

**Diophantine equations after an
idea of Lehman**

Wim Couwenberg

6 November 2002

1 Diophantine equations

In general: equations for which we want to find integral solutions.

In this talk: a single polynomial equation in two variables defined over the integers.

The Lehman paper:

R. Lehman, *Factoring large integers*,
Mathematics of Computation 28 (1974),
637-646.

2 Integer factorization

Let $N > 0$ be some positive integer. Factoring N comes down to finding integer solutions to the equation $xy = N$ with $x, y > 1$.

2.1 Trial division

Try $x \in \{2, 3, \dots, \lfloor \sqrt{N} \rfloor\}$. This requires $O(\sqrt{N})$ divisions.

2.2 Fermat's method

Suppose N is odd. Let $m = \lceil \sqrt{N} \rceil$ and try if $x^2 - N$ is a square y^2 for $x \in \{m, m + 1, \dots\}$. Then $\gcd(x - y, N)$ leads to a factor of N . Nice if both factors are about the same size but worst case this method requires $O(N)$ trials.

3 Lehman's method

Resembles Fermat's method. Take $k \geq 1$ and $x = \lceil 2\sqrt{kN} \rceil$. Check if $(x + m)^2 - 4kN$ is a square for some $m \geq 0$. Then we only need to check $k \leq N^{1/3}$ and m for which $m^2k \leq N^{1/3}$ to find a factor $> N^{1/3}$, if any. So N can be factored in $O(N^{1/3})$ operations. We will prove this by a fairly generic method of "normal approximation."

3.1 Normal approximation

Let $p, q \geq 0$ be integers, not both zero, such that $\gcd(p, q) = 1$. Then $(x, y) \mapsto px + qy$ maps \mathbb{Z}^2 onto \mathbb{Z} and $px + qy = z$ defines a line ℓ_z perpendicular to (p, q) at "height" z .

Idea: choose z integral such that ℓ_z intersects the curve $xy = N$ but ℓ_{z-1} does not. Then either the intersection gives a factorization of N or there are no integral points on the curve between the two intersections.

To increase the width of the “no-solution” zone we will use additional lines ℓ_{z+m} for $m \geq 1$ and call this m the “width.”

For ℓ_z we find intersections

$$x = \frac{z \pm \sqrt{z^2 - 4pqN}}{2p}$$

$$y = \frac{z \mp \sqrt{z^2 - 4pqN}}{2q}$$

so choose $z = \lceil 2\sqrt{pqN} \rceil$ and consider ℓ_{z+m} for $m \geq 0$. The equations above show that the normal approximation (p, q) with height at most $z + m$ will find this factorization if

$$|pX - qY| \leq 2\sqrt{m}(pqN)^{1/4}.$$

If (q_1, p_1) and (q_2, p_2) are two successive continued fraction convergents of (X, Y) and $p_i, q_i > 0$ then

$$|p_1 X - q_1 Y| \leq \min\left(\frac{X}{q_2}, \frac{Y}{p_2}\right) \leq \left(\frac{N}{p_2 q_2}\right)^{1/2}.$$

Suppose $N = XY$ is a factorization of N for which $N^{1/3} < X \leq Y$. There exists a convergent (q, p) of (X, Y) such that

$$|pX - qY| \leq N^{1/3} \text{ and } pq \leq N^{1/3}.$$

We will find the factorization XY if we consider all values for $k = pq$ up to $N^{1/3}$ and for each k try widths m at least up to

$$m \geq \frac{N^{1/6}}{4\sqrt{k}}.$$

Use trial division to find any factors $\leq N^{1/3}$. Factorization takes at most $O(N^{1/3})$ arithmetic operations.

4 Some other quadrics

We apply normal approximation to some well known quadric equations.

4.1 An elliptic equation

For integers $0 < d < N$ consider the equation

$$x^2 + d y^2 = N.$$

We proceed as in the discussion of Lehman's method. For normal approximation (p, q) the line ℓ_z touches at height

$$z = \sqrt{\frac{N(q^2 + d p^2)}{d}}.$$

Checking heights $\lfloor z \rfloor, \dots, \lfloor z \rfloor - m$ will spot an integral solution (X, Y) if

$$|q X - p d Y| \leq \sqrt{m}(dN(q^2 + d p^2))^{1/4}.$$

Now there exists a convergent (p, q) of (X, dY) such that

$$|qX - pdY| \leq N^{1/3} \text{ and } q^2 + dp^2 \leq dN^{1/3}.$$

Now it suffices to check normal approximations (p, q) with $q^2 + dp^2 \leq dN^{1/3}$ and width $-m$ up to at least

$$m \geq \frac{N^{1/6}}{\sqrt{d(q^2 + dp^2)}}.$$

Again we find a solution (if any) in $O(N^{1/3})$ arithmetic operations.

