasevich to prove the undecidability of Hilbert's 10th problem. He analyzes a particular case $x^2 - (a^2 - 1)y^2 = 1$, where a is a natural number. He showed that solutions of such equation may grow exponentially and it was suffices to show that every computably enumerable set is diophantine. The solvability of this case and only such a case of Pell's equation has been already formalized in HOL Light [6] and Metamath [7]. However, in the case we can skip a complicated construction of a non trivial solution that is used for the general case, since pair $\langle a, 1\rangle$ is a solution.

A. Formalization in the Mizar System

In our formalization, we show that there exists a solution of Pell's equation for the general case, based on the approach used in the textbook [4] that is is very detailed. Nevertheless, fitting this approach to the limitations of a proof-checker system forced us to rebuild significantly the informal reasoning. In several situations we have to use an equivalent approach, to allow the use of already formalized facts in the Mizar Mathematical Library. Finally, we have to extract fragments of proofs as lemmas to highlight the main ideas of main theorems.

III. Formalization Details

In this section, we show the details of our formalization. We focus on two main theorems that determine the cardinality of the set that contains each solution of Pell's equation and dependencies between individual solutions. We show also the details of an important lemma that is used in the proof of the first theorem.

A. Basic lemma

The informal approach that is considered in the textbook [4] and is used to provide existence of at least one solution of Pell's equation is based on the following lemma:

Supported by Polish National Science Center grant decision nºDEC-2015/19/D/ST6/01473.

Lemma 3.1: If a natural number D is not the square of a natural number, then there exist infinitely many different pairs of integers x, y satisfying the inequalities $y \neq 0$ and $|x^2 - Dy^2| < 2\sqrt{D} + 1$.

which we formalized as follows:

```
theorem Th9:
  D is non square implies {[x,y] where x, y is Integer:
    y<>0 & |.x^2-D*y^2.|<2*sqrt D +1 &
      0<x-y*sqrt D} is infinite
```

It is important to note that in our reformulation we proved a slightly stronger theorem, where the pairs in the considered set satisfy an additional condition `0<x-y*sqrt D`. The condition can easily be deduced from the information collected in original informal proof of Lemma 3.1 and significantly facilitates application of the lemma in the main theorem. We distinguished three main stages in the proof of Lemma 3.1.

The first stage can be described as a remark that for each natural number n greater than 1 there exists a pair of integers x, y such that $0 < x - y\sqrt{D} < \frac{1}{n}$ with $0 < |y| \leq n$. We extract this stage as a theorem:

```
theorem Th6:
  D is non square & n > implies ex x,y be Integer st
    y<>0 & |.y.|<=n & 0<x-y*sqrt D<1/n
```

where the main idea of the proof can be described in the following way.

Let us consider a finite sequence $F : \{1, 2, \ldots, n+1\} \mapsto \mathbb{R}$ associate to any natural number $1 \leq i \leq n+1$ the floor $[(i-1)\sqrt{D} + 1]$. We have $0 < F(i) - (i-1)\sqrt{D} \leq 1$ for every $1 \leq i \leq n+1$. Moreover, $\sqrt{D}$ is an irrational number, hence $F(i) - (i-1)\sqrt{D} \neq F(j) - (j-1)\sqrt{D}$ for every $1 \leq i < j \leq n+1$. Then applying the pigeonhole principle (commonly called *Dirichlet's box principle*) it can be seen that there exist natural numbers i, j such that $i \neq j$ and $|(F(i) - (i-1)\sqrt{D}) - (F(j) - (j-1)\sqrt{D})| < \frac{1}{n}$, where as items we take the numbers $F(i) - (i-1)\sqrt{D}$ and as containers we take intervals: $\left]0, \frac{1}{n}\right], \left]\frac{1}{n}, \frac{2}{n}\right], \ldots, \left]\frac{n-1}{n}, 1\right]$. Now the proof of `Th6` is straightforward if we take $x := j - i$, $y := F(j) - F(i)$ or $x := i - j$, $y := F(i) - F(j)$.

