Math 4527 (Number Theory 2) Lecture #9 of 38 \sim February 8, 2021

Pell's Equation (Part 3): The Super Magic Box

- The Super Magic Box
- **•** Factoring Via Continued Fractions

This material represents §6.3.3 from the course notes.

The Super Magic Box, I

Theorem (Pell's Equation, Part 1)

Let D be a positive squarefree integer. Then the following hold:

- 1. Let r be an integer with $r^2 < D$. If x and y are positive integers with $x^2 - Dy^2 = r$, then x/y is a continued fraction megers with $x -$
convergent to \sqrt{D} .
- 2. The equation $x^2 Dy^2 = 1$ always has a nontrivial integer solution.
- 3. The ring Z[√ \sqrt{D}] has a well-defined fundamental unit $u = x_1 + y_1 \sqrt{D}$. Furthermore, if w is an arbitrary unit in $\mathbb{Z}[\sqrt{D}]$, then $w=\pm u^{n}$ for some integer n (possibly negative).
- 4. If $u = x_1 + y_1$ \sqrt{D} is the fundamental unit in $\mathbb{Z}[\sqrt{D}]$ D is the fundamental unit in $\mathbb{Z}[\sqrt{D}]$, then if we define $x_n + y_n\sqrt{D} = (x_1 + y_1\sqrt{D})^n$ for nonnegative integers n, then (x_n, y_n) is a solution to $x^2 - Dy^2 = \pm 1$, and these are all of the solutions up to changing the signs of x_n or y_n .

The Super Magic Box, II

Theorem (Pell's Equation, Part 2)

Let $D > 0$ be squarefree, with $\sqrt{D} = [a_0, \ldots, a_n, \alpha_{n+1}]$, and take $p_n/q_n = [a_0, a_1, \ldots, a_n]$ to be the nth convergent. Define the sequences A_n and C_n by setting $A_0 = 0$ and $C_0 = 1$, and for $n \ge 1$ set $A_{n+1} = a_n C_n - A_n$ and $C_{n+1} = (D - A_{n+1}^2)/C_n$.

- 5. The continued fraction expansion of \sqrt{D} is periodic and of the form $[a_0, a_1, a_2, \cdots, a_{k-1}, 2a_0]$ with $a_0 = [\sqrt{D}]$.
- 6. The sequences A_n and C_n are integer-valued, $\alpha_n = (A_n + \sqrt{D})/C_n$, $p_n p_{n-1} - Dq_n q_{n-1} = (-1)^n A_{n+1}$ and $p_n^2 - Dq_n^2 = (-1)^{n+1}C_{n+1}$.
- 7. With notation as in (5), the fundamental unit of $\mathbb{Z}[\sqrt{2}]$ on as in (5), the fundamental unit of $\mathbb{Z}[\sqrt{D}]$ is $p_{k-1} + q_{k-1} \sqrt{D}$. Its norm is -1 when k is odd and its norm is $+1$ when k is even.

Last time I proved items (1)-(5). Today I will go through the morass of algebra for (6) and then finish off (7).

- The purpose here is to go through the algebra that justifies how the super magic box works, so that we can then do a bunch of examples of the magic box calculations.
- Just to emphasize, the purpose of all of this is to give a computationally efficient procedure for computing the computationally efficient procedure for computing the continued fraction expansion of \sqrt{D} by simplifying the calculations that are required. (It is sort of like converting polynomial long division into synthetic division.)

The Super Magic Box, IV

6a. Define A_n and C_n via $A_0 = 0$ and $C_0 = 1$, and for $n \ge 1$ set $A_{n+1} = a_n C_n - A_n$ and $C_{n+1} = (D - A_{n+1}^2)/C_n$. Then A_n and C_n are integer-valued.

Proof:

- We induct on n. The base case $n = 0$ is trivial.
- For the inductive step, clearly A_{n+1} is an integer.
- For C_{n+1} , plugging in for A_{n+1} and expanding yields $C_{n+1} = \frac{D - (a_n C_n - A_n)^2}{C}$ $\frac{(C_n - A_n)^2}{C_n} = (2a_nA_n - a_n^2C_n) + \frac{D - A_n^2}{C_n}$ $\frac{n}{C_n}$ and the fraction at the end is simply C_{n-1} .
- Thus A_{n+1} and C_{n+1} are integers.

