ELEMENTARY NUMBER THEORY II

DAVID PIERCE

These are notes from Math 366 in the METU Mathematics Department, spring semester, 2007/8. Class meets Tuesdays at 13.40 for two hours and Fridays at 13.40 (originally 12.40) for one hour, and continues. So these notes are unfinished.

I typeset these notes after class, from memory and from handwritten notes prepared before class. I do some polishing and slight rearrangement.

The main published reference for the course is [1], which has apparently been on reserve in the library since the last time this course was offered several years ago. I have that text only in the form of a photocopy of chapters 6–11, used by Ayşe Berkman when she was a student. The text is a rough guide only, and I may change its terminology and notation.

All special symbols used in these notes are found at the head of the index.

For continued fractions, the text [2] used for Math 365 is useful, as is [4]. I also consult [3] and [6], and occasionally other works.

Class was cancelled Friday, February 29, because I was in İstanbul for my *doçentlik* exam. Ayşe taught for me on the following Tuesday, since I was sick with a gastrointestinal infection from the trip.

On Friday, March 21, I spent the class solving homework problems: the first examination was the following Monday evening.

Class on Tuesday, April 8, was only one hour, because of a special seminar that day (on teaching conic sections).

Contents

Date: April 18, 2008.

List of Figures

List of Topics

1. February 19, 2008 (Tuesday)

We begin with some **Diophantine equations** (that is, polynomial equations in which all constants and variables are integers).

Problem 1. *Solve*

$$x^2 + y^2 = z^2 \tag{1}$$

(that is, find all solutions).

Solution. The following are equivalent:

(i) (a, b, c) is a solution;
(ii) $(|a|, |b|, |c|)$ is a solution;
(iii) (na, nb, nc) is a solution, where $n \neq 0$;
(iv) (b, a, c) is a solution.

Also, (1) is equivalent to

$$x^2 = (z + y)(z - y).$$

Suppose (a, b, c) is a solution of (1) such that $a, b, c > 0$ and $\gcd(a, b, c) = 1$. Then (a, b, c) may be called a **primitive solution,** and all solutions can be obtained from primitive solutions. Observe that not both a and b are even. Also, if $a, b \equiv 1 \pmod 2$, then $c^2 \equiv a^2 + b^2 \equiv 2 \pmod 4$, which is absurd. So exactly one of a and b is even. Say a is even. Then b and c are odd, and

$$\left(\frac{a}{2}\right)^2 = \left(\frac{c+b}{2}\right)\left(\frac{c-b}{2}\right).$$

Also $(c + b)/2$ and $(c - b)/2$ are co-prime, since their sum is c and their difference is b. Hence each must be a square; say

$$\frac{c+b}{2} = n^2, \qquad \frac{c-b}{2} = m^2,$$

where $n, n > 0$. Then

$$c = n^2 + m^2, \qquad b = n^2 - m^2, \qquad a = 2nm.$$

Moreover, n and m are co-prime, and exactly one of them is odd (since c is odd).

Conversely, suppose n and m are co-prime, exactly one of them is odd, and $0 < m < n$. Then the triple $(2nm, n^2 - m^2, n^2 + m^2)$ solves (1). Moreover, every common prime factor of $n^2 - m^2$ and $n^2 + m^2$ is a factor of the sum $2n^2$ and the difference $2m^2$, hence of n and m. So there is no common prime factor, and the triple is a *primitive* solution.

We conclude that there is a one-to-one correspondence between:

(i) pairs (m, n) of co-prime integers, where $0 < m < n$, and exactly one of m and n is odd;
(ii) primitive solutions (a, b, c) to (1), where a is even.

The correspondence is $(x, y) \mapsto (2xy, y^2 - x^2, y^2 + x^2)$. □

Problem 2. *Solve*

$$x^4 + y^4 = z^4. \tag{2}$$

Solution. Let (a, b, c) be a solution, where $a, b, c > 0$, and $\gcd(a, b, c) = 1$. Then (a^2, b^2, c^2) is a primitive **Pythagorean triple** (that is, solution to (1)). We may assume a is even, and so

$$a^2 = 2mn, \quad b^2 = n^2 - m^2, \quad c^2 = n^2 + m^2.$$

In particular,
$$m^2 + b^2 = n^2.$$
Since $\gcd(a, b) = 1$, and every prime factor of m divides a, we have $\gcd(m, b) = 1$. Hence (m, b, n) is a primitive Pythagorean triple. Also m is even, since b is odd. Hence
$$m = 2de, \quad b = e^2 - d^2, \quad n = e^2 + d^2$$
for some d and e. Then
$$a^2 = 2mn = 4de(e^2 + d^2).$$
But $\gcd(d, e) = 1$, so $e^2 + d^2$ is prime to both d and e. Therefore each of d, e, and $e^2 + d^2$ must be square: say
$$d = r^2, \quad e = s^2, \quad e^2 + d^2 = t^2.$$
This gives $t^2 = e^2 + d^2 = s^4 + r^4$; that is, (s, r, t) is a solution to
$$x^4 + y^4 = z^2. \tag{3}$$
But (a, b, c^2) is also a solution to this; moreover,
$$1 \leqslant |t| \leqslant t^2 = e^2 + d^2 = n \leqslant n^2 < n^2 + m^2 = c^2.$$
We never used that c^2 is a square. Thus, for every solution to (3) with positive entries, there is a solution with positive entries in which the third entry is smaller. This is absurd; therefore there is no such solution to (3), or to (2). □

We used here Fermat's method of **infinite descent.**

In Elementary Number Theory I, we proved that the Diophantine equation
$$x^2 + y^2 + z^2 + w^2 = n$$
is soluble for every positive integer n.

Problem 3. *Find those n for which*
$$x^2 + y^2 = n$$
is soluble.

Solution. Let S be the set of such n. Since
$$\begin{aligned}(a^2 + b^2)(c^2 + d^2) &= |a + b\mathrm{i}|^2|c + d\mathrm{i}|^2 \\ &= |(a + b\mathrm{i})(c + d\mathrm{i})|^2 \\ &= |(ac - bd) + (ad + bc)\mathrm{i}|^2 \\ &= (ac - bd)^2 + (ad + bd)^2,\end{aligned}$$
S is closed under multiplication. We ask now: Which primes are in S?

All squares are congruent to 0 or 1 *modulo* 4. Hence elements of S are congruent to 0, 1, or 2 *modulo* 4. Therefore S contains no primes that are congruent to $3 \pmod 4$.

However, S does contain 2, since $2 = 1^2 + 1^2$.

