Diophantine Equation $y^2 + 2 = x^3$

Introduction

In the 1650's French mathematician Pierre de Fermat wanted to show to the English mathematical community that he is better than they are. Therefore he declared that he knew the proof of assertion that the only integer solutions of the equation $y^2 + 2 = x^3$ are $(3, \pm 5)$. Nobody of Fermat's contemporaries solved this problem and a correct proof was given 150 years later. The program of this lecture will be an elementary proof of this assertion.

Suppose we are interested in solutions of this equation in rational numbers. It is interesting that there is so-called *duplication formula* (which can be derived from adding point (x, y) with itself in a group which contains all points of cubic (elliptic) curve $y^2 = x^3 - 2$): if $(x, y) \in \mathbb{Q}^2$ is a solution of the equation $y^2 + 2 = x^3$, then another solution is also

$$\left(\frac{x^4 + 16x}{4y^2}, \frac{x^6 - 40x^3 - 32}{8y^3} \right).$$

So we obtain a sequence of solutions:

$$(3, 5), \left(\frac{129}{10^2}, -\frac{383}{10^3} \right), \left(\frac{2\,340\,922\,881}{7\,660^2}, \frac{113\,259\,286\,337\,279}{7\,660^3} \right), \ldots.$$

Using Nagell-Lutz theorem we can prove that in this way we obtain infinite number of distinct solutions of this equation. But in integers there are only two solutions, so let's prove this assertion.

First Step

Our equation is equivalent to

$$\left(y + \sqrt{-2}\right)\left(y - \sqrt{-2}\right) = x^3,$$

which is an equation in the integral domain $\mathbb{Z}[\sqrt{-2}] = \{a + b\sqrt{-2}; a, b \in \mathbb{Z}\}$. Main part of the solution is proving the fact that if $(y+\sqrt{-2})(y-\sqrt{-2}) = x^3$ for $x, y \in \mathbb{Z}$, then $y + \sqrt{-2}$ is a cube in $\mathbb{Z}[\sqrt{-2}]$, i.e. there exist $a, b \in \mathbb{Z}$ such that $y + \sqrt{-2} = (a + b\sqrt{-2})^3$. Then

$$y + \sqrt{-2} = (a^3 - 6ab^2) + (3a^2b - 2b^3)\sqrt{-2}.$$

Comparing the real and imaginary parts we obtain

$$\begin{aligned} y &= a^3 - 6ab^2 = a \cdot (a^2 - 6b^2), \\ 1 &= 3a^2b - 2b^2 = b \cdot (3a^2 - 2b^2). \end{aligned}$$

Hence $b \mid 1$ so $b = \pm 1$. It follows that $a = \pm 1$. Substituting these numbers into the first equation, we find that $y = \pm 5$. Because $x^3 = (\pm 5)^2 + 2$, then $x = 3$, hence the only solutions of given equation are $(x, y) = \pm 5$.

Possible Problems

Let's turn our attention to the equality

$$k \cdot l = x^3.$$

If it is an equality in the integral domain $(\mathbb{Z}, +, \cdot)$, then, by the theorem about unique factorization into primes, we may write $x = \pm p_1^{\alpha_1} \cdot \ldots \cdot p_s^{\alpha_s}$ and

$$k \cdot l = (\pm 1)^3 p_1^{3\alpha_1} \cdot \ldots \cdot p_s^{3\alpha_s},$$

where the set of primes $\{p_1, \ldots, p_s\}$ and exponents $\alpha_1, \ldots, \alpha_s \in \mathbb{N}$ are unique. If also $gcd(k, l) = 1$ (i.e. k, l are relatively prime), then by uniqueness of factorization we obtain

$$k = \pm p_{i_1}^{3\alpha_{i_1}} \cdot \ldots \cdot p_{i_c}^{3\alpha_{i_c}}, l = \pm p_{j_1}^{3\alpha_{j_1}} \cdot \ldots \cdot p_{j_d}^{3\alpha_{j_d}}$$