Note that we also have shown that $C_{n+1} = 2a_n A_n - a_n^2 C_n + C_{n-1}$. We will use this later.

The Super Magic Box, V

6b. Define A_n and C_n via $A_0 = 0$ and $C_0 = 1$, and for $n \ge 1$ set $A_{n+1} = a_n C_n - A_n$ and $C_{n+1} = (D - A_{n+1}^2)/C_n$. Then $\alpha_n = (A_n + \sqrt{D})/C_n$.

Proof:

We induct on n . The base case $n=0$ is $\alpha_0=$ √ $\overline{D} = \frac{0 + \sqrt{D}}{1}$ $\frac{1}{1}$. For the inductive step, suppose $\alpha_n = \frac{A_n + \sqrt{D}}{C_n}$ √ $\frac{1}{C_n}$. Then $\alpha_{n+1} = \frac{1}{n}$ $\frac{1}{\alpha_n - a_n} = \frac{C_n}{-A_{n+1}}$ $-A_{n+1} +$ √ D $=\frac{A_{n+1}+A_{n+1}+B_{n+1}+$ √ D $\frac{(D - A_{n+1}^2)/C_n}{(D - A_{n+1}^2)/C_n}$ $A_{n+1} +$ √ D $\frac{C_{1} + C_{2}}{C_{n+1}}$ as claimed.

6cd. Define A_n and C_n via $A_0 = 0$ and $C_0 = 1$, and for $n \ge 1$ set $A_{n+1} = a_n C_n - A_n$ and $C_{n+1} = (D - A_{n+1}^2)/C_n$. Then $p_n p_{n-1} - Dq_n q_{n-1} = (-1)^n A_{n+1}$ and $p_n^2 - Dq_n^2 = (-1)^{n+1} C_{n+1}.$

Proof:

- We induct on n. The base cases $n = 0$ and $n = 1$ are straightforward calculations.
- For the inductive step, suppose that $p_np_{n-1} - Dq_nq_{n-1} = (-1)^n A_{n+1}, p_n^2 - Dq_n^2 = (-1)^{n+1} C_{n+1}.$
- Recall that $p_{n+1} = a_{n+1}p_n + p_{n-1}$ and $q_{n+1} = a_{n+1}q_n + q_{n-1}$.

The Super Magic Box, VII

6c. Define A_n and C_n via $A_0 = 0$ and $C_0 = 1$, and for $n \ge 1$ set $A_{n+1} = a_n C_n - A_n$ and $C_{n+1} = (D - A_{n+1}^2)/C_n$. Then $p_n p_{n-1} - Dq_n q_{n-1} = (-1)^n A_{n+1}$ and $p_n^2 - Dq_n^2 = (-1)^{n+1} C_{n+1}.$

Proof (continued):

• We have
$$
p_n p_{n-1} - Dq_n q_{n-1} = (-1)^n A_{n+1}
$$
,
\n $p_n^2 - Dq_n^2 = (-1)^{n+1} C_{n+1}$, $p_{n+1} = a_{n+1} p_n + p_{n-1}$ and
\n $q_{n+1} = a_{n+1} q_n + q_{n-1}$. Then

$$
p_{n+1}p_n - Dq_{n+1}q_n = (a_{n+1}p_n + p_{n-1})p_n - D(a_{n+1}q_n + q_{n-1})(q_n)
$$

= $a_{n+1}(p_n^2 - Dq_n^2) + (p_np_{n-1} - Dq_nq_{n-1})$
= $a_{n+1}(-1)^{n+1}C_{n+1} + (-1)^nA_{n+1}$
= $(-1)^{n+1}A_{n+2}$

The Super Magic Box, VIII

6d. Define A_n and C_n via $A_0 = 0$ and $C_0 = 1$, and for $n \ge 1$ set $A_{n+1} = a_n C_n - A_n$ and $C_{n+1} = (D - A_{n+1}^2)/C_n$. Then $p_n p_{n-1} - Dq_n q_{n-1} = (-1)^n A_{n+1}$ and $p_n^2 - Dq_n^2 = (-1)^{n+1} C_{n+1}.$