Suppose $p \equiv 1 \pmod 4$. Then -1 is a quadratic residue *modulo* p, so
$$-1 \equiv a^2 \pmod p$$
for some a, where we may assume $|a| < p/2$. Hence
$$a^2 + 1 = tp$$

for some positive t. This means $tp \in S$. Let k be the least positive number such that $kp \in S$. Since

$$0 < t = \frac{a^2+1}{p} < \frac{(p/2)^2+1}{p} = \frac{p}{4} + \frac{1}{p} < p,$$

we have $0 < k \leqslant t < p$. By assumption,

$$kp = b^2 + c^2 \tag{4}$$

for some b and c. There are d and e such that

$$d \equiv b, \quad e \equiv c \pmod{k}; \qquad |d|, |e| \leqslant \frac{k}{2}.$$

Then $d^2 + e^2 \equiv b^2 + c^2 \equiv 0 \pmod{k}$, so

$$d^2 + e^2 = km \tag{5}$$

for some m, where

$$0 \leqslant m = \frac{d^2+e^2}{k} \leqslant \frac{2(k/2)^2}{k} = \frac{k}{2} < k.$$

But multiply (4) and (5), getting

$$\begin{aligned} k^2mp &= (b^2+c^2)(d^2+e^2) \\ &= (bd-ce)^2 + (be+cd)^2 \\ &= (bd+ce)^2 + (be-cd)^2. \end{aligned}$$

Since

$$bd + ce \equiv b^2 + c^2 \equiv 0, \quad be - cd \equiv bc - cb \equiv 0 \pmod{k},$$

we can divide by k^2, getting

$$mp = \left(\frac{bd+ce}{k}\right)^2 + \left(\frac{be-cd}{k}\right)^2.$$

This implies $mp \in S$. By minimality of k, we have $m = 0$. Therefore $d^2 + e^2 = 0$, so $d = 0 = e$. Then $b, c \equiv 0 \pmod{k}$, so

$$k^2 \mid kp,$$

and therefore $k \mid p$. This means $k = 1$, so $p \in S$.

Finally, suppose $n \in S$ and $p \mid n$. Then $n = a^2 + b^2$ for some a and b, so

$$a^2 + b^2 \equiv 0 \pmod{p}.$$

If $p \mid a$, then $p \mid b$, so $p^2 \mid n$, which means n is not square-free. If $p \nmid a$, then a is invertible *modulo* p, so $1 + (b/a)^2 \equiv 0 \pmod{p}$, which means -1 is a quadratic residue *modulo* p, and so $p = 2$ or else $p \equiv 1 \pmod{4}$.

The conclusion is that S contains just those numbers of the form n^2m, where m is square-free and has no prime factors congruent to 3 *modulo* 4. $\square$

2. February 22, 2008 (Friday)

Solving (1) in integers is related to finding integrals like

$$\int \frac{\mathrm{d}\,\theta}{2+3\sin\theta}.$$

Indeed,

$$x^2+y^2=z^2 \iff \left(\frac{x}{z}\right)^2+\left(\frac{y}{z}\right)^2=1 \text{ or } x=y=z=0.$$

So finding Pythagorean triples corresponds to solving

$$x^2+y^2=1$$

in *rationals.* To do so, since the equation defines the unit circle, consider also the line through $(-1,0)$ with slope t, so that its Y-intercept is also t, as in Figure 1: this line is

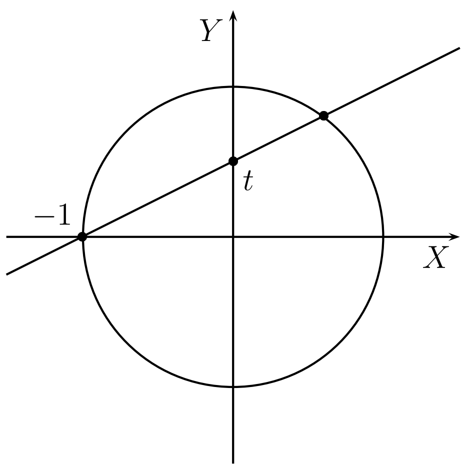

Figure 1. Finding the rational points of the circle

given by

$$y=tx+t. \tag{6}$$

The circle and the line meet at $(-1,0)$ and also (x,y), where

$$x^2+(tx+t)^2=1,$$
$$(1+t^2)x^2+2t^2x+t^2-1=0,$$
$$x^2+\frac{2t^2}{1+t^2}\cdot x-\frac{1-t^2}{1+t^2}=0.$$

The constant term in the left member of the last equation is the product of the roots; one of the roots is -1; so we get

$$(x,y)=\left(\frac{1-t^2}{1+t^2},\frac{2t}{1+t^2}\right).$$

If t is rational, then so are the coordinates of this point, which is therefore a **rational point** of the circle. Conversely, if x and y are rational, then so is t, by (6). Hence the function

$$t\mapsto\left(\frac{1-t^2}{1+t^2},\frac{2t}{1+t^2}\right)$$

is a one-to-one correspondence, with inverse

$$(x, y) \mapsto \frac{y}{x+1},$$

between $\mathbb{Q}$ and the set of rational points (other than $(-1, 0)$) of the unit circle.

Hence we can conclude that every integral solution of (1) is a multiple of

$$(1 - t^2, 2t, 1 + t^2).$$

Taking $t = m/n$ and multiplying by n^2, we get

$$(n^2 - m^2, 2mn, n^2 + m^2).$$

* * * * *

We can convert $\sqrt{2}$ into a *continued fraction* as follows:

$$\sqrt{2} = 1 + (\sqrt{2} - 1) = 1 + \frac{1}{\sqrt{2}+1} = 1 + \frac{1}{2 + (\sqrt{2} - 1)} = 1 + \cfrac{1}{2 + \cfrac{1}{\sqrt{2}+1}} = \cdots$$

In the general procedure, given a real number x, we define a_n and ξ_n recursively as follows, where square brackets denote the greatest-integer function:

$$\begin{aligned} a_0 &= [x], & \xi_0 &= x - a_0; \\ a_1 &= \left[\frac{1}{\xi_0}\right], & \xi_1 &= \frac{1}{\xi_0} - a_1; \end{aligned}$$

and generally

$$a_n = \left[\frac{1}{\xi_{n-1}}\right], \quad \xi_n = \frac{1}{\xi_{n-1}} - a_n;$$

where ξ_{n-1} must be non-zero for a_n to be defined. Then

$$x = a_0 + \xi_0 = a_0 + \frac{1}{a_1 + \xi_1} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \xi_2}} = \cdots$$

These are **continued fractions.** Taking $x = \sqrt{3}$, we get

$$\begin{aligned} & & a_0 &= 1, & \xi_0 &= \sqrt{3} - 1, \\ \frac{1}{\xi_0} &= \frac{\sqrt{3}+1}{2}, & a_1 &= 1, & \xi_1 &= \frac{\sqrt{3}-1}{2}, \\ \frac{1}{\xi_1} &= \sqrt{3} + 1, & a_2 &= 2, & \xi_2 &= \sqrt{3} - 1, \end{aligned}$$

and now the process repeats:

$$\xi_n = \begin{cases} \sqrt{3} - 1, & \text{if } n \text{ is even;} \\ \dfrac{\sqrt{3}-1}{2}, & \text{if } n \text{ is odd;} \end{cases} \qquad a_n = \begin{cases} 1, & \text{if } n = 0, \text{ or } n \text{ is odd;} \\ 2, & \text{if } n \text{ is positive and even.} \end{cases}$$

It appears that

$$\sqrt{3} = 1 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{\ddots}}}}} \tag{7}$$

But to make this precise, we need some notion of *convergence*. To define this, we introduce some notation. Here square brackets do *not* denote the greatest-integer function:

$$\begin{aligned}
[a_0] &= a_0,\\
[a_0; a_1] &= a_0 + \frac{1}{a_1},\\
[a_0; a_1, a_2] &= a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2}},
\end{aligned}$$

and so forth, so that

$$[a_0; a_1, \dots, a_{n+1}] = [a_0; a_1, \dots, a_{n-1}, a_n + \frac{1}{a_{n+1}}].$$

Here we must have $a_n \neq 0$ when $n > 0$; we shall assume also $a_n > 0$ when $n > 0$. We can also use the notation in the infinite case. For example, from $\sqrt{3}$, we have obtained $[1; 1, 2, 1, 2, \dots]$, which we can write as

$$[1; \overline{1, 2}].$$

But again, we have not yet established that this notation defines a particular number.