To improve the main idea of `Th6` we formulate two theorems: the existence of such finite sequence F and a dedicated case of the pigeonhole principle:

```
theorem Th4:
  ex F be FinSequence of NAT st len F=n+1 &
    (for k st k in dom F holds F.k=[\ (k-1)*sqrt D/]+1) &
    (D is non square implies F is one-to-one)


theorem Th5:
  for a,b be Real, F be FinSequence of REAL st
    n>1 & len F=n+1 & (for k st k in dom F holds a<F.k<=b)
  holds
    ex i,j be Nat st i in dom F & j in dom F & i<>j &
      F.i<=F.j & F.j-F.i<(b-a)/n
```

Note that we present theorems as well as the majority of theorems in the paper without proofs which can be found in the proof script `PELLS_EQ.miz`.

The second stage can be formulated as an observation that there exists a pair of integers that fulfills property formulated in Lemma 3.1. However, the justification of its existence is "informally" repeated in the last stage as the sentence *In virtue of what we have proved before there exists at least one pair of integers x, y satisfying [...].* Therefore, to avoid repetition, we formulate a theorem that based on the assumption as well as the properties of the pair x, y formulated in `Th6` we can prove an additional property:

```
theorem Th7:
  D is non square & n<>0 & |.y.|<=n & 0<x-y*sqrt D <1/n
    implies |.x^2-D*y^2.|<=2*sqrt D+1/(n^2)
```

Then justification of this stage is a simple consequence of theorems labeled by `Th6`, `Th7`.

The justification of the third stage can be considered as a *complete* proof of Lemma 3.1 that refers to the earlier stages. Note that the justification has the form of an indirect proof, where the whole thesis of Lemma 3.1 is taken as an indirect assumption. A formal justification of the stage can be described as follows.

Let as define a set S of pairs considered in the Lemma 3.1 and suppose contrary to our claim that S is finite. Let us consider a function $f : S \mapsto \mathbb{R}$ that assigns $x - y\sqrt{D}$ for each pair $\langle x, y \rangle \in S$. We have that the range of f, denoted by R is finite since S is finite by the assumption and nonempty by `Th8`. Consequently, the *infimum* of R is a member of R and is positive as each element of R. Further, there exists a natural number n such that $\frac{1}{n}$ is less than the *infinium* of R. Then from `Th6` and `Th7` there exists a pair of integers x, y such that $y \neq 0$, $|x^2 - Dy^2| < 2\sqrt{D} + 1$, and $0 < x - y\sqrt{D} < \frac{1}{n}$. But the number $x - y\sqrt{D}$ is a member of R and is less than the *infimum*, which is impossible.

This finishes the justification of the third stage and consequently, the justification of Lemma 3.1.

B. Solvability of Pell's equation

The first main theorem that we take into consideration in our formalization is originally formulated as follows:

Theorem 3.1: If a natural number D is not the square of a natural number, then the equation $x^2 - Dy^2 = 1$ has infinitely many solutions in natural numbers x, y.

Since the theorem is one of the main results in our formalization, we have put an additional effort to obtain a readable formulation

```
theorem Th14:
  for D be non square Nat holds
    the set of all ab where ab is positive Pell's_solution of D
    is infinite
```

The informal justification can naturally be divided into two main stages. The first one states that Pell's equation has a solution in positive natural numbers and the second one that based on a given solution x, y we can construct another solution x', y' where $x' > x$, $y' > y$.