Proof (continued):

• We have
$$
p_n p_{n-1} - Dq_n q_{n-1} = (-1)^n A_{n+1}
$$
,
\n $p_n^2 - Dq_n^2 = (-1)^{n+1} C_{n+1}$, $C_{n+1} = 2a_n A_n - a_n^2 C_n + C_{n-1}$,
\n $p_{n+1} = a_{n+1} p_n + p_{n-1}$ and $q_{n+1} = a_{n+1} q_n + q_{n-1}$. Then

$$
p_{n+1}^2 - Dq_{n+1}^2 = (a_{n+1}p_n + p_{n-1})^2 - D(a_{n+1}q_n + q_{n-1})^2
$$

= $a_{n+1}^2(p_n^2 - Dq_n^2) + 2a_{n+1}(p_n p_{n-1} - Dq_n q_{n-1}) + (p_{n-1}^2 - Dq_{n-1}^2)$
= $a_{n+1}^2(-1)^{n+1}C_{n+1} + 2a_{n+1}(-1)^nA_{n+1} + (-1)^nC_n$
= $(-1)^{n+1}[a_{n+1}^2C_{n+1} - 2a_{n+1}A_{n+1} - C_n]$
= $(-1)^{n+2}C_{n+2}$

The Super Magic Box, IX

7. If $\sqrt{D} = [a_0, \overline{a_1, a_2, \ldots, a_{k-1}, 2a_0}]$ and $p_{k-1}/q_{k-1} = [a_0, a_1, \ldots, a_{k-1}]$, then the fundamental unit of $\mathbb{Z}[\sqrt{D}]$ is $p_{k-1}+q_{k-1}\sqrt{D}.$ Its norm is -1 when k is odd and its norm is $+1$ when k is even.

Proof:

- $\overset{\text{\tiny{L+1}}}{\text{Suppose that }} \sqrt{D} = [a_0, \overline{a_1, a_2, \ldots, a_{k-1}, 2a_0}].$
- Then since the expansion is periodic, we have $a_0 + \sqrt{D} = [2a_0, a_1, \ldots, a_{k-1}, a_0 + \sqrt{D}],$ so $\alpha_{k+1} =$ √ $D - a_0$. By (6b), this means $\frac{A_k + A_k}{C_k}$ √ D $\frac{1}{C_k} = -a_0 +$ √ D, and so since √ D is irrational the only way this can occur is when $\mathcal{C}_k = 1.$ Then by (6d), $p_{k-1}^2 - Dq_{k-1}^2 = (-1)^k C_k = (-1)^k$. Thus, $p_{k-1} + q_{k-1}$ \sqrt{D} is a unit in $\mathbb{Z}[\sqrt{D}]$ and its norm is $(-1)^k$.

7. If
$$
\sqrt{D} = [a_0, \overline{a_1, a_2, \dots, a_{k-1}, 2a_0}]
$$
 and
\n $p_{k-1}/q_{k-1} = [a_0, a_1, \dots, a_{k-1}],$ then the fundamental unit of
\n $\mathbb{Z}[\sqrt{D}]$ is $p_{k-1} + q_{k-1}\sqrt{D}$. Its norm is +1 when *r* is even and
\nits norm is -1 when *r* is odd.

Proof (continued):

- Conversely, suppose that p_n+q_n \sqrt{D} is a unit in $\mathbb{Z}[\sqrt{D}]$ D] so that $p_n^2 - Dq_n^2 = \pm 1$.
- By (1), p_n/q_n is a convergent to \sqrt{D} .
- Then by (6), we have $\rho_n^2 Dq_n^2 = (-1)^{n+1} \textit{\textsf{C}}_{n+1}$ and so we must have $C_{n+1} = 1$ and $(-1)^{n+1}$ equal to the norm of $p_n + q_n \sqrt{D}$.

7. If
$$
\sqrt{D} = [a_0, \overline{a_1, a_2, \dots, a_{k-1}, 2a_0}]
$$
 and
\n $p_{k-1}/q_{k-1} = [a_0, a_1, \dots, a_{k-1}],$ then the fundamental unit of
\n $\mathbb{Z}[\sqrt{D}]$ is $p_{k-1} + q_{k-1}\sqrt{D}$. Its norm is +1 when *r* is even and
\nits norm is -1 when *r* is odd.