3. February 26, 2008 (Tuesday)

The process of obtaining the sequences $(a_n : n \in \omega)$ and $(\xi_n : n \in \omega)$ from x as above can be compared with the **Euclidean algorithm:** To find $\gcd(155, 42)$, we compute

$$\begin{aligned}
155 &= 42 \cdot 3 + 29,\\
42 &= 29 \cdot 1 + 13,\\
29 &= 13 \cdot 2 + 3,\\
13 &= 3 \cdot 4 + 1,\\
3 &= 1 \cdot 3 + 0.
\end{aligned}$$

We can rewrite this as

$$\begin{aligned}
\left[\frac{155}{42}\right] &= 3, & \frac{155}{42} - 3 &= \frac{29}{42}, \\
\left[\frac{42}{29}\right] &= 1, & \frac{42}{29} - 1 &= \frac{13}{29}, \\
\left[\frac{29}{13}\right] &= 2, & \frac{29}{13} - 2 &= \frac{3}{13}, \\
\left[\frac{13}{3}\right] &= 4, & \frac{13}{3} - 4 &= \frac{1}{3}, \\
\left[\frac{3}{1}\right] &= 3, & \frac{3}{1} - 3 &= 0.
\end{aligned}$$

Thus, when $x = 155/42$, then the sequence of a_n is just $(3, 1, 2, 4, 3)$, and

$$\frac{155}{42} = 3 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{4 + \cfrac{1}{3}}}}.$$

Thus we can write every fraction as a (finite) **continued fraction** $[a_0; a_1, \ldots, a_n]$, where the a_k are integers, and all of them are positive except perhaps a_0. Such a continued fraction is called **simple.** We shall work only with simple continued fractions. But the continued fraction obtained for irrational x does not terminate.

The kth **convergent** of $[a_0; a_1, \ldots]$ is $[a_0; a_1, \ldots, a_k]$. For example, the convergents of $[1; \overline{1, 2}]$ are

$$1, \quad 2, \quad \frac{5}{3}, \quad \frac{7}{4}, \quad \frac{19}{11}, \quad \frac{26}{15}, \quad \frac{71}{41}, \quad \frac{97}{56}, \quad \ldots$$

by a tedious computation to be made easier in a moment. How are these convergents as approximations of $\sqrt{3}$? We have

$$\begin{aligned}
\left(\frac{5}{3}\right)^2 &= \frac{25}{9}, & 25 - 3 \cdot 9 &= -2, \\
\left(\frac{7}{4}\right)^2 &= \frac{49}{16}, & 49 - 3 \cdot 16 &= 1, \\
\left(\frac{19}{11}\right)^2 &= \frac{361}{121}, & 361 - 3 \cdot 121 &= -2, \\
\left(\frac{26}{15}\right)^2 &= \frac{676}{225}, & 676 - 3 \cdot 225 &= 1, \\
\left(\frac{71}{41}\right)^2 &= \frac{5041}{1681}, & 5041 - 3 \cdot 1681 &= -2.
\end{aligned}$$

We shall define p_k and q_k so that

$$\frac{p_k}{q_k} = [a_0; a_1, \ldots, a_k], \tag{8}$$

the kth convergent of $[a_0; a_1, \dots]$. We start with

$$\begin{aligned}
\frac{p_0}{q_0} &= a_0, & p_0 &= a_0, & q_0 &= 1;\\
\frac{p_1}{q_1} &= a_0 + \frac{1}{a_1} = \frac{a_0 a_1 + 1}{a_1}, & p_1 &= p_0 a_1 + 1, & q_1 &= a_1;\\
\frac{p_2}{q_2} &= a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2}} = \frac{a_0 a_1 a_2 + a_0 + a_2}{a_1 a_2 + 1}, & p_2 &= p_1 a_2 + p_0, & q_2 &= q_1 a_2 + q_0.
\end{aligned}$$

Following this pattern, we define

$$p_{k+2} = p_{k+1} a_{k+2} + p_k, \qquad q_{k+2} = q_{k+1} a_{k+2} + q_k.$$

Theorem 1. *Equation* (8) *holds for all k in ω.*

Proof. Use induction. The claim holds when $k = 0$. By assuming the claim for some k, we can compute $[a_0; a_1, \dots, a_{k+3}]$ from it, replacing a_{k+2} with $a_{k+2} + 1/a_{k+3}$:

$$\begin{aligned}
[a_0; a_1, \dots, a_{k+3}] = \frac{p_{k+1} \cdot \left(a_{k+2} + \dfrac{1}{a_{k+3}}\right) + p_k}{q_{k+1}\left(a_{k+2} + \dfrac{1}{a_{k+3}}\right) + q_k} &= \frac{p_{k+1} a_{k+2} a_{k+3} + p_{k+1} + p_k a_{k+3}}{q_{k+1} a_{k+2} a_{k+3} + q_{k+1} + q_k a_{k+3}}\\
&= \frac{p_{k+2} a_{k+3} + p_{k+1}}{q_{k+2} a_{k+3} + q_{k+1}}.
\end{aligned}$$

By induction, we have (8) for all k. □

Is p_k/q_k in lowest terms?

Theorem 2. *The integers p_k and q_k are co-prime; in fact,*

$$\frac{p_{k+1}}{q_{k+1}} - \frac{p_k}{q_k} = \frac{(-1)^k}{q_{k+1} q_k},$$

equivalently,

$$p_{k+1} q_k - p_k q_{k+1} = (-1)^k.$$

Proof. Again use induction. We have

$$\frac{p_1}{q_1} - \frac{p_0}{q_0} = \frac{1}{a_1} = \frac{(-1)^0}{q_1 q_0},$$

so the claim holds when $k = 0$. Suppose it holds for some k. Then

$$p_{k+2} q_{k+1} - p_{k+1} q_{k+2} = (p_{k+1} a_{k+2} + p_k) q_{k+1} - p_{k+1}(q_{k+1} a_{k+2} + q_k) = p_k q_{k+1} - p_{k+1} q_k,$$

which is $= -(-1)^k$ or $(-1)^{k+1}$. Thus the claim holds for all k. □

Corollary. *$\{p_{2n}/q_{2n}\}$ is increasing, and $\{p_{2n+1}/q_{2n+1}\}$ is decreasing, and*

$$\frac{p_0}{q_0} < \frac{p_2}{q_2} < \dots < \frac{p_3}{q_3} < \frac{p_1}{q_1}.$$

The two sequences converge to the same limit. If the convergents are obtained as above from x, then their limit is x.

Now we are justified in writing (7), for example.

*   *   *   *   *

With these tools, we turn now to the **Pell equation,**

$$x^2 - dy^2 = 1. \tag{9}$$

We first take care of some trivial cases:

(i) If $d < -1$, then $(x, y) = (\pm 1, 0)$.
(ii) If $d = -1$, then (x, y) is $(\pm 1, 0)$ or $(0, \pm 1)$.
(iii) If $d = 0$, then $x = \pm 1$, while y is anything.
(iv) If d is a positive square, as a^2, then $1 = (x + ay)(x - ay)$, so $x \pm ay$ are alike ± 1, and therefore $y = 0$ and $x = \pm 1$.

