Pell's equation (II)

[redacted]

4/1/2022

1 Introduction

We attempt to achieve a complete formalization of the existence of solutions to Pell's equation as proved in [1]. Namely,

$$x^2 - dy^2 = 1 \ldots (*)$$

The main theorem we attempt to formalize is that if d is nonsquare, then $(*)$ always has a non-trivial solution. We formalize this in three steps. Each step corresponds to 1 file.

Theorem 1. Let a be an irrational number, and $Q > 1$ be an integer. Then there exist p, q integers with $1 \leq q < Q$ such that $|p - qa| < \frac{1}{Q}$.

Corollary 2 (Dirichlet). For any a irrational, there are infinitely many $\frac{p}{q}$ such that $|a - \frac{p}{q}| < \frac{1}{q^2}$.

Theorem 3. For any nonsquare d there is a nontrivial solution to

$$x^2 - dy^2 = 1.$$

In this project, we achieved a complete formalization of theorem 1 and corollary 2 and an almost complete formalization to Theorem 3. We will discuss our implementation.

Theorem 1

In our last project, we attempted to formalize this Theorem but did not succeed. In this project, we realize our formalization failed for multiple reasons only one of which was noted before:

- The use of the Pigeonhole Principle on pairs (p, q) rather than just on a single variable q.
- The argument presented before was strictly incomplete. In that. we considered the set $\mathbb{Z} \cap [1, Q]$, but actually we need to consider the set $\mathbb{Z} \cap [1, Q)$ since we need $q < Q$. This means the entire pigeonhole argument fails until we figure out how to remove one of the intervals.When trying to fix this, we realized that we use the condition that a is irrational. (We did not assume this before.) Let's present a complete argument, one of which we formalized.

We consider Q numbers $\{ka\}$ for $k \in \mathbb{Z} \cap [0, Q)$. We divide the interval $[0, 1)$ into Q intervals of length $1/Q$. Now, it is easy to see that if $\exists k$ such that $\{ka\}$ lies in the first or last subinterval, then we are done by subtracting 0 or subtracting from 1 respectively. Fortunately, the formula for p, q when $\{ka\}$ lies in the first interval is covered by the general formula after pigeonhole principle since we can consider 0 as $\{0 * a\}$. So there really is just one edge-case, when $\{ka\}$ lies in the last interval.

Hence, we dealt with this case separately as our first step to proving Theorem 1. That allows us to remove one sub-interval and so there are now Q numbers but $Q - 1$ sub-intervals. So we can pigeonhole, but there is a catch:

$$\{\{0 * a\}, \{a\}, \ldots \{Qa\}\} = \{0, 1/Q, \ldots (Q-1)/Q\}.$$

Clearly, this cannot happen because a is irrational. However, this case has not been excluded yet, and to apply pigeonhole principle and then arrive at this very case would have been very hard to formalize. Fortunately, by a small change in order of arguments, the formalization is doable.

Instead, we first formalized `imp_irrational_approx` as a separate lemma that Theorem 1 is equivalent to

Theorem 1'. Let a be an irrational number, and $Q > 1$ be an integer. Then there exist p, q integers with $1 \leq q < Q$ such that $|p - qa| \leq \frac{1}{Q}$.

This is because if we have equality $|p-qa| = 1/Q$, then a cannot be irrational. In fact we proved something more general `irrational_ne_rational_abs` that

$$x, y \in \mathbb{Z}, y \neq 0, |p - q * a| = x/y \implies \texttt{false}.$$

We will use this later as well. Now, in the formalization of Theorem 1, we can simply use the Pigeonhole Principle without having to make the difficult observation earlier. Instead, we consider the subintervals

$$\left[0, \frac{1}{Q}\right], \left[\frac{1}{Q}, \frac{2}{Q}\right], \ldots, \left[\frac{Q-2}{Q}, \frac{Q-1}{Q}\right]$$

and we are done. This is our completely new formalization of Theorem 1 coupled with the fact that we pigeonholed on a single variable.

Corollary 2

We start by formalizing $|p - qa|$ is strictly positive because we will divide by $|p - qa|$ in many inequalities henceforth. This is just an application of `irrational_ne_rational_abs` and the fact that the absolute value is non-negative. Next, we formalize a lemma `div_q_eq` useful in numerous proofs onwards,

$$1 \leq q \implies |a - p/q| = |p - qa|/q.$$

And, additionally an inequality version `div_q_lt`, for $b \in \mathbb{R}$, $1 \leq q$,

$$|p - qa| < bq \iff |a - p/q| < b.$$

The challenge is, the main result states there are *infinitely* many p/q. We formalize this in three steps:

1. Given such p, q there is p', q' satisfying the assumptions of Corollary 2 such that $|p' - q'a| < |p - qa|$. Indeed, we find an integer Q such that $1/Q < |a - p/q|$. This requires the Archimedean property and `div_q_eq`. Then we use Theorem 1 to find p', q' so that

$$|a - p'/q'| < 1/Qq' < 1/(q')^2$$

and

$$|p' - q'a| < 1/Q < |a - p/q| < |p - qa|.$$

So this is indeed a smaller pair.