Proof (finally):

- But if $C_{n+1} = 1$, since all remainders are between 0 and 1, we But if $C_{n+1} = 1$, since an remainders are between 0 and 1
must have $\alpha_{n+1} = \sqrt{D} - \lfloor \sqrt{D} \rfloor = \alpha_0$. By periodicity, the only way this can occur is if $n+1$ is a multiple of k.
- The fundamental unit corresponds to the smallest possible value of n, which (per the calculation above) is $n = k - 1$.
- Thus, the fundamental unit of $\mathbb{Z}[\sqrt{2}]$ ndamental unit of $\mathbb{Z}[\sqrt{D}]$ is indeed $\rho_{k-1} + q_{k-1}\sqrt{D}$ as claimed, and its norm is $(-1)^k$ as calculated earlier.

Pell's Equation Continued Fractions, XXIV

We have reduced the seemingly quite difficult problem of solving Pell's equation $x^2 - Dy^2 = \pm 1$ to the very approachable problem Fen's equation $x^2 - Dy^2 = \pm 1$ to the very approachable of computing the continued fraction expansion of \sqrt{D} .

- We can organize these calculations quite a bit more efficiently using the sequences A_n and C_n from (6).
- The point is that these sequences automatically encode the remainder term, because $\alpha_n = (A_n + \sqrt{D})/C_n$.
- \bullet Thus, what we can do is just compute these sequences A_n and C_n recursively: starting with $A_0 = 0$ and $C_0 = 1$, for $n \ge 1$ we set $A_{n+1} = a_n C_n - A_n$ and $C_{n+1} = (D - A_{n+1}^2)/C_n$. √
- Furthermore, we have $a_n = \lfloor \alpha_n \rfloor = \lfloor (A_n +$ $D)/C_n$.
- Once we have the terms a_n from the continued fraction expansion, we can then compute the convergent terms p_n and q_n using the magic box procedure from a few lectures ago.

The Super Magic Box, XI

If we combine all of these things, we get the "super magic box":

- The rows in the table are A_n , C_n , a_n , p_n , q_n , and $p_n^2 Dq_n^2$.
- We compute the sequences a_n , A_n , C_n via the recurrences $A_{n+1} = a_n C_n - A_n$ $C_{n+1} = (D - A_{n+1}^2)/C_n$ $a_{n+1} = \frac{A_{n+1} + a_0}{C_{n+1}}$ with initial conditions $A_0=0,~\mathcal{C}_0=1,$ and $a_0=\lfloor$ √ D .
- Once we reach a term with $C_k = 1$ we stop, since we will have finished computing the full continued fraction expansion in the previous step.
- We can then evaluate the convergents p_n/q_n via $p_n = a_n p_{n-1} + p_{n-2}$ and $q_n = a_n q_{n-1} + q_{n-2}$ with initial conditions $p_{-1} = 1$, $p_0 = a_0$, $q_{-1} = 0$, $q_0 = 1$.

<u>Example</u>: Find the fundamental unit in $\mathbb{Z}[\sqrt{2}]$ 14] using the super magic box.

• Here is the start of the super magic box calculation:

<u>Example</u>: Find the fundamental unit in $\mathbb{Z}[\sqrt{2}]$ 14] using the super magic box.

• Here is the completed super magic box:

Thus, $\sqrt{14} = [3, \overline{1, 2, 1, 6}]$ and the fundamental unit in Thus, $\sqrt{14} = [3, 1, 2, 1, 0]$ and the $\mathbb{Z}[\sqrt{14}]$ is $15 + 4\sqrt{14}$ with norm 1.