Henceforth we assume d is a positive non-square. Then (9) still has the solution $(\pm 1, 0)$; but perhaps it has others too. Indeed, in case $d = 3$, we found (on p. 9) solutions $(49, 16)$ and $(676, 225)$, with a possibility of finding more if the pattern continues.

Suppose (a, b) and (s, t) are solutions to (9). Then

$$a^2 - db^2 = 1, \qquad s^2 - dt^2 = 1,$$

so multiplication gives

$$1 = (a^2 - db^2)(s^2 - dt^2) = (as \pm dbt)^2 - d(at \pm bs)^2, \tag{10}$$

so $(as \pm dbt, at \pm bs)$ is a solution. We can repeat this process on (a, b) as follows. We can define the ordered pair (a_n, b_n) of integers by

$$a_n + b_n\sqrt{d} = (a + b\sqrt{d})^n.$$

Then also $a_n - b_n\sqrt{d} = (a - b\sqrt{d})^n$, so

$${a_n}^2 - d{b_n}^2 = (a_n + b_n\sqrt{d})(a_n - b_n\sqrt{d}) = (a + b\sqrt{d})^n(a - b\sqrt{d})^n = (a^2 - b^2 d)^n = 1,$$

and (a_n, b_n) is a solution. If $a + b\sqrt{d} > 1$, then these solutions (a_n, b_n) must all be distinct.

We ask now: Is there *one* solution (a, b) such that $a + b\sqrt{d} > 1$?

Lemma 1. *If d is a positive non-square, then, for some positive k, the equation*

$$x^2 - dy^2 = k \tag{11}$$

has infinitely many solutions.

Proof. Let $(p_n/q_n \colon n \in \omega)$ be the sequence of convergents for $\sqrt{d}$. When n is odd, we have

$$\begin{aligned} &0 < \frac{p_n}{q_n} - \sqrt{d} < \frac{p_n}{q_n} - \frac{p_{n+1}}{q_{n+1}} = \frac{1}{q_{n+1}q_n} < \frac{1}{{q_n}^2}, \\ &0 < \frac{p_n}{q_n} + \sqrt{d} < \frac{2p_n}{q_n}; \end{aligned}$$

multiplying gives

$$0 < \frac{{p_n}^2}{{q_n}^2} - d < \frac{2p_n}{{q_n}^3}, \qquad 0 < {p_n}^2 - d{q_n}^2 < \frac{2p_n}{q_n} < \frac{2p_1}{q_1}.$$

Thus there are finitely many possibilities for ${p_n}^2 - d{q_n}^2$, so one of them must be realized infinitely many times. □

If (a, b) solves (9), and each of a and b is positive, then let us refer to (a, b) as a **positive** solution.

Lemma 2. *If d is a positive non-square, then the equation (9) has a positive solution.*

Proof. By the previous lemma, we may let k be a positive number such that (11) has infinitely many solutions. But there are just finitely many pairs (a, b) such that $0 \leqslant a < k$ and $0 \leqslant b < k$. Hence there must be one such pair for which (11) together with the congruences

$$x \equiv a, \quad y \equiv b \pmod{k}$$

have infinitely many solutions. Let (m, n) and (s, t) be two such solutions. Then by the identity in (10), we have

$$k^2 = (m^2 - dn^2)(s^2 - dt^2) = (ms - dnt)^2 - d(mt - ns)^2.$$

But we have also

$$ms - dnt \equiv m^2 - dn^2 \equiv 0, \quad mt - ns \equiv mn - nm \equiv 0 \pmod{k}.$$

So we can divide by k^2 to get

$$1 = \left(\frac{ms - dnt}{k}\right)^2 - d\left(\frac{mt - ns}{k}\right)^2.$$

Hence $(|ms - dnt|/k, |mt - ns|/k)$ is a positive solution to (9). □

Theorem 3. *If d is a positive non-square, let (a, b) be the positive solution (ℓ, m) of (9) for which $\ell + m\sqrt{d}$ is minimized. Then the equation (9) has just the solutions (s, t), where $(|s|, |t|) = (a_n, b_n)$ for some non-negative integer n, where $a_n + b_n\sqrt{d} = (a + b\sqrt{d})^n$.*

Proof. Let (a, b) be as in the statement. (It exists by Lemma 2.) Then $a + b\sqrt{d} > 1$, so the powers of $a + b\sqrt{d}$ grow arbitrarily large. We know that all of the (a_n, b_n) are solutions of (9). Let (s, t) be an arbitrary positive solution. Then

$$(a + b\sqrt{d})^n \leqslant s + t\sqrt{d} < (a + b\sqrt{d})^{n+1}$$

for some non-negative n. Since $(a + b\sqrt{d})(a - b\sqrt{d}) = 1$, and $a + b\sqrt{d}$ is positive, so is $a - b\sqrt{d}$. We can therefore multiply by the nth power of this, getting

$$1 \leqslant (s + t\sqrt{d})(a - b\sqrt{d})^n < a + b\sqrt{d}.$$

But we have

$$(s + t\sqrt{d})(a - b\sqrt{d})^n = \ell + m\sqrt{d}$$

for some ℓ and m, and then also $(s - t\sqrt{d})(a + b\sqrt{d})^n = \ell - m\sqrt{d}$. Hence $\ell^2 - m^2 d = 1$, so (ℓ, m) is a solution of (9). But we have

$$1 \leqslant \ell + m\sqrt{d} < a + b\sqrt{d}.$$

Hence $0 \leqslant \ell - m\sqrt{d} \leqslant 1$, so neither ℓ nor m can be negative. By minimality of $a + b\sqrt{d}$, we must have $\ell + m\sqrt{d} = 1$, so $(s, t) = (a_n, b_n)$. □

4. March 4, 2008 (Tuesday)

If F_1 is a subfield of a field F_2, then F_2 is a vector-space over F_1: the dimension is denoted by

$$[F_2 : F_1].$$

If K is a field such that $\mathbb{Q} \subseteq K$, and $[K : \mathbb{Q}] = 2$, we say K is a **quadratic field.**

Suppose K is a quadratic field. In particular, there is x in $K \smallsetminus \mathbb{Q}$. Then 1 and x are linearly independent over $\mathbb{Q}$, so $\{1, x\}$ must be a basis of K over $\mathbb{Q}$. In particular,

$$x^2 + bx + c = 0$$

for some b and c in $\mathbb{Q}$, so

$$x = \frac{-b \pm \sqrt{b^2 - 4c}}{2}.$$

Then $\sqrt{b^2 - 4c} \in K \smallsetminus \mathbb{Q}$. We can write $b^2 - 4c$ as $s^2 d$, where $s \in \mathbb{Q}$ and d is a square-free integer different from 1. Then $\sqrt{d} \in K \smallsetminus \mathbb{Q}$, so $\{1, \sqrt{d}\}$ is a basis of K, and

$$K = \{x + y\sqrt{d} \colon x, y \in \mathbb{Q}\}.$$

Also, K is the smallest subfield of $\mathbb{C}$ that contains $\sqrt{d}$; so we can denote K by

$$\mathbb{Q}(\sqrt{d}).$$

It is an exercise to check that, conversely, we always have

$$\mathbb{Q}(\sqrt{d}) = \{x + y\sqrt{d} \colon x, y \in \mathbb{Q}\}.$$

In particular, if $a, b \in \mathbb{Q}$, and $b \neq 0$, then, assuming d is not a square, we have

$$\frac{1}{a + b\sqrt{d}} = \frac{a - b\sqrt{d}}{a^2 - b^2 d} = \frac{a}{a^2 - b^2 d} - \frac{b}{a^2 - b^2 d}\sqrt{d}.$$

So non-zero elements of $\{x + y\sqrt{d} \colon x, y \in \mathbb{Q}\}$ have multiplicative inverses.

