Section 20. $x^2 - Ny^2 = 1$

Note. In this section, we discuss more general Diophantine equations. We give a bit of history, and then focus on Pell's equation $x^2 - Ny^2 = 1$.

Note/Definition. We take a general *Diophantine equation* as a polynomial equation with integer coefficients and a finite number of unknowns. A *solution* is a collection of integer values for the variables which make the equation true. In Section 3. Linear Diophantine Equations we considered first degree polynomial equations in two variables, and classified the solutions in Theorem 3.1. In Section 18. Sums of Two Squares we considered Diophantine equations of the form $x^2 + y^2 = n$ and classified when solutions exist in Theorem 18.1. In Section 19. Sums of Four Squares we considered equations of the form $x^2 + y^2 + z^2 + w^2 = n$ and showed that a solution exists for all $n \geq 0$ in Lagrange's Four-Square Theorem (Theorem 19.1).

Note. In 1900 at the International Congress of Mathematicians in Paris, David Hilbert (January 23, 1862–February 14, 1943) stated 10 unsolved problems that should gain the attention of mathematicians in the 20th century. The list later grew to 23 problems and was published in *The Bulletin of the American Mathematical Society* in 1902. The tenth-problem concerns solutions to Diophantine equations:

> Given a Diophantine equation with any number of unknown quantities and with rational integral numerical coefficients: To devise a process according to which it can be determined in a finite number of operations whether the equation is solvable in rational integers.

In 1970, Yuri Matiyasevich proved that no such process exists. His proof is based on earlier work of Martin Davis, Hilary Putnam, and (crucially) Julia Robinson. The result is variously called Matiyasevich's Theorem or the Matiyasevich-Robinson-Davis-Putnam Theorem (or simply the MRDP Theorem). The reference for Matiyasevich's main paper is: "Enumerable Sets are Diophantine," *Dokl. Adad. Nauk SSSR*, **191** 279–282 (1970, in Russian), and *Soviet Math. Doklady*, **11**, 354–357 (1970, English translation). A readable explanation of the proof appears in Martin Davis' "Hilbert's Tenth Problem is Unsolvable," *The American Mathematical Monthly*, **80**(3), 233–269 (1973); a copy is available online on the MAA website. An example of a general result concerning this type of problem is the following.

Theorem 20.A. Let

$$F(x, y) = a_n x^n + a_{n-1} x^{n-1} y + a_{n-2} x^{n-2} y^2 + \cdots + a_0 y^n,$$

and suppose that $F(x, 1) = 0$ has no repeated roots. Then the equation $F(x, y) = c$, where c is an integer, has only finitely many solutions if $n \geq 3$.

Note. Dudley does not give a reference for Theorem 20.A. Yann Bugeaud mentions the result in his "On Some Results of Alan Baker," *Resonance*, 735–748 (July 2018), available online on the Indian Academy of Sciences website (notice Theorem 2.5 in this paper is our Theorem 20.A; accessed 3/28/2022), where he credits it to "Feldman, 1971." The result seems to follow from Feldman's much more general result which, in turn, builds on work of Gelfond, Baker, and Thue. The reference for Felman's paper is: N. I. Feldman, "An Effective Power Sharpening of a Theorem

of Liouville," *Izv. Akad. SSSR, Ser. Math.*, **35**, 973–990 (1971; in Russian). For a discussion on solving Diophantine equations and another reference to Feldman's work, see R. J. Stroeker's "How to Solve a Diophantine Equation: A Number-Theoretic Excursion," *The American Mathematical Monthly*, **91**(7) 385–392 (1984); this is available through JSTOR.

Note. We now concentrate on Diophantine equations of two variables of the form $ax^n+by^n = c$. We see by Theorem 20.A that such equations have only finitely many solutions for $n \geq 3$. In the case $n = 1$, we saw when solutions do not exist and when they do (in which case there were infinitely many, and we found these solutions) in Theorem 3.1. The general case for $n = 2$ is "too complicated for us to treat here" (Dudley, page 156). We saw a special case in Section 16. Pythagorean Triangles. We now consider another special case, equations of the form $x^2 - Ny^2 = 1$. We'll see in Theorem 20.1 that when such an equation has a solution, it has infinitely many solutions.