The first part of the first stage can be described as a theorem:

```
theorem Th10:
  D is non square implies ex k,a,b,c,d be Integer st 0 <> k &
    a^2-D*b^2 = k = c^2-D*d^2 &
    a,c are_congruent_mod k & b,d are_congruent_mod k &
    (|.a.|<>|.c.| or |.b.|<>|.d.|)
```

The proof of the theorem follows the idea of the textbook and includes constructions of successive infinite subsets of the set that is indicated in Lemma .3.1. Denote by S indicated there set. Note that for each pair x, y that belongs to S the expression $|x^2 - Dy^2|$ can have a finite number of nonzero natural values bounded by $2\sqrt{D}+1$. Consequently, there exists an infinite subset Z of S for which x^2-Dy^2 is equal to a fixed number k. Further, for each pair of Z we can assign a pair of remainders obtained by dividing by k. Note that there exist at most k^2 possible pairs of remainders. Therefore, there exists an infinite subset R of Z for which the pair of remainders is equal to a fixed one. Moreover we can choose two pairs a, b and c, d that belong to R that fulfil $|a| \neq |c|$ or $|b| \neq |d|$, since both equations can only occur in 4 cases.

To imitate the selection processes of an infinite subset, we use theorem from Mizar Mathematical Library labeled by CARD_2:101 in the Mizar article [8].

```
theorem :: CARD_2:101
  for F be Function st dom F is infinite & rng F is finite
  ex x st x in rng F & F"{x} is infinite;
```

It is important to note that we have to construct all necessary functions and justify their basic properties to use this theorem. In consequence, our formal justification of Th10 has almost 100 steps and is 5.19 times longer than the corresponding part of the informal one, if we compare the number of characters. Note that the proportion, called *de Bruijn factor*, calculated for whole our formalization is 3.62. However, the proportion is not so weak for each fragment of our formalization.

Let us focus on the reasoning contained in the remaining part of the first stage that can be summarized as

```
theorem Th11:
  D is non square implies ex x,y be Nat st x^2-D*y^2=1 & y<>0
```

The formal proof of this fact is comparable with the informal one. Therefore we will not focus on its details. However, in this case, we obtain de Bruijn factor equals 0.97.

As in the case of theorem Th9, a fragment of the original proof of Theorem 3.1 that corresponds to the second stage is used directly as a the proof of Th14. However, to be able to formulate Th14, we have to introduce two necessary definitions in our formalization.

First we define a solution of a given Pell's equation as each pair it of integers

```
definition
  let D be Nat;
  mode Pell's_solution of D -> Element of [:INT,INT:]
    means (it`1)^2 - D * (it`2)^2 = 1
```

where it`1 denotes the first coordinate of it and it`2 denotes the second ones.

Next, we define the concept of positive solutions of Pell's equation. A pair of real numbers is positive if both coordinates are positive and we formalize the adjective as follows:

```
definition
  let D1,D2 be real-membered non empty set;
  let p be Element of [:D1,D2:];
  attr p is positive means :Def2:
    p`1 is positive & p`2 is positive;
end;
```

Furthermore, to use the type positive Pell's_solution of D in the formulation of Th14, it is necessary to show non-emptiness for this type that is that exists at least one object of a given type. Obviously, we can justify this condition based on Th11 if D is a positive integer that is not a perfect square. We express this observation in the Mizar system as follows:

```
registration
  let D be non square Nat;
  cluster positive for Pell's_solution of D;
```

Based on this approach, we can start to prove Th14 based on the reasoning in the second stage. The main idea of the reasoning is expressed by the sentence:

If the equality $x^2 - Dy^2 = 1$ holds for natural numbers x, y then, clearly, $(2x^2-1)^2 - D(2xy)^2 = 1$ with $2xy > y$.

Obviously, its shows in a simple way that we can increase any number of times the second coordinate of a solution, generating in consequence infinitely many pairwise different solutions in natural numbers. Such kind of demonstration that a given set has infinite cardinality is typical in informal practice. However, a formalization could not strictly reflect it and we reflect the idea as follows:

Let P denotes the set of all pairs that correspond to positive solutions of $x^2 - Dy^2 = 1$. Suppose, contrary to our claim, that P is finite. By Th11 the set P is also non empty. Consequently, the set of the second coordinates of each pair that belongs to P, denoted by P_2 is also non empty and finite. Further, the *supremum* of P_2 is a member of P_2. Then there exists a positive pair of integers x, y such that $x^2 - Dy^2 = 1$ and y is the *supremum* of P_2. It is clear that $\langle 2x^2-1, 2xy\rangle$ is a member of P since $(2x^2-1)^2 - D(2xy)^2 = 1$. But then $2xy$ is a member of P_2 that is greater than the *supremum* of P_2, which is impossible.