for suitable mutually distinct $i_1, \ldots, i_c, j_1, \ldots, j_d \in \{1, \ldots, s\}$. Because 1 and -1 are cubes in $\mathbb{Z}$, there exist $k_1, l_1 \in \mathbb{Z}$ such that

$$k = k_1^3, l = l_1^3.$$

If $\mathbb{Z}[\sqrt{-2}]$ is also unique factorization domain (UFD), it is a chance that we'll be able to finish the proof that $y + \sqrt{-2}$ is a cube in $\mathbb{Z}[\sqrt{-2}]$. But for integral domains $\mathbb{Z}[\sqrt{D}]$, where $D \in \mathbb{Z} \setminus \{0, 1\}$ is square-free, the situation about equality $k \cdot l = x^3$ is relatively complicated.

For example in $\mathbb{Z}[\sqrt{-23}]$ (which isn't UFD) we have

$$\left(2 + \sqrt{-23}\right)\left(2 - \sqrt{-23}\right) = 3^3$$

(which is very similar to the equation $(y + \sqrt{-2})(y - \sqrt{-2}) = x^3$), but $2 \pm \sqrt{-23}$ and 3 are irreducibles (i.e. they have similar properties as primes in $\mathbb{Z}$), hence $2 + \sqrt{-23}$ and $2 - \sqrt{-23}$ are relatively prime but there are no $k_1, l_1 \in \mathbb{Z}[\sqrt{-23}]$ such that $2 + \sqrt{-23} = k_1^3$ and $2 - \sqrt{-23} = l_1^3$.

Some problems may be also caused by units (invertible elements) in $\mathbb{Z}[\sqrt{D}]$. For example in $\mathbb{Z}[\sqrt{2}]$ (which is by the way UFD) it is true that all units are $\pm(1 + \sqrt{2})^n$ for $n \in \mathbb{Z}$. Surely $(1 + \sqrt{2})(1 + \sqrt{2})^2 = (1 + \sqrt{2})^3$, but $1 + \sqrt{2}$ isn't a cube in $\mathbb{Z}[\sqrt{2}]$: if

$$1 + \sqrt{2} = (a + b\sqrt{2})^3 = a^3 + 3a^2b\sqrt{2} + 6ab^2 + 2b^3\sqrt{2},$$

then

$$1 = a \cdot (a^2 + 6b^2), 1 = b \cdot (3a^2 + 2b^2),$$

hence $a \mid 1$ and $b \mid 1$, so $a = \pm 1, b = \pm 1$ and after substituting into the first equation we obtain $1 = \pm 7$ which is impossible. Further $1 + \sqrt{2}$ and $(1 + \sqrt{2})^2$ are relatively prime because both numbers are invertible in $\mathbb{Z}[\sqrt{2}]$.

So we see that during analysis of the equation $(y + \sqrt{-2})(y - \sqrt{-2}) = x^3$, we must be careful of units in $\mathbb{Z}[\sqrt{-2}]$ and it will be useful to prove that $\mathbb{Z}[\sqrt{-2}]$ is UFD. For doing this, we need to define some terms. We have already used informally some of these terms.

Norm Map and Its Usefulness

Definition. Let R be an integral domain. If $a, b \in R$, we'll say that a *divides* b and write $a \mid b$ if there exists some $c \in R$ such that $ac = b$.
Any divisor of 1 is called *unit*.
We'll say that a and b are *associates* and write $a \sim b$ if there exists a unit $u \in R$ such that $a = bu$.
We'll say that $\pi \in R \setminus \{0\}$ is *irreducible* if π is not a unit and for any factorization $\pi = bc$, either b or c is a unit.
We'll say that $a, b \in R$ are *relatively prime* if

$$\forall r \in R\colon r \mid a \text{ and } r \mid b \Rightarrow r \text{ is a unit.}$$

Definition. Let R be an integral domain. If there is a map $N\colon R\setminus\{0\} \to \mathbb{N}$ such that:

(i) $\forall a, b \in R \setminus \{0\}\colon N(ab) = N(a)N(b)$,

(ii) $N(a) = 1 \Leftrightarrow a$ is a unit,

this map is called a *norm map*.