A rational number is an integer if and only if it satisfies an equation

$$x + c = 0,$$

where $c \in \mathbb{Z}$. This is a trivial observation, but it motivates the following definition. An element of a quadratic field is an **integer** of that field if it is an integer in the old sense, or else it satisfies an equation

$$x^2 + bx + c = 0,$$

where $b, c \in \mathbb{Z}$. Henceforth, integers in the old sense can be called **rational integers.** More generally, an **algebraic integer** is the root of an equation

$$x^n + a_1 x^{n-1} + \cdots + a_{n-1} x + a_n = 0,$$

where $a_i \in \mathbb{Z}$; but we shall not go beyond the quadratic case, $n = 2$.

* * * * *

The integers of $\mathbb{Q}(\mathrm{i})$, that is, $\mathbb{Q}(\sqrt{-1})$, are called the **Gaussian integers.** The subset $\{x + y\mathrm{i} \colon x, y \in \mathbb{Z}\}$ of $\mathbb{Q}(\mathrm{i})$ is denoted by

$$\mathbb{Z}[\mathrm{i}].$$

Theorem 4. *The Gaussian integers compose the set* $\mathbb{Z}[\mathrm{i}]$.

Proof. Let $\alpha = m + n\mathrm{i}$. Then $(\alpha - m)^2 = (n\mathrm{i})^2 = -n^2$, so $\alpha^2 - 2m\alpha + m^2 + n^2 = 0$, and α is a Gaussian integer.

Suppose conversely α is a Gaussian integer. Then $\alpha^2 + b\alpha + c = 0$ (by definition) for some b and c in $\mathbb{Z}$. Hence

$$\alpha = \frac{-b \pm \sqrt{b^2 - 4c}}{2}.$$

We must have $\alpha \in \mathbb{Q}(\mathrm{i})$. So $\pm(b^2 - 4c)$ is a square in $\mathbb{Z}$. Say $b^2 - 4c = \pm e^2$. Then

$$b^2 \mp e^2 = 4c \equiv 0 \pmod 4; \qquad b \equiv e \pmod 2.$$

Also,

$$\alpha = \frac{-b \pm e}{2} \quad \text{or} \quad \alpha = \frac{-b \pm e\mathrm{i}}{2}.$$

If $b \equiv e \equiv 0 \pmod 2$, then α is in $\mathbb{Z}$ or $\mathbb{Z} \oplus \mathbb{Z}\mathrm{i}$. If $b \equiv e \equiv 1 \pmod 2$, then $4 \nmid b^2 + e^2$, so $b^2 - e^2 = 4c$, which means $b^2 - 4c = e^2$, so that $\alpha \in \mathbb{Z}$. □

It is an exercise to check that $\mathbb{Z}[\mathrm{i}]$ is a ring. But multiplicative inverses may fail to exist in $\mathbb{Z}[\mathrm{i}]$. For example, $2 \in \mathbb{Z}[\mathrm{i}]$, but $1/2 \notin \mathbb{Z}[\mathrm{i}]$.

The **norm** on $\mathbb{Q}(\mathrm{i})$ is the function given by

$$\mathrm{N}(a + b\mathrm{i}) = a^2 + b^2 = |a + b\mathrm{i}|^2; \tag{12}$$

so its values are non-negative rational numbers, and

$$\mathrm{N}(\alpha\beta) = \mathrm{N}(\alpha)\,\mathrm{N}(\beta).$$

Note that

$$\frac{1}{a + b\mathrm{i}} = \frac{a - b\mathrm{i}}{a^2 + b^2} = \frac{a - b\mathrm{i}}{\mathrm{N}(a + b\mathrm{i})}.$$

Hence

$$\frac{1}{a + b\mathrm{i}} \in \mathbb{Z}[\mathrm{i}] \iff \mathrm{N}(a + b\mathrm{i}) = \pm 1 \iff \mathrm{N}(a + b\mathrm{i}) = 1.$$

So $a + b\mathrm{i}$ is a unit of $\mathbb{Z}[\mathrm{i}]$ if and only if $a^2 + b^2 = 1$, and the unit Gaussian integers are ± 1 and $\pm\mathrm{i}$.

5. March 7, 2008 (Friday)

An **integral domain** (*tamlık alanı*), or simply a **domain,** is a sub-ring of a field. For us, the field will usually be $\mathbb{C}$. As an example, $\mathbb{Z}[\mathrm{i}]$ is an integral domain. A *Euclidean domain* is a domain in which the Euclidean algorithm works. This means we can perform division with remainder, where the remainder is "smaller" than the divisor; and a sequence of remainders of decreasing size must terminate. Since decreasing sequences of natural numbers must terminate, we shall use natural numbers to measure size. So, formally, a domain R is a **Euclidean domain** if there is a function $x \mapsto \mathrm{d}(x)$, the **degree,** from $R \smallsetminus \{0\}$ into $\mathbb{N}$ such that, for all α and β in R, if $\beta \neq 0$, then the system

$$\alpha = \beta x + y \;\&\; \mathrm{d}(y) < \mathrm{d}(\beta)$$

is soluble in R.

Gaussian integers have a size, namely the absolute value, but this need not be a rational integer. The square is, however. So we let $\mathrm{d}(x)$ be the norm $\mathrm{N}(x)$ as in (12).

Theorem 5. $\mathbb{Z}[\mathrm{i}]$ *with* $x \mapsto \mathrm{N}(x)$ *is a Euclidean domain.*

Proof. Given α and β in $\mathbb{Z}[\mathrm{i}]$, where $\beta \neq 0$, we must solve

$$\alpha = \beta x + y \ \& \ \mathrm{N}(y) < \mathrm{N}(\beta)$$

The Gaussian-integral multiples of β compose a square **lattice** (*kafes*) in $\mathbb{C}$, as in Figure 2. Then α is in one of the squares whose vertices are among these multiples. The closest

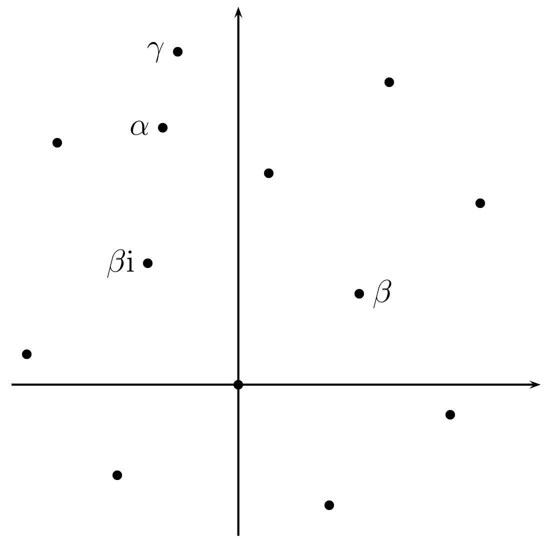

Figure 2. A lattice of Gaussian multiples

vertex to α is some γ such that

$$|\alpha - \gamma| \leqslant \frac{\sqrt{2}}{2}|\beta|, \qquad \mathrm{N}(\alpha - \gamma) \leqslant \frac{1}{2}\,\mathrm{N}(\beta).$$

So our solution is $(\gamma/\beta, \alpha - \gamma)$. □

Doing the proof more algebraically, we have $\alpha/\beta = r + s\mathrm{i}$ for some r and s in $\mathbb{Q}$. There are m and n in $\mathbb{Z}$ such that $|r - m|, |s - n| \leqslant 1/2$. Then

$$\mathrm{N}(\alpha - \beta(m + n\mathrm{i})) = \mathrm{N}(\beta)\,\mathrm{N}(\frac{\alpha}{\beta} - (m + n\mathrm{i})) = \mathrm{N}(\beta)\,\mathrm{N}(r - m + (s - n)\mathrm{i}) \leqslant \frac{1}{2}\,\mathrm{N}(\beta).$$

Now we can find *greatest common divisors* in $\mathbb{Z}[\mathrm{i}]$. In any domain, a **greatest common divisor** of two elements α and β, not both 0, is a common divisor that is divisible by every other common divisor. This greatest common divisor need not be unique. Two greatest common divisors divide each other and so are called **associates.** Conversely, the associate of a greatest common divisor is a greatest common divisor.