4.2 A hyperbolic equation

For integers $d > 0$, not a square, and $N > 0$ consider the equation

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

Now there can be infinitely many solutions.

How fast can we find one? Suppose that there are no solutions with $|y| \leq Y$ for some $Y > 0$.

If the (not necessarily integral) point (X, Y) is on the curve then $(X, -dY)$ is a normal.

To check for solutions with $|y| > Y$ use normal approximation $(p, -q)$ with $p > 0$ minimal such that

$$\frac{dY}{X} \leq \frac{q}{p} < \sqrt{d}.$$

Such a pair can easily be determined from the continued fraction expansion of \sqrt{d} .

Then for a certain height ℓ_z touches the curve *above* the bound Y . In this way we can successively raise the lower bound Y until we find a solution.

An example. Consider $x^2 - 61y^2 = 103$.

Applying our approximation scheme we find the following successive bounds:

Y	(p, q)	width
0	(1, 0)	0
1	(1, 5)	0
2	(1, 7)	0
5	(3, 23)	1
13	(5, 39)	1
69	(21, 164)	1
222	(79, 617)	0

At this point we find the solution (2882, 369).

In general it takes $O(\log Y)$ operations to find a solution with $0 < y \leq Y$.

5 Squareful integer factorization

Suppose $N > 0$ is a squareful integer i.e. $d^2|N$ for some integer $d > 1$. Given this additional information, can we speed up factorization?

Apply normal approximation to the curve $x^2y = N$. For a normal approximation (p, q) the line ℓ_z touches at height

$$z = \frac{3}{2}(2p^2qN)^{1/3}.$$

For $m \geq 0$, puiseux expansion shows that the line ℓ_{z+m} intersects approximately at

$$x \approx \frac{2z \pm 2\sqrt{mz}}{3p}$$

$$y \approx \frac{z \mp 2\sqrt{mz}}{3q}$$

so we expect to find a solution if

$$|px - q2y| \leq 2\sqrt{mz} = \sqrt{6m}(2p^2qN)^{1/6}.$$

First check, by trial divisions, if there is a factorization $N = X^2Y$ with $X \leq N^{1/4}$ or $Y \leq N^{1/4}$. If not, then there is a convergent (q, p) of $(X, 2Y)$ such that

$$|pX - q2Y| \leq N^{1/4} \text{ and } pq^2 \leq 2N^{1/4}.$$

So we expect to find a factorization if we check all normal approximations (p, q) with $pq^2 \leq 2N^{1/4}$ and for each such approximation check m up to at least about

$$m \geq \frac{N^{1/6}}{6(2p^2q)^{1/3}}.$$

This involves $O(N^{1/4} \log N)$ arithmetic operations. Hence a squareful integer can be factored in $O(N^{1/4+\epsilon})$ bit operations.

6 A Thue equation

As a final example we consider the Thue equation

$$x^3 - 2y^3 = 1.$$

It has the small solutions $(1, 0)$, $(-1, -1)$.

Using normal approximation one can check for solutions with $|y| \leq N$ very quickly as we will see.

As for the hyperbolic equation we apply a scheme to increase a lower bound $Y > 0$ for solutions with $y > Y$. If (X, Y) is on the curve, then $(X^2, -2Y^2)$ is a normal. In each step we look for $(p, -q)$ with p minimal such that

$$\frac{2Y^2}{X^2} \leq \frac{q}{p} < 2^{1/3}.$$

It turns out that for this normal approximation we can intersect ℓ_1 at height one to increase the lower bound Y .

Let's try it out, start at $Y = 1$:

Y	(p, q)
1	(1, 1)
4	(27, 34)
56	(504, 635)
4782	(913235, 1150604)
1343193	(2448641198, 3085094589)
4280199447	(186454048314072, 234917380309015)
3018338689940635	...

The bound Y increases very rapidly indeed! If Y, Y' are successive bounds from this table then $Y' > Y^{3/2}$. This hints (correctly) at $O(\log \log Y)$ arithmetic operations to check for solutions with $1 \leq y \leq Y$.