2. There exists an initial pair satisfying Corollary 2. i.e. the set of such p/q is non-empty. We applied Theorem 1 to $Q = 2$ and verified it satisfies Corollary 2.

3. The set $\mathcal{S} = \{(p, q) \in \mathbb{Z} \times \mathbb{Z} \mid p, q \text{ satisfy Corollary 2}\}$ is infinite. We approach this by contradiction. If this set is finite then we can construct a function f that maps $(p, q) \in \mathcal{S}$ to $|p - qa|$. In particular, $\mathrm{Im}(f)$ has a minimum by `fintype.exists_min` which requires $\mathcal{S}$ to be nonempty (step 2). But then, by step 1, we're done because there does exist a strictly smaller pair $(p', q') \in \mathcal{S}$ wrt f.

Theorem 3

Before formalizing Theorem 3, we prove some lemmas.
`inf_bounded_pairs`: $\forall (p, q) \in \mathcal{S}, |N(p - q\sqrt{d})| \leq 2\sqrt{d} + 1$. The norm was imported in the first file, so this is equivalent to

$$|p - q\sqrt{d}||p + q\sqrt{d}| \leq \frac{1}{q}\left(2q\sqrt{d} + \frac{1}{q}\right) \leq 2\sqrt{d} + 1.$$

where we use Corollary 2 for bounding $|p - q\sqrt{d}|$ and the triangle inequality combined with Corollary 2 for bounding $|p + q\sqrt{d}|$. But we showed $\mathcal{S}$ is infinite so this essentially says there are infinitely many pairs (p, q) with bounded norm. Hence this motivates us to define $\mathcal{A} := \{x \in \mathbb{Z} \mid |x| \leq 2\sqrt{d} + 1\}$.

`A_finite`: We show that $\mathcal{A}$ is bounded above and below and finish off with `bdd_below.finite_of_bdd_above`. $\lfloor -2\sqrt{d} - 1 \rfloor$ is a lower bound while $\lceil 2\sqrt{d} + 1 \rceil$ is an upper bound. Note our use of integer lower and upper bounds to ease formalization.

`pell_ne_zero`: $1 \leq y \implies x^2 - dy^2 \neq 0$. This is needed extensively in proving Theorem 3, because after pigeonhole we might end up with norm 0, and then we cannot divide. To prove this we reduced it to the equation $\sqrt{d} = (x/y)^2$ which is impossible since we have $\sqrt{d}$ is irrational from the

library.

We are ready to pigeonhole and prove Theorem 3. Firstly, $\mathcal{S}$ is infinite but $\mathcal{A}$ is finite. Through the norm map, $N : \mathcal{S} \to \mathcal{A}$. So by infinite pigeonhole, there are infinitely many pairs $(p, q) \in \mathcal{S}$ all with the same norm N. Clearly $N \neq 0$ due to `pell_ne_zero`. Hence, we can define $\mathcal{B}$ as the set of remainders mod N. Next, we need to show $\mathcal{B}$ is finite in order to pigeonhole again. Indeed, by the same strategy as before, we show 0 is a lower bound and $|N|$ is an upper bound. Next, we construct $\mathcal{C} = \mathcal{B} \times \mathcal{B}$ which is also finite because `set.finite.prod` states the product of finite sets is finite. Now we construct a map g that reduces the pairs (p, q) with norm N to $(p\%N, q\%N)$ i.e. reduction modulo N. By infinite pigeonhole again, there exist two pairs $(p_1, q_1) \neq (p_2, q_2)$ such that $(p_1\%N, q_1\%N) = (p_2\%N, q_2\%N)$ and $N(p_1 - q_1\sqrt{d}) = N(p_2 - q_2\sqrt{d}) = N$. This is as far as we got in our formalization which is on the right track.

What's left are the following exercises. We've already formalized the hard part of the proof which was the double application of PHP. The rest is a matter of rewrites and piecing the pieces together.

- $N \mid p_1p_2 - dq_1q_2$ and $N \mid p_1q_2 - p_2q_1$ and hence the quotient x, y respectively are integers.
- $\dfrac{p_1 - q_1\sqrt{d}}{p_2 - q_2\sqrt{d}} = \dfrac{p_1p_2 - dq_1q_2 + (p_1q_2 - p_2q_1)\sqrt{d}}{N}$. This would imply $x + y\sqrt{d}$ has norm 1 and hence solves Pell's equation by `pell_to_norm`.
- $\dfrac{p_1 - q_1\sqrt{d}}{p_2 - q_2\sqrt{d}} \neq 1$ because $(p_1, q_1) \neq (p_2, q_2)$. Hence the solution is indeed nontrivial because we show that if $y = 0$ then $x = 1$ and hence $x + y\sqrt{d} = 1$.

References

1. MATH60041 Number Theory Lecture notes by Dr. David Helm