Now using a norm we'll prove the existence of the factorization into primes in $\mathbb{Z}[\sqrt{-2}]$.

Proposition. *Let R be an integral domain with a norm map N. Then every nonunit element $a \in R \setminus \{0\}$ can be written as a product of irreducible elements.*

Proof. Let S be the set of all nonunit elements of $R \setminus \{0\}$ that cannot be written as a product of irreducibles. If $S \neq \emptyset$, take $a \in S$ with the least norm. Since a isn't irreducible, then there exist nonunits $b, c \in R \setminus \{0\}$ such that $a = bc$. Then $N(a) = N(b)N(c)$ and $N(b) > 1, N(c) > 1$, hence $N(b) < N(a)$ and $N(c) < N(a)$. Then $b, c \notin S$ and we can write b and c as a product of irreducibles. Hence we can write also $a = b \cdot c$ as a product of irreducibles, which is a contradiction because $a \in S$. Then $S = \emptyset$ and the proposition is proved. □

Corollary. *Let $D \in \mathbb{Z} \setminus \{0, 1\}$ be square-free. Then every nonunit and nonzero element of integral domain $\mathbb{Z}[\sqrt{D}] = \{a + b\sqrt{D}; a, b \in \mathbb{Z}\}$ can be written as a product of irreducible elements.*

Proof. Define $N\colon \mathbb{Z}[\sqrt{D}]\setminus\{0\} \to \mathbb{N}$ as follows: for any $a+b\sqrt{D} \in \mathbb{Z}[\sqrt{D}]\setminus\{0\}$ put

$$N(a + b\sqrt{D}) = |a^2 - b^2 D|.$$

We must verify that this map satisfies conditions (i) and (ii) for a norm map.

(i) For $a+b\sqrt{D}, c+d\sqrt{D} \in \mathbb{Z}[\sqrt{D}] \setminus \{0\}$:

$$\begin{aligned} N[(a+b\sqrt{D})(c+d\sqrt{D})] &= N[(ac+bdD)+(ad+bc)\sqrt{D}] = \\ &= |(ac+bdD)^2-(ad+bc)^2D| = \\ &= |(a^2-b^2D)(c^2-d^2D)| = \\ &= N(a+b\sqrt{D})N(c+d\sqrt{D}), \end{aligned}$$

so the condition (i) is satisfied.

(ii) $a+b\sqrt{D}$ is a unit $\Rightarrow$
$\exists c+d\sqrt{D}\colon (a+b\sqrt{D})(c+d\sqrt{D})=1 \Rightarrow$
$\exists c+d\sqrt{D}\colon N(a+b\sqrt{D})N(c+d\sqrt{D})=N(1)=1 \Rightarrow$
$\exists c+d\sqrt{D}\colon N(a+b\sqrt{D})=1$ and $N(c+d\sqrt{D})=1 \Rightarrow$
$N(a+b\sqrt{D})=1$.
Conversely, if $N(a+b\sqrt{D})=1$, then $a^2-b^2D=\pm 1$, hence

$$\frac{1}{a^2-b^2D}\cdot(a-b\sqrt{D}) \in \mathbb{Z}[\sqrt{D}] \setminus \{0\}.$$

Since

$$\left[\frac{1}{a^2-b^2D}\cdot(a-b\sqrt{D})\right]\cdot(a+b\sqrt{D}) = \frac{a^2-b^2D}{a^2-b^2D} = 1,$$

$a+b\sqrt{D}$ is a unit.