• Here is the super magic box calculation for $D = 29$:

From this calculation we can see that the fundamental unit of From this calculation we can see that the function $\mathbb{Z}[\sqrt{29}]$ is 70 + 13 $\sqrt{29}$ and it has norm $-1.$

- Since the fundamental unit 70 + 13[√] 29 has norm −1, the smallest nontrivial solution is given by its square.
- omanese nomanuar solation is given by its square.
We compute $(70 + 13\sqrt{29})^2 = 9801 + 1820\sqrt{29}$, yielding the solution $(x, y) = (9801, 1820)$.
- Notice that the super magic box calculation is quite short and easy to do by hand, quite unlike a brute-force search for solutions to $x^2 - 29y^2 = 1!$

Example (Audience Participation): Pick an integer D with $15 \le D \le 99$ that is not contained in the set $\{19, 22, 61\}$ and compute the fundamental unit in $\mathbb{Z}[\sqrt{D}]$ using the magic box. Example (Audience Participation): Pick an integer D with $15 \le D \le 99$ that is not contained in the set $\{19, 22, 61\}$ and compute the fundamental unit in $\mathbb{Z}[\sqrt{D}]$ using the magic box.

As it turns out, we can use the ideas from the super magic box algorithm to give an integer factorization algorithm, as first proposed by Lehmer and Powers in 1931. So suppose that D is some large composite integer.

- The idea, as with other factorization algorithms such as the quadratic sieve, is to find a solution to the congruence $x^2 \equiv y^2 \pmod{D}$ where $x \not\equiv y \pmod{D}$.
- \bullet The claim then is that we can find a factorization of D by computing $gcd(x + y, D)$.
- Note that we can calculate this gcd very efficiently via the Euclidean algorithm (it runs in linear time in its input size), so if we can find a solution to the congruence, we can rapidly find the factorization.

More precisely we have the following:

Lemma (Sieve Factoring Lemma)

Suppose that $x^2 \equiv y^2 \pmod{D}$ where $x \not\equiv \pm y \pmod{D}$. Then $1 < \gcd(x + y, D) < D$, so $\gcd(x + y, D)$ is a nontrivial factor of D.

Proof:

- By hypothesis, $(x + y)(x y)$ is divisible by D.
- But the gcd of $x + y$ and D cannot be 1, since then necessarily D would divide $x - y$.
- \bullet The gcd also cannot be D, since then necessarily D would divide $x + y$.
- This means $1 < \gcd(x + y, D) < D$, and so $\gcd(x + y, D)$ is a nontrivial common divisor of n.

We can use the continued fraction convergents in the super magic box algorithm to try to find a solution to $x^2 \equiv y^2 \pmod{D}$.

- The idea is that, as we have shown, $p_n^2 Dq_n^2 = (-1)^{n+1} C_{n+1}$, and so modulo D we see $p_n^2 \equiv (-1)^{n+1}C_{n+1}$ (mod D).
- \bullet So, if we are able to find a convergent such that n is odd and C_{n+1} is a perfect square, we will obtain a congruence of the form $p_n^2 \equiv k^2 \pmod{D}$.
- By our factorization lemma, this will give us a factorization of D as long as $p_n \not\equiv \pm k$ (mod D).

Example: Use the super magic box to factor $D = 1271$.

• Here is the super magic box calculation for $D = 1271$:

						4	5	.
$A_n = a_{n-1}C_{n-1} - A_{n-1}$			35		14	29	31	\cdots
$C_n = (D - A_n^2)/C_{n-1}$			46	25 ₁	43	10	31	.
$a_n = (A_n + a_0)/C_n $		35				6		.
$p_n = a_n p_{n-1} + p_{n-2}$		35	36	71	107	713	1533	\cdots
$q_n = a_n q_{n-1} + q_{n-2}$	0				3	20	43	\cdots
$p_n^2 - 1271q_n^2$			25	-43	10	-31	31	\cdots

Note $C_2 = 25$ is a perfect square. Therefore, $p_1^2 = 36^2$ will be congruent to \mathcal{C}_n modulo D , so we see 36 $^2\equiv 5^2$ (mod 1271).

• We can easily compute $gcd(36 + 5, 1271) = 41$, and so we obtain the factorization $1271 = 41 \cdot 31$.

Of course, this method requires some amount of luck to find a factorization quickly, since there is no guarantee that we will find a term with $(-1)^{n+1}C_{n+1}$ a perfect square early in the calculation.