Problem 4. *In $\mathbb{Z}[\mathrm{i}]$, find a greatest common divisor of $7 + 6\mathrm{i}$ and $-1 + 7\mathrm{i}$.*

Solution. We can compute thus:

$$\frac{7 + 6\mathrm{i}}{-1 + 7\mathrm{i}} = \frac{(7 + 6\mathrm{i})(-1 - 7\mathrm{i})}{50} = \frac{35 - 55\mathrm{i}}{50} = \frac{7 - 11\mathrm{i}}{10} = 1 - \mathrm{i} + \frac{-3 - \mathrm{i}}{10},$$

$$7 + 6\mathrm{i} = (-1 + 7\mathrm{i})(1 - \mathrm{i}) + \frac{(-1 + 7\mathrm{i})(-3 - \mathrm{i})}{10} = (-1 + 7\mathrm{i})(1 - \mathrm{i}) + 1 - 2\mathrm{i},$$

$$\frac{-1 + 7\mathrm{i}}{1 - 2\mathrm{i}} = \frac{(-1 + 7\mathrm{i})(1 + 2\mathrm{i})}{5} = -3 + \mathrm{i}.$$

So $1-2\mathrm{i}$ is a greatest common divisor of $7+6\mathrm{i}$ and $-1+7\mathrm{i}$. The others are obtained by multiplying by the units of $\mathbb{Z}[\mathrm{i}]$, namely ± 1 and $\pm\mathrm{i}$. So the gcd's are $\pm(1-2\mathrm{i})$ and $\pm(2+\mathrm{i})$. □

6. March 11, 2008 (Tuesday)

All Euclidean domains are *principal-ideal domains,* and all principal-ideal domains are **unique-factorization domains;** therefore $\mathbb{Z}[\mathrm{i}]$ is a unique-factorization domain. But we can prove this directly, using that

$$\mathrm{N}(\xi\eta) = \mathrm{N}(\xi)\,\mathrm{N}(\eta).$$

First, an element of any domain, other than 0 or a unit, is **irreducible** if its only divisors are itself and units. Suppose α is a reducible Gaussian integer. Then

$$\alpha = \beta\gamma$$

for some β and γ, neither of which is a unit. But then $\mathrm{N}(\beta)$ and $\mathrm{N}(\gamma)$ are greater than 1, so

$$1 < \mathrm{N}(\beta) < \mathrm{N}(\alpha), \qquad 1 < \mathrm{N}(\gamma) < \mathrm{N}(\alpha).$$

Since there is no infinite strictly decreasing sequence of natural numbers, the process of factorizing the factors of α as products of non-units must terminate. Thus α can be written as a product of irreducible factors.

The definition of unique-factorization domain requires that irreducible factorizations must be unique. This means, if

$$\alpha_0\alpha_1\cdots\alpha_m = \beta_0\beta_1\cdots\beta_n,$$

where each α_i and each β_j are irreducible, then each α_i must be an associate of some β_j. To prove this for $\mathbb{Z}[\mathrm{i}]$, it is enough to show that each irreducible Gaussian integer is *prime.* In any domain, an element α (not 0 or a unit) is **prime,** provided

$$\alpha \mid \beta\gamma \mathrel{\&} \alpha \nmid \beta \implies \alpha \mid \gamma.$$

In $\mathbb{Z}[\mathrm{i}]$, suppose α is irreducible, and $\alpha \mid \beta\gamma \mathrel{\&} \alpha \nmid \beta$. Then the greatest common divisors of α and β are just the units, and we have

$$\alpha\xi + \beta\eta = 1$$

for some ξ and η in $\mathbb{Z}[\mathrm{i}]$. But then

$$\alpha\gamma\xi + \beta\gamma\eta = \gamma,$$

and since α divides the two summands on the left, it divides γ.

We now ask: What are the primes of $\mathbb{Z}[\mathrm{i}]$? Suppose π is one of them. Then π is not a unit, so $\mathrm{N}(\pi)$ has rational-prime factors. But

$$\pi\overline{\pi} = \mathrm{N}(\pi).$$

Therefore, since π is prime, we have

$$\pi \mid p$$

for some rational-prime factor of $\mathrm{N}(\pi)$. If q is another rational prime, then $ap + bq = 1$ for some rational integers a and b. Since $\pi \nmid 1$, it must be that $\pi \nmid q$. Thus p is unique.

We now consider three cases:

(i) $p = 2$;
(ii) $p \equiv 3 \pmod 4$;
(iii) $p \equiv 1 \pmod 4$.

We have

$$2 = (1 + \mathrm{i})(1 - \mathrm{i}).$$

Also, $1 \pm \mathrm{i}$ must be irreducible, since $\mathrm{N}(1 \pm \mathrm{i}) = 2$ (so if $1 \pm \mathrm{i} = \alpha\beta$, then α or β must have norm 1 and so be a unit). So we have the unique prime factorization of 2. Also $1+\mathrm{i}$ and $1-\mathrm{i}$ are associates. Hence the only prime divisors of 2 are the four associates

$$1 + \mathrm{i}, \quad 1 - \mathrm{i}, \quad -1 + \mathrm{i}, \quad -1 - \mathrm{i}.$$

Now suppose $p \equiv 3 \pmod 4$, and $\pi \mid p$. Then $\mathrm{N}(\pi) \mid \mathrm{N}(p)$, that is,

$$\pi\overline{\pi} \mid p^2.$$

So $\pi\overline{\pi}$ is either p^2 or p. But the latter is impossible, since $\mathrm{N}(\pi) = x^2 + y^2 \equiv 0, 1$, or $2 \pmod 4$. Therefore

$$\pi\overline{\pi} = p^2.$$

But $\pi\overline{\pi}$ is a prime factorization, so it is unique. Therefore π and $\overline{\pi}$ are associates of p and hence of each other. In short, p is a Gaussian prime.