Since N is a norm map, by previous proposition every nonunit element $a \in \mathbb{Z}[\sqrt{D}] \setminus \{0\}$ can be written as a product of irreducible elements. □

Corollary. *Every nonunit and nonzero element of $\mathbb{Z}[\sqrt{-2}]$ can be written as a product of irreducible elements. The only units in $\mathbb{Z}[\sqrt{-2}]$ are ± 1.*

Proof. We have proved that

$$N : \mathbb{Z}[\sqrt{-2}] \setminus \{0\} \to \mathbb{N}, \quad N(a+b\sqrt{-2}) = a^2+2b^2$$

is a norm map. Hence $a+b\sqrt{-2}$ is a unit $\Leftrightarrow a^2+2b^2=1 \Leftrightarrow a^2=1$ and $b^2=0 \Leftrightarrow a=\pm 1$ and $b=0$. □

The existence of the factorization into primes in $\mathbb{Z}[\sqrt{-2}]$ is proved, we want to prove also uniqueness.

Unique Factorization Domains, Euclidean Domains and Norm

Definition. Let R be an integral domain. We'll say that R is a *unique factorization domain* (UFD) if two conditions are fulfilled:

(i) every nonunit $a \in R \setminus \{0\}$ can be written as a product of irreducibles,

(ii) this factorization is unique in the sense that if $a = \pi_1 \cdot \ldots \cdot \pi_r = \tau_1 \cdot \ldots \cdot \tau_s$ are two such decompositions, then $r = s$ and after suitable permutation, $\pi_i \sim \tau_i$.

We have proved that condition (i) is fulfilled for all $\mathbb{Z}[\sqrt{D}]$ where $D \in \mathbb{Z} \setminus \{0, 1\}$ is square-free. But condition (ii) is fulfilled only sometimes:

- if $D < 0$ then (ii) holds true only for $D \in \{-1, -2\}$ but if $D \in \{-3, -7, -11, -19, -43, -67, -163\}$ it can be repaired if we take $\mathbb{Z}[\frac{1+\sqrt{D}}{2}]$ and these are all negative values when we obtain UFD in this way (but the proof of this assertion is very hard);
- if $D > 0$, the situation is different: it is conjectured that we can obtain UFD for infinitely many values of D but it is still an open problem.

Because the uniqueness of factorization in $\mathbb{Z}$ follows from the theorem on the division with remainder, we will prove the uniqueness of factorization in $\mathbb{Z}[\sqrt{-2}]$ in a similar way. First we must define a division with remainder generally in integral domains.

Definition. An integral domain R is a *Euclidean domain* if there is a map $\varphi\colon R \setminus \{0\} \to \mathbb{N}$ such that

$$\forall a \in R, b \in R \setminus \{0\}\colon \exists q, r \in R : a = bq + r \text{ where } r = 0 \text{ or } \varphi(r) < \varphi(b).$$

Proposition. $\mathbb{Z}[\sqrt{-2}]$ *is a Euclidean domain.*

Proof. We take the norm N as a map φ. For $a \in \mathbb{Z}[\sqrt{-2}], b \in \mathbb{Z}[\sqrt{-2}] \setminus \{0\}$ we consider $a/b = a\bar{b}/b\bar{b}$, where $\bar{b}$ is complex conjugation of b. Notice that $a\bar{b} \in \mathbb{Z}[\sqrt{-2}]$ and $b\bar{b} = N(b) \in \mathbb{N}$, so

$$\frac{a}{b} = \frac{a\bar{b}}{b\bar{b}} = c + d\sqrt{-2} \in \mathbb{Q}[\sqrt{-2}].$$