Note. A solution of $x^2 - Ny^2 = 1$ can be used to get a rational approximation of $\sqrt{N}$. Since the graph of $x^2 - Ny^2 = 1$ is a hyperbola centered at $(0, 0)$ with asymptotes $y = \pm x/\sqrt{N}$, then for x "large", we have $y \approx x/\sqrt{N}$ or $\sqrt{N} \approx x/y$. Indian astronomer and mathematician Brahmagupta (598–670) extensively studied the equation in his *Brāhmasphuṭasiddhānta*. The computation of solutions requires the manipulation of large numbers. Others who made additional contributions (also in India) are Bhaskara II (1114–1185) and Narayana (1340–1400). These Indian

results were unknown in Europe, when Fermat took an interest in the problem in 1657. In writing on the equation, Leonhard Euler mistakenly attributed some work to John Pell (which was actually results of another, William Brouckner), and the equation became known as Pell's Equation (or sometimes the Pell-Fermat Equation); Dudley prefers the term "Fermat's Equation." These historical notes (and the few that follow) are based on the MacTutor History of Mathematics Archive webpage on Pell's Equation and the Wikipedia page on Pell's Equation.

Note. One way to solve $x^2 - Ny^2 = 1$ (for a given N) is to make a table of values of $1 + Ny^2$ and look for a square. For example, to solve $x^2 - 2y^2 = 1$, we consider:

y	$1 + 2y^2$
1	3
2	9

So a solution is given by $x = 3$ and $y = 2$. For $x^2 - 3y^2 = 1$, we consider:

y	$1 + 3y^2$
1	4
2	13
3	28
4	49

So two solutions are $x = 2$ and $y = 1$, and $x = 7$ and $y = 4$. The ease of these two examples is misleading! With $n = 13$, the smallest solution is $x = 649$ and $y = 180$; with $n = 61$, the smallest solutions is $x = 1{,}766{,}319{,}049$ and $y = 226{,}153{,}980$ (in fact, Bhaskara II (1114–1185) found this solution for $n = 61$).

Note. Notice that for $N \leq 0$, there are unique solutions, mostly involving $y = 0$ and $x = \pm 1$. So we will only consider positive N. Notice that if N itself is a square, say $N = m^2$, then the Pell equation becomes $1 = x^2 - Ny^2 = x^2 - m^2n^2 = (x - my)(x + my)$. So we must have that both $(x - my)$ and $(x + my)$ are $+1$ or are -1, from which we find that $x = \pm 1$ and $y = 0$. So we only address $N > 0$ and N not a square.

Note 20.B. With $N > 0$ and N not a square, there is always a solution to $x^2 - Ny^2 = 1$, other than the solution $x = \pm 1$ and $y = 0$. This was first shown by Joseph-Louis Lagrange (January 25, 1736–April 10, 1813) in 1766. This appears in his "Solution d'un Problème d'Arithmétique," in Joseph Alfred Serret (ed.), *Œuvres de Lagrange*, volume 1, pp. 671-731 (1867), and in his *Additions to Euler's Elements of Algebra* (1771). We take this result as true, without attempting a proof.

Definition. An irrational number of the form $\alpha = r + s\sqrt{n}$, where r and s are integers, it gives a solution of the Pell equation $x^2 - Ny^2 = 1$ if $r^2 - Ns^2 = 1$.

Note. An example of an irrational number that gives a solution to $x^2 - 2y^2 = 1$ is $\alpha = r + s\sqrt{2} = 3 + 2\sqrt{2}$ since $r^2 - Ns^2 = (3)^2 - 2(2)^2 = 1$. Also, for $x^2 - 7y^2 = 1$, $\alpha = r + s\sqrt{7} = 8 + 3\sqrt{7}$ gives a solution since $r^2 - Ns^2 = (8)^2 - 7(3)^2 = 1$. The choice of the terminology "gives a solution" will be made a bit more clear in our main theorem.