Finally, suppose $p \equiv 1 \pmod 4$. Then -1 is a quadratic residue *modulo* p, so $-1 \equiv x^2 \pmod p$ for some x, that is, $p \mid 1 + x^2$, and therefore

$$p \mid (1 + x\mathrm{i})(1 - x\mathrm{i}).$$

But $(1 \pm x\mathrm{i})/p$ is *not* a Gaussian integer. Therefore p must not be a Gaussian prime. Consequently, if π is a prime factor of p, then $\mathrm{N}(\pi) = p$, that is,

$$\pi\overline{\pi} = p.$$

This is a prime factorization. Moreover, π and $\overline{\pi}$ are not associates. Indeed, $\pi = x + y\mathrm{i}$ for some rational integers x and y, so that

$$\frac{\pi}{\overline{\pi}} = \frac{(x + y\mathrm{i})^2}{p} = \frac{x^2 - y^2 + 2xy\mathrm{i}}{p}.$$

If this is a Gaussian integer, then $p \mid 2xy$, so $p \mid xy$ (since p is odd), so $p < x^2 + y^2 = p$, which is absurd. We have now shown:

Theorem 6. *The Gaussian primes are precisely the associates of the following:*

(i) $1 + \mathrm{i}$;
(ii) *the rational primes p, where $p \equiv 3 \pmod 4$;*
(iii) *α, where $\mathrm{N}(\alpha)$ is a rational prime p such that $p \equiv 1 \pmod 4$ (and two such non-associated α exist for every such p).*

If n is a positive rational integer, then the Diophantine equation

$$x^2 + y^2 = n \tag{13}$$

is soluble if and only if the equation

$$\mathrm{N}(\xi) = n \tag{14}$$

is soluble, where ξ is a Gaussian integer. Moreover, there is a bijection $(x, y) \mapsto x + y\mathrm{i}$ between the solution-sets. We now have an alternative proof, using general theory, that, when n is a rational prime p, then (13) has a solution if and only if $p = 2$ or $p \equiv 1 \pmod 4$.

Indeed, if $n = p \equiv 1 \pmod 4$, then (14) has exactly 8 solutions: the associates of π for some prime π, and the associates of $\overline{\pi}$. Then the solutions when $n = p^2$ are the associates of π^2, of $\pi\overline{\pi}$, and of $\overline{\pi}^2$, so there are 12 solutions. But if $p \neq q \equiv 1 \pmod 4$, then there are 16 solutions when $n = pq$.

Lemma 3. *The number of solutions of* (13) *is* $4(a-b)$, *where*

$$a = |\{x \in \mathbb{N} \colon x \mid n \;\&\; n \equiv 1\}|, \qquad b = |\{x \in \mathbb{N} \colon x \mid n \;\&\; n \equiv 3\}|,$$

the modulus being 4.

Proof. Exercise. □

Theorem 7. *Let* π *be the circumference of the unit circle; then*

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots$$

Proof. The area of a circle of radius r is πr^2. Hence

$$\pi r^2 \approx |\{\xi \in \mathbb{Z}[\mathrm{i}] \colon 1 \leqslant |\xi| \leqslant r\}| = \sum_{n=1}^{r^2} |\{\xi \in \mathbb{Z}[\mathrm{i}] \colon \mathrm{N}(\xi) = n\}|.$$

(See Figure 3.) By Lemma 3, to this number, each positive $4m+1$ contributes 4 for each

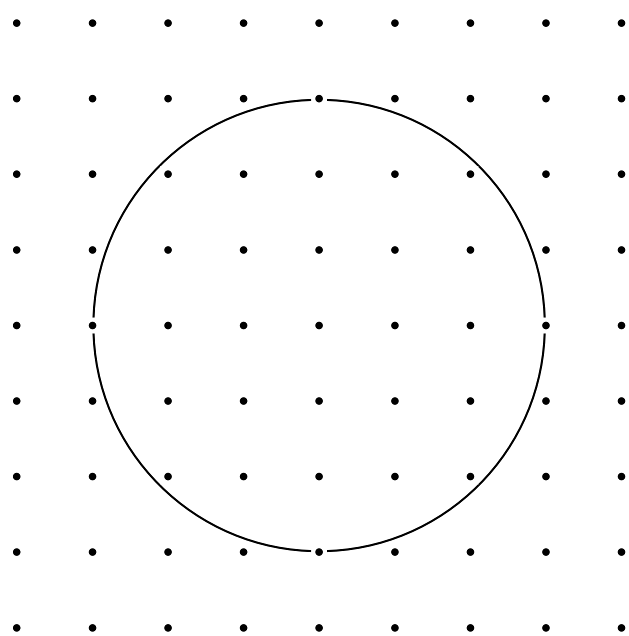

Figure 3. Estimating the area of a circle

of its multiples between 1 and r^2, while each positive $4m+3$ takes away 4 for each such multiple. Therefore

$$\frac{\pi r^2}{4} \approx \sum_{n=0}^{\infty} \left(\left[\frac{r^2}{4n+1} \right] - \left[\frac{r^2}{4n+3} \right] \right) = r^2 - \left[\frac{r^2}{3} \right] + \left[\frac{r^2}{5} \right] - \left[\frac{r^2}{7} \right] + \cdots$$

Dividing by r^2 and taking the limit yields the claim. (For details, see [5].) □

* * * * *

Recall the Pell equation (9),

$$1 = x^2 - dy^2 = (x + y\sqrt{d})(x - \sqrt{d}). \tag{15}$$

This factorization suggests looking at $\mathbb{Q}(\sqrt{d})$. Let us assume d is square-free.

We may write K for $\mathbb{Q}(\sqrt{d})$; here K is for the German *Körper* "body", the name in most languages (besides English) for a field.

On K we define $\xi \mapsto \xi'$ by

$$(a + b\sqrt{d})' = a - b\sqrt{d}.$$

When $d < 0$, this is complex conjugation. We then define:

(i) $\mathrm{Tr}(\alpha) = \alpha + \alpha'$, the **trace** of α;
(ii) $\mathrm{N}(\alpha) = \alpha\alpha'$, the **norm** of α.

These are rational numbers. Indeed, if $\alpha = a + b\sqrt{d}$, then

$$\mathrm{Tr}(\alpha) = 2a, \qquad \mathrm{N}(\alpha) = a^2 - b^2 d.$$

Also, α is a root of

$$(x - \alpha)(x - \alpha') = x^2 - (\alpha + \alpha')x + \alpha\alpha' = x^2 - \mathrm{Tr}(\alpha)x + \mathrm{N}(\alpha).$$

If $\alpha \notin \mathbb{Q}$, then this must be the **minimal polynomial** of α over $\mathbb{Q}$, that is, the polynomial of least degree with rational coefficients, and leading coefficient 1, of which α is a root. (This must exist, since the ring $\mathbb{Q}[x]$ of polynomials is a Euclidean domain with respect to degree.) Therefore, if $\alpha \in \mathbb{Q}(\sqrt{d}) \smallsetminus \mathbb{Q}$, then the following are equivalent:

(i) $\alpha^2 - m\alpha - n = 0$ for some rational integers m and n;
(ii) $\mathrm{Tr}(\alpha)$ and $\mathrm{N}(\alpha)$ are rational integers.

So we have two equivalent conditions for being a an integer of $\mathbb{Q}(\sqrt{d})$. The set of these integers can be denoted by

$$\mathfrak{O}_K.$$

This is a ring, hence an integral domain, since if $\mathrm{Tr}(\alpha_i)$ and $\mathrm{N}(\alpha_i)$ are in $\mathbb{Z}$, then so are $\mathrm{Tr}(\alpha_0 + \alpha_1)$ and $\mathrm{N}(\alpha_0 + \alpha_1)$ and $\mathrm{Tr}(\alpha_0\alpha_1)$ and $\mathrm{N}(\alpha_0\alpha_1)$ (exercise).