We choose $m, n \in \mathbb{Z}$ as close as possible to c and d so $|m - c| \leq \frac{1}{2}$ and $|n - d| \leq \frac{1}{2}$. Let $q = m + n\sqrt{-2}$, so we write $a = bq + r$ and $r = a - bq$. If $r \neq 0$, then

$$\begin{aligned}
N(r) &= N(a - bq) = \\
&= N[b(c + d\sqrt{-2}) - b(m + n\sqrt{-2})] = \\
&= N[b((c - m) + (d - n)\sqrt{-2})] = \\
&= b[(c - m) + (d - n)\sqrt{-2}] \cdot \overline{b[(c - m) + (d - n)\sqrt{-2}]} = \\
&= b[(c - m) + (d - n)\sqrt{-2}] \cdot \bar{b}[(c - m) - (d - n)\sqrt{-2}] = \\
&= b\bar{b}[(c - m)^2 + 2(d - n)^2] \leq N(b)\left(\frac{1}{4} + \frac{1}{2}\right) < \\
&< N(b),
\end{aligned}$$

hence $\mathbb{Z}[\sqrt{-2}]$ is a Euclidean domain. □

It is well-known fact that every Euclidean domain is UFD but it wasn't probably well-known for Fermat and his contemporaries so we will prove the uniqueness of the factorization in $\mathbb{Z}[\sqrt{-2}]$ in another way.

Proposition. *Let R be a Euclidean domain and $a, b \in R$ are relatively prime. Then there exist $x, y \in R$ such that $ax + by = 1$.*

Proof. If a or b is a unit, we can take $x = a^{-1}, y = 0$ or $x = 0, y = b^{-1}$. Otherwise $N(a) > 1, N(b) > 1$ and we'll divide with remainder. We want to obtain remainder with norm 1.
$a = bq_1 + r_1$, where $N(r_1) < N(b)$ ($r_1 \neq 0$ because otherwise $b \mid a$, which is cotradiction with an assumption that a, b are relatively prime). If $N(r_1) = 1$, we are lucky, otherwise
$b = r_1q_1 + r_2$, where $N(r_2) < N(r_1)$ ($r_2 \neq 0$ because otherwise $r_1 \mid b, r_1 \mid a$ and r_1 isn't a unit - cotradiction).
Continuing in this procedure we obtain after finite steps $r_k \in R$ with $N(r_k) = 1$, hence r_k is a unit. Working backwards we see that $r_k = ax_1 + by_1$ for some $x_1, y_1 \in R$, so

$$1 = ax_1r_k^{-1} + by_1r_k^{-1}$$

and the proposition is proved. □

Proposition. *Let R be a Euclidean domain, let $\pi \in R$ be irreducible and $a, b \in R$. If $\pi \mid ab$ then $\pi \mid a$ or $\pi \mid b$.*

Proof. If $\pi \mid a$ we are lucky, so suppose $\pi \mid ab$ and $\pi \nmid a$. If $r \mid \pi$ and $r \mid a$ for $r \in R$, then exist $c, d \in R$ such that $rc = \pi, rd = a$. If c is a unit, then $r = \pi c^{-1}$, so $\pi c^{-1}d = a$, then $\pi \mid a$ - contradiction. Because π is irreducible, $rc = \pi$ and c isn't a unit, r must be a unit. We supposed that $r \mid \pi, r \mid a$ and we proved that then r is a unit. Hence π and a are relatively prime and by previous proposition there exist $x, y \in R$ such that $\pi x + ay = 1$. Then

$$\pi bx + aby = b.$$

Since $\pi \mid ab$, then $\pi \mid (\pi bx + aby) = b$ which we wanted to prove. □

Theorem. *Let R be a Euclidean domain with a norm map. Then R is UFD.*

Proof. We have already proved that in every integral domain with a norm map every nonzero and nonunit element can be written as a product of irreducible elements. Hence it suffice to prove the uniqueness. Suppose we have $a \in R$ that has two factorizations into irreducibles:

$$a = \pi_1 \cdot \ldots \cdot \pi_r = \tau_1 \cdot \ldots \cdot \tau_s, \text{ where } r \geq s.$$