C. The shape of all solutions of Pell's equation

The second main theorem that we take into consideration in our formalization is originally formulated as follows:

Theorem 3.2: If t_0, u_0 is the least solution of the equation $x^2 - Dy^2 = 1$ in natural numbers, then in order that a pair of natural numbers t, u be a solution of this equation it is necessary and sufficient for the equality $t + u\sqrt{D} = (t_0 + u_0\sqrt{D})^n$ to hold for a natural number n.

It is worth pointing out that the sentence *the least solution* in the context of a pair is quite confusing and requires an explanation. Note that in this context a pair of natural numbers x_0, y_0 is the least solution that satisfies $x^2 - Dy^2 = 1$ if and only if for each pair of natural numbers x_1, y_1 that also satisfies the equation holds $x_0 \leq x_1$ and $y_0 \leq y_1$. Obviously, the order that is used here is partial and the least element does not have to exist. However, it has been shown that the order is total on the set of solutions in natural numbers for a given Pell's equation. Therefore, imitating the original approach we prove the following two theorems:

```
theorem Th18:
  D is non square implies
    (p is positive iff p`1+p`2*sqrt D>1)

theorem Th19:
  1<p1`1+p1`2*sqrt D<p2`1+p2`2*sqrt D
  & D is non square
    implies p1`1<p2`1& p1`2<p2`2
```

where variables p1, p2 are Pell′s_solution of D. Additionally, we define a function that associates the least positive solution of the equation $x^2 - Dy^2 = 1$ with non square natural number D as:

```
definition
  let D be non square Nat;
  func min_Pell′s_solution_of D ->
        positive Pell′s_solution of D means :Def3:
  for p be positive Pell′s_solution of D holds
    it`1 <= p`1 & it`2 <= p`2;
```

where based on the registration that there exists a positive Pell′s_solution of D as well as theorems Th18, Th19 we prove that such solution exists and is unique.

Using the introduced functor, we can formulate Theorem 3.2 as follows:

```
theorem Th21:
  for D be non square Nat
    for p be Element of [:INT,INT:] holds
      p is positive Pell′s_solution of D
    iff ex n be Nat st p`1 + p`2 * sqrt D =
        ((min_Pell′s_solution_of D)`1 +
         (min_Pell′s_solution_of D)`2*sqrt D)|^n
```

The formulation of Th21 naturally suggests the division of its proof into two parts that justify the necessary and sufficient conditions, respectively. Moreover, the least solution of Pell's equation that is used in the sufficient condition can be simply replaced by any other solution, keeping the correctness of the justification. Therefore, we extract the condition as a theorem:

```
theorem Th20:
  for D be non square Nat, a,b be Integer, n be Nat
    p be positive Pell′s_solution of D
      n>0 & a+b*sqrt D=(p`1+ p`2*sqrt D)|^n
  holds [a,b] is positive Pell′s_solution of D
```

Note that the proof is immediate if we observe that based on the equality $a + b\sqrt{D} = (c + d\sqrt{D})^n$ we can provide that $a - b\sqrt{D} = (c - d\sqrt{D})^n$ and consequently $a^2 - Db^2 = (c^2 - Dd^2)^n$ under the condition that a, b, c, d are integer numbers and D is non square natural number (for more detail see the justification of theorem Th17 in our formalization).

Next, let us focus on the necessary condition, where the originally formulated justification is indirect. A formal justification of the condition is available in our formalization and can be described as follows.

