PELL'S EQUATION AND UNITS IN $\mathbb{Z}[\sqrt{D}]$

DEFINITION: The equation $x^2 - Dy^2 = 1$ for some fixed positive integer D that is not a perfect square, where the variables x, y range through integers is called a **Pell's equation**. We say that a solution (x_0, y_0) is a **positive solution** if x_0, y_0 are both positive integers. We say that one positive solution (x_0, y_0) is **smaller** than another positive solution (x_1, y_1) if $x_0 < x_1$; equivalently, $y_0 < y_1$.

(1) Warmup with Pell's equation:
  (a) Verify that $(9, 4)$ is a solution to Pell's equation with $D = 5$.
  (b) Fix some D. Show that if (x_0, y_0) is a solution to Pell's equation, then $(\pm x_0, \pm y_0)$ are solutions to Pell's equation with the same D.
  (c) What two trivial solutions does every Pell's equation have?
  (d) Explain how to recover all solutions from just the positive solutions.

(a) $9^2 - 5 \cdot 4^2 = 81 - 5 \cdot 16 = 1\checkmark$.
(b) $(\pm x_0)^2 - D(\pm y_0)^2 = x_0^2 - Dy_0^2 = 1$.
(c) $(\pm 1, 0)$.
(d) By throwing in $(\pm 1, 0)$ and taking $\pm$ each coordinate.

(2) By trial and error find the smallest positive solutions to Pell's equation with $D = 2$, $D = 3$, and $D = 5$.

For $D = 2$ we find $(3, 2)$. For $D = 3$ we find $(2, 1)$, For $D = 5$ we find $(9, 4)$.

(3) Suppose that D is a perfect square. Show that the equation $x^2 - Dy^2 = 1$ has no positive solutions.

If $D = d^2$ with $d > 0$, then $x^2 - Dy^2 = (x - dy)(x + dy)$. For any positive integers x, y, we have $x + dy > 1$, and $x - dy \in \mathbb{Z}$, so the product cannot be 1.

DEFINITION: Let D be a positive integer that is not a perfect square. We define the **quadratic ring** of D to be

$$\mathbb{Z}[\sqrt{D}] := \{a + b\sqrt{D} \mid a, b \in \mathbb{Z}\} \subseteq \mathbb{R}.$$

DEFINITION: For the quadratic ring $\mathbb{Z}[\sqrt{D}]$ we define the **norm** function

$$N : \mathbb{Z}[\sqrt{D}] \to \mathbb{Z} \qquad N(a + b\sqrt{D}) = a^2 - b^2 D.$$

Note that $N(a + b\sqrt{D}) = (a + b\sqrt{D})(a - b\sqrt{D})$.

LEMMA: For the quadratic ring $\mathbb{Z}[\sqrt{D}]$ the norm function satisfies the multiplicative property $N(\alpha\beta) = N(\alpha)N(\beta)$.

(4) Warmup with $\mathbb{Z}[\sqrt{D}]$:
  (a) Show[1] that $\mathbb{Z}[\sqrt{D}]$ is a ring.
  (b) Show that every element in $\mathbb{Z}[\sqrt{D}]$ has a unique expression in the form $a+b\sqrt{D}$.

(a) We check the conditions for a subring: Let $a+b\sqrt{D}, c+d\sqrt{D} \in \mathbb{Z}[\sqrt{D}]$. Then,
- $1 = 1+0\sqrt{D} \in \mathbb{Z}[\sqrt{D}]$
- $(a+b\sqrt{D})-(c+d\sqrt{D}) = (a-c)+(b-d)\sqrt{D} \in \mathbb{Z}[\sqrt{D}]$, and
- $(a+b\sqrt{D})(c+d\sqrt{D}) = (ac+bdD)+(ad+bc)\sqrt{D} \in \mathbb{Z}[\sqrt{D}]$.

(b) If $a+b\sqrt{D} = c+d\sqrt{D}$ and $(a,b)\neq(c,d)$, then $a-c=(d-b)\sqrt{D}$. If $a\neq c$, then we must have $b\neq d$, so either way, $b\neq d$. Then $\sqrt{D}=\frac{a-c}{d-b}$, which contradicts that $\sqrt{D}$ is irrational. Thus, $a+b\sqrt{D}=c+d\sqrt{D}$ implies $(a,b)=(c,d)$.

(5) Norms, units, and Pell's equation:
  (a) Prove the Lemma above.
  (b) Show that an element of $\mathbb{Z}[\sqrt{D}]$ is a unit (has a multiplicative inverse) if and only if its norm is ±1.
  (c) Show that the set of units of $\mathbb{Z}[\sqrt{D}]$ forms a group under multiplication.
  (d) Show that the set of elements $a+b\sqrt{D}\in\mathbb{Z}[\sqrt{D}]$ such that (a,b) is a solution to the Pell's equation $x^2-Dy^2=1$ forms a group under multiplication.

(a) Set $\alpha=a+b\sqrt{D}$, $\beta=c+d\sqrt{D}$. Then $\alpha\beta=(ac+bdD)+(ad+bc)\sqrt{D}$ so

$$\begin{aligned}N(\alpha\beta) &= (ac+bdD)^2-(ad+bc)^2D\\ &= a^2c^2+2abcdD+b^2d^2D^2-a^2+d^2D-2abcdD-b^2c^2D\\ &= a^2c^2+b^2d^2D^2-a^2d^2D-b^2c^2D.\end{aligned}$$

On the other hand,

$$N(\alpha)N(\beta)=(a^2-b^2D)(c^2-d^2D)=a^2c^2-a^2d^2D-b^2c^2D+b^2d^2D^2.$$

(b) If α is a unit so $\alpha\beta=1$ for some β, then

$$1=N(1)=N(\alpha\beta)=N(\alpha)N(\beta),$$

so $N(\alpha)$ is a unit in $\mathbb{Z}$, hence is ±1. Conversely, if $\alpha=a+b\sqrt{D}$ and $N(\alpha)=\pm1$, then $(a+b\sqrt{D})(a-b\sqrt{D})=\pm1$, so $(a+b\sqrt{D})(\pm(a-b\sqrt{D}))=1$, and α is a unit.

(c) The product of two elements of norm 1 has norm 1, by the lemma. The element 1 has norm 1, which serves as the identity. By the previous part, an element of norm 1 has an inverse, which must have norm 1 by the lemma.

THEOREM: Let D be a positive integer that is not a perfect square. Consider the Pell's equation $x^2-Dy^2=1$. Let (a,b) be the smallest positive solution (assuming that some positive solution exists). Then every positive solution (c,d) can be obtained by the rule

$$c+d\sqrt{D}=(a+b\sqrt{D})^k$$

[1]Recall: to check that a subset of a ring is a subring, it suffices to show that it contains the multiplicative identity and is closed under subtraction and multiplication.