Then $\tau_1 \mid \pi_1 \cdot \ldots \cdot \pi_r$ and by previous proposition, $\tau_1 \mid \pi_i$ for some i and since both are irreducible, they must be associate. Without loss of generality we can let π_i be π_1, so there is a unit $u_1 \in R$ such that $\tau_1 = \pi_1 u_1$. Then

$$\pi_1\pi_2 \cdot \ldots \cdot \pi_r = \pi_1 u_1 \tau_2 \cdot \ldots \cdot \tau_s \Rightarrow \pi_1(\pi_2 \cdot \ldots \cdot \pi_r - u_1\tau_2 \cdot \ldots \cdot \tau_s) = 0 \Rightarrow$$
$$\pi_2 \cdot \ldots \cdot \pi_r - u_1\tau_2 \cdot \ldots \cdot \tau_s = 0 \Rightarrow \pi_2 \cdot \ldots \cdot \pi_r = u_1\tau_2 \cdot \ldots \cdot \tau_s.$$

Following the same process we can pair up $u_1\tau_2$ with its associate and we can continue to do this until we have paired up each of the irreducible factors τ_i with associate π_j. If $r > s$, we obtain $\pi_{s+1} \cdot \ldots \cdot \pi_r = 1$ which is impossible because π_j aren't units. So $r = s$ and the theorem is proved. □

Let us mention that even a stronger theorem is true: every Euclidean domain is UFD, but proof isn't so short and needs ideals which weren't discovered at the time of Fermat.

Corollary. $\mathbb{Z}[\sqrt{-2}]$ *is UFD.*

Finishing the Solution of $y^2 + 2 = x^3$

We can now continue to solve the equation $y^2 + 2 = x^3$ which is equivalent to $(y + \sqrt{-2})(y - \sqrt{-2}) = x^3$. If y is even then x is also, but then $x^3 \equiv 0 \pmod 4$ whereas $y^2 + 2 \equiv 2 \pmod 4$. So y and x are both odd.

If $r \mid y+\sqrt{-2}$ and $r \mid y-\sqrt{-2}$ then $r \mid [(y+\sqrt{-2}) - (y-\sqrt{-2})] = 2\sqrt{-2}$ so $N(r) \mid N(2\sqrt{-2}) = 8$. But $N(r) \mid N(y + \sqrt{-2}) = y^2 + 2$ which is odd. It is possible only if $N(r) = 1$, so r is a unit and then $y + \sqrt{-2}$ and $y - \sqrt{-2}$ are relatively prime.

Since there is factorization into irreducibles in $\mathbb{Z}[\sqrt{-2}]$ and $x \notin \{0, \pm 1\}$, then $x = \pi_1 \cdot \ldots \cdot \pi_r$ ($\pi_1, \ldots, \pi_r$ are irreducibles in $\mathbb{Z}[\sqrt{-2}]$) and

$$x^3 = \pi_1^3 \cdot \ldots \cdot \pi_r^3 = (y + \sqrt{-2})(y - \sqrt{-2}).$$

Since $\mathbb{Z}[\sqrt{-2}]$ is UFD, then there exist mutually different $i_1, \ldots, i_c \in \{1, \ldots, s\}$ such that

$$y + \sqrt{-2} = u \cdot \pi_{i_1}^3 \cdot \ldots \cdot \pi_{i_c}^3,$$

where u is a unit of $\mathbb{Z}[\sqrt{-2}]$, because otherwise $y + \sqrt{-2}$ and $y - \sqrt{-2}$ couldn't be relatively prime. As the only units of $\mathbb{Z}[\sqrt{-2}]$ are ± 1, which are both cubes, then

$$y + \sqrt{-2} = (\pm\pi_{i_1} \cdot \ldots \cdot \pi_{i_c})^3 = (a + b\sqrt{-2})^3$$

for some $a, b \in \mathbb{Z}$. We proved at the beginning of this lecture that then we have $y = \pm 5$ and $x = 3$. We have proved that the only solutions of the equation $y^2 + 2 = x^3$ in integers are $(3, \pm 5)$.