7. March 14, 2008 (Friday)

Moreover, $\mathrm{N}(\alpha\beta) = \mathrm{N}(\alpha)\,\mathrm{N}(\beta)$. This is simply because $(\alpha\beta)' = \alpha'\beta'$.

Immediately,

$$\mathbb{Z}[\sqrt{d}] = \{x + y\sqrt{d} \colon x, y \in \mathbb{Z}\} \subseteq \mathfrak{O}_K.$$

How about the reverse? Suppose $\alpha = a + b\sqrt{d} \in \mathfrak{O}_K$. Then $2a, a^2 - b^2 d \in \mathbb{Z}$. Consider two cases:

(i) If $a \in \mathbb{Z}$, then $b^2 d \in \mathbb{Z}$, so $b \in \mathbb{Z}$ (since d is square-free), which means $\alpha \in \mathbb{Z}[\sqrt{d}]$.
(ii) Suppose $a \notin \mathbb{Z}$. Then $2a$ is odd, so, *modulo* 4, we have $4a^2 \equiv (2a)^2 \equiv 1$. But also $4a^2 - 4b^2 d \equiv 0$, so that $(2b)^2 d \equiv 4b^2 d \equiv 4a^2 \equiv 1$. Since $(2b)^2 \equiv 0$ or 1, we conclude $(2b)^2 \equiv 1$, hence $d \equiv 1$.

But now suppose $d \equiv 1$. We have shown, if $\alpha \notin \mathbb{Z}[\sqrt{d}]$, that $2a$ and $2b$ are odd, so that

$$\alpha = a - b + b + b\sqrt{d} = a - b + 2b \cdot \frac{1 + \sqrt{d}}{2} \in \mathbb{Z}\Big[\frac{1 + \sqrt{d}}{2}\Big].$$

Conversely, if $\alpha = (1 + \sqrt{d})/2$, then $(2\alpha - 1)^2 = d$, so $4\alpha^2 - 4\alpha + 1 - d = 0$, hence $\alpha^2 - \alpha + (1 - d)/4 = 0$, which means $\alpha \in \mathfrak{O}_K$ (since $d \equiv 1 \pmod 4$). Thus:

Theorem 8. *The ring of integers of K is given by*

$$\mathfrak{O}_K = \begin{cases} \mathbb{Z}[\sqrt{d}], & \text{if } d \equiv 2 \text{ or } 3 \pmod 4; \\ \mathbb{Z}\Big[\dfrac{1+\sqrt{d}}{2}\Big], & \text{if } d \equiv 1 \pmod 4. \end{cases}$$

* * * * *

Assuming $a, b, c \in \mathbb{Q}$, let

$$f(x,y) = ax^2 + bxy + cy^2; \tag{16}$$

this is a **binary quadratic form.** We shall investigate the rational-integral solutions of

$$f(x,y) = m,$$

where $m \in \mathbb{Q}$. The Pell equation (15) is a special case. We can factorize f over a quadratic field by completing the square:

$$\begin{aligned} f(x,y) &= a\Big(x^2 + \frac{b}{a}\cdot xy + \frac{b^2}{4a^2}\cdot y^2\Big) - \Big(\frac{b^2}{4a} - c\Big)y^2 \\ &= a\Big(x + \frac{b}{2a}\cdot y\Big)^2 - \Big(\frac{b^2}{4a} - c\Big)y^2 \\ &= \frac{1}{a}\Big(ax + \frac{b}{2}\cdot y\Big)^2 - \frac{1}{a}\Big(\frac{b^2}{4} - ac\Big)y^2 \\ &= \frac{1}{a}\Big[\Big(ax + \frac{b}{2}\cdot y\Big)^2 - \frac{D}{4}\cdot y^2\Big] \end{aligned}$$

where $D = b^2 - 4ac$, the **discriminant** of f. Then

$$\begin{aligned} f(x,y) &= \frac{1}{a}\Big(ax + \frac{b}{2}\cdot y + \frac{\sqrt{D}}{2}\cdot y\Big)\Big(ax + \frac{b}{2}\cdot y - \frac{\sqrt{D}}{2}\cdot y\Big) \\ &= \frac{1}{a}\Big(ax + \frac{b+\sqrt{D}}{2}\cdot y\Big)\Big(ax + \frac{b-\sqrt{D}}{2}\cdot y\Big). \end{aligned}$$

We can write D as $s^2 d$, where $s \in \mathbb{Q}$, but d is a square-free rational integer. Working in $\mathbb{Q}(\sqrt{d})$, letting

$$\alpha = a, \qquad \beta = \frac{b+\sqrt{D}}{2} = \frac{b+s\sqrt{d}}{2},$$

we have

$$f(x,y) = \frac{1}{a}(\alpha x + \beta y)(\alpha' x + \beta' y) = \frac{1}{a}\,\mathrm{N}(\alpha x + \beta y).$$

Moreover, α and β are linearly independent over $\mathbb{Q}$; that is, the only rational solution to $\alpha x + \beta y = 0$ is $(0, 0)$.

For any α and β in K (which is $\mathbb{Q}(\sqrt{d})$), let us denote the set $\{\alpha x + \beta y : x, y \in \mathbb{Z}\}$ of all rational-integral linear combinations of α and β by

$$\mathbb{Z}\alpha + \mathbb{Z}\beta \quad \text{or} \quad \langle\alpha, \beta\rangle.$$

If α and β are linearly independent over $\mathbb{Q}$, then $\langle\alpha, \beta\rangle$ is a **lattice** in K: that is, $\langle\alpha, \beta\rangle$ is a *free abelian subgroup* of K, and the number of generators is the dimension $[K : \mathbb{Q}]$.

For example, as a group, $\mathfrak{O}_K$ is the lattice $\langle 1, \omega\rangle$, where

$$\omega = \begin{cases} \sqrt{d}, & \text{if } d \equiv 2 \text{ or } 3 \pmod 4; \\ \dfrac{1+\sqrt{d}}{2}, & \text{if } d \equiv 1 \pmod 4). \end{cases}$$

Henceforth ω will always have this meaning.

In general, if Λ is a lattice in K, let

$$\operatorname{End}(\Lambda) = \{\xi \in \mathbb{C} \colon \xi\Lambda \subseteq \Lambda\}.$$

This set is a sub-ring of K and and can be understood as the ring of **endomorphisms** of the abelian group Λ. That is, the function $\xi \mapsto \alpha\xi$ is an endomorphism of Λ if and only if $\alpha \in \operatorname{End}(\Lambda)$. For example, if $\Lambda = \langle 1, \mathrm{i}\rangle$ in $\mathbb{Q}(\mathrm{i})$, then $\operatorname{End}(\Lambda) = \langle 1, \mathrm{i}\rangle$.

But suppose $\Lambda = \langle 1, \tau\rangle$, where

$$\tau = \frac{-1+\sqrt{-7}}{4}.$$

Then $(4\tau+1)^2 = -7$, so $16\tau^2 + 8\tau + 8 = 0$, or $2\tau^2 + \tau + 1 = 0$. Suppose $x + y\tau \in \operatorname{End}(\Lambda)$. Equivalently, Λ contains both $x + y\tau$ and $(x+y\tau)\tau$. But

$$(x+y\tau)\tau = x\tau + y\tau^2 = x\tau + y\frac{-\tau-1}{2} = -\frac{y}{2} + \Big(x - \frac{y}{2}\Big)\tau.$$

So y must be even. Conversely, this is enough to ensure $x + y\