Denote by $\langle x, y\rangle$ the least positive solution of a given Pell's equation, and suppose that $\langle t, u\rangle$ is a positive solution of the equation where $t + u\sqrt{D} \neq (x - y\sqrt{D})^n$ for each natural number n. Then there exists n (e.g. $\left[\frac{log_{10}(x+y\sqrt{D})}{log_{10}(t+u\sqrt{D})}\right]$) such that

$$(t + u\sqrt{D})^n < x + y\sqrt{D} < (t + u\sqrt{D})^{n+1}. \quad (1)$$

Obviously, there exists a pair of natural numbers t_n, u_n such that $t_n + u_n\sqrt{D} = (t + u\sqrt{D})^n$. By Th17 we have that $\langle t_n, u_n\rangle$ is a positive solution. Combining this with inequalities (1) multiplied by $t_n + u_n\sqrt{D}$ we obtain that

$$1 < (x + y\sqrt{D}) \cdot (t_n - u_n\sqrt{D}) = (xt_n - Dyu_n) + \sqrt{D}(yt_n - xu_n) < t + u\sqrt{D}. \quad (2)$$

Moreover, it is easy to check that $xt_n - Dyu_n, yt_n - xu_n > 0$ and $(xt_n - Dyu_n)^2 - D(yt_n - xu_n)^2 = 1$, hence $\langle xt_n - Dyu_n, yt_n - xu_n\rangle$ is a positive solution of considered equation. Then, combining Th19 with (2) we obtain that $xt_n - Dyu_n < x$ and $yt_n - xu_n < y$, which contradicts that $\langle x, y\rangle$ is the least.

This contradiction finally ends a formal justification of Theorem Th21.

IV. Conclusions

Our formalization has so far focused on Pell's equation, their solvability as well as the cardinality and shape of all possible solutions. Now we are working on the first stage of Matiyasevich's theorm.

We show that we can express Pell's equation and prove their properties in the Mizar environment obtaining a human readable formalization. Our effort allowed us to formulate great majority of theorems that precisely describe selected stages of informal deductions. Moreover, our work has provided many additional pieces of information that have been used implicitly in the textbook. Finally, the formalization can also be used as a basic course of formalization for inexperienced Mizar users.

References

[1] A. Grabowski, "Efficient rough set theory merging," *Fundamenta Informaticae*, vol. 135, no. 4, pp. 371–385, October 2014. doi: 10.3233/FI-2014-1129

[2] K. Pąk, "Readable formalization of Euler's partition theorem in Mizar," in *Intelligent Computer Mathematics – International Conference, CICM 2015, Washington, DC, USA, July 13–17, 2015, Proceedings*, 2015. doi: 10.1007/978-3-319-20615-8_14 pp. 211–226.

[3] G. Bancerek, C. Bylinski, A. Grabowski, A. Korniłowicz, R. Matuszewski, A. Naumowicz, K. Pąk, and J. Urban, "Mizar: State-of-the-art and beyond," in *Intelligent Computer Mathematics - International Conference, CICM 2015*, ser. LNCS, vol. 9150. Springer, 2015. doi: 10.1007/978-3-319-20615-8_17 pp. 261–279.

[4] W. Sierpiński, *Elementary theory of numbers*, A. Schinzel, Ed. Mathematical Institute of the Polisch Academy of Science, 1964.

[5] M. Acewicz and K. Pąk, "The Pell's equation," *Formalized Mathematics*, vol. 25, no. 3, 2017. doi: 10.1515/forma-2017-0019

[6] J. Harrison, "The HOL Light system REFERENCE," 2014, http://www.cl.cam.ac.uk/ jrh13/hol-light/reference.pdf.

[7] N. D. Megill, "Metamath: A Computer Language for Pure Mathematics," 2007, http://us.metamath.org/downloads/metamath.pdf.

[8] G. Bancerek, "Cardinal arithmetics," *Formalized Mathematics*, vol. 1, no. **3**, pp. 543–547, 1990.