Note 20.B. The square root of any positive integer that is not a perfect square, is irrational. This is Proposition 9 in Book X of Euclid's *Elements of Geometry*; see my online notes for Introduction to Modern Geometry (MATH 4157/5157) on Section 2.4. Books VII and IX. Number Theory (notice that last note). The proof that $\sqrt{2}$ is irrational is in my online notes for Mathematical Reasoning (MATH 3000) on Section 6.3. Divisibility: The Fundamental Theorem of Arithmetic in Theorem 6.31. This proof can be modified to also show that any nonsquare has an irrational square root. We need this fact in the proof of our first lemma.

Lemma 20.1. If $N > 0$ is not a square, then $x + y\sqrt{N} = r + s\sqrt{N}$ if and only if $x = r$ and $y = s$.

Note. The next lemma can be verified by simply multiplying out both sides of the claimed equation.

Lemma 20.2. For any integers a, b, c, d, N,

$$(a^2 - Nb^2)(c^2 - Nd^2) = (ac + Nbd)^2 - N(ad + bc)^2.$$

Lemma 20.3. If α gives a solution of $x^2 - Ny^2 = 1$, then so does $1/\alpha$.

Lemma 20.4. Let α and β give solutions of $x^2 - Ny^2 = 1$, then so does $\alpha\beta$.

Exercise 20.3. Two solutions of $x^2 - 8y^2 = 1$ are $(x, y) = (3, 1)$ and $(17, 6)$. Apply Lemma 20.4 to find another.

Solution. We can use the solution to find irrational numbers that "give a solution." We take $\alpha = (3) - (1)\sqrt{8} = 3 - \sqrt{8}$ and $\beta = (17) - (6)\sqrt{8} = 17 - 6\sqrt{8}$. We then have that

$$\alpha\beta = (3-\sqrt{8})(17-6\sqrt{8}) = ((3)(17)+(8)(1)(6))+((3)(6)+(1)(17))\sqrt{8} = 99+35\sqrt{8}$$

gives a solution by Lemma 20.4. So, by the definition of "gives a solution" we have that $\boxed{(x, y) = (99, 35)}$ is a solution to $x^2 - 8y^2 = 1$.

Lemma 20.5. If α gives a solution of $x^2 - Ny^2 = 1$, then so does α^k for any integer k, positive, negative, or zero.

Lemma 20.6. Suppose that a, b, c, d are nonnegative and that $\alpha = a + b\sqrt{N}$ and $\beta = c + d\sqrt{N}$ give solutions of $x^2 - Ny^2 = 1$. Then $\alpha < \beta$ if and only if $a < c$.

Note. Fix N and consider the equation $x^2 - Ny^2 = 1$. Consider the set of all real numbers that give a solution to the equation, $G = \{r + s\sqrt{N} \mid r^2 - ns^2 = 1\}$. We know by Lemma 20.6 that the elements of this set can be ordered based on the size of the values r. So we can use this ordering to find a smallest element of the set G (it will always be $1 + 0\sqrt{N}$), a second smallest element, and so forth. It is the second smallest element that we are interested in (notice that this element corresponds to the smallest *nontrivial* solution).

Definition. For equation $x^2 - Ny^2 = 1$, and the set $G = \{r + s\sqrt{N} \mid r^2 - ns^2 = 1\}$ of all real number that give a solution, the smallest number θ in set G greater than one is a *generator* for $x^2 - Ny^2 = 1$.

Note. Notice that by Note 20.A, a generator always exists (provided N is not a square). Our main theorem explains the sense in which θ is a generator of solutions. In turn, this is also how $\alpha = r + s\sqrt{N}$ "gives solutions."

Theorem 20.1. If θ is the generator for $x^n - Ny^2 = 1$, then all nontrivial solutions of the equation with x and y positive are given by θ^k, $k = 1, 2, \ldots$. That is, if $x = r$, $y = s$ is a solution then $\alpha = r + s\sqrt{N}$ is some positive power of θ.

Revised: 3/29/2022