- However, all we really need are two terms whose squares are congruent modulo $D.$ Since $|\mathcal{C}_{n+1}|<\sqrt{D},$ this means if we congruent modulo *D*. Since $|C_{n+1}| \le \sqrt{D}$, this means if we compute $2\sqrt{D}$ terms of the continued fraction expansion, we will be guaranteed to find two values of $(-1)^{n+1}C_{n+1}$ that are congruent modulo D, and thus we will obtain two convergents whose numerators satisfy $p_m^2 \equiv p_n^2 \pmod{D}$.
- That is not very efficient by itself, since trial division would That is not very emclent by itsen, since that divisionly take \sqrt{D} steps. And it might happen that the numerators of these terms have $p_m \equiv \pm p_n$ (mod D), in which case we will need to search for other tuples until we find a pair such that $p_m \not\equiv \pm p_n$ (mod D).

Continued Fraction Factorization, VI

However, if we combine these ideas with those of the quadratic sieve, we can improve the speed of this procedure.

- Instead of trying to find a single value of a for which a^2 modulo *n* is a square, we instead compute a number of different values of a such that a^2 modulo n has all of its prime divisors in a small fixed set.
- Then, by taking products of some of these values, one can obtain a congruence of the form $a^2 \equiv b^2$ (mod n) with $a \neq \pm b$ (mod *n*).
- For example, modulo 2077, if we search for powers that have small prime divisors we will find 46 $^2\equiv 3^113^1$ and $59^2 \equiv 2^2 3^3 13^1$. Multiplying them yields the equality $(46 \cdot 59)^2 \equiv (2^1 3^2 1 3^1)^2$, which is the same as $637^2 \equiv 234^2$.
- Then gcd(637 234, 2077) = 31, giving a divisor of 2077.

In general, this kind of search requires (i) finding many squares whose factorizations only involve small primes, and then (ii) finding a product of such factorizations that has a square value.

- Goal (i) we can achieve using convergents arising from the magic box, because we know that all of the values of C_n will have $|\mathcal{C}_n| < 2\sqrt{D}+1$: these are "small" relative to the modulus D, and are likely to have nice factorizations more often than larger values.
- Goal (ii) can be done efficiently with linear algebra: the idea is to find a nonzero linear dependence between the vectors of prime-factorization exponents, considered modulo 2.

For example, suppose we wanted to find a set of elements among 6, 10, 30, 150 whose product is a perfect square.

- We first find the prime factorizations $6=2^13^15^0,$ $10 = 2¹3⁰5¹$, $30 = 2¹3¹5¹$, $150 = 2¹3¹5²$.
- Then we take the four vectors of exponents $\langle 1, 1, 0 \rangle$, $\langle 1, 0, 1 \rangle$, $\langle 1, 1, 1 \rangle$, $\langle 1, 1, 2 \rangle$ and search for a linear combination of these vectors whose entries are all even.
- In this case, we can see that $\langle 1, 1, 0 \rangle + \langle 1, 1, 2 \rangle = \langle 2, 2, 2 \rangle$, corresponding to the product $6 \cdot 150 = 900 = 30^2$.
- There are simple linear-algebra procedures for finding such a linear combination by row-reducing an appropriate matrix (which is quite computationally efficient, especially because we are only working with entries in binary).

If we tune all of the computations suitably well, one may show that the resulting sieving algorithm will find a factorization of D in approximately $e^{\sqrt{2} \ln n \ln \ln n}$ time.

- This is quite a lot faster, asymptotically, than trial division This is quite a fot faster, asymptotically, than that divisible (which takes roughly \sqrt{D} time) or other algorithms like Pollard's ρ -algorithm (which heuristically takes $D^{1/4}$ time).
- There is an improvement on the quadratic sieve called the general number field sieve, which runs in roughly $e^{1.95(\ln n)^{1/3}(\ln \ln n)^{2/3}}$ time.
- The principle of the general number field sieve is similar to the quadratic sieve, but instead of working in $\mathbb Q$ it works in more general number fields like $\mathbb{Q}(\sqrt{2})=\{a+b\sqrt{2}:\ a,b\in\mathbb{Q}\}.$

We finished our proofs about the solutions to Pell's equations. We described the super magic box algorithm and used it to compute fundamental solutions for various D.

We discussed an integer factorization algorithm arising from the super magic box.

Next lecture: Miscellaneous Diophantine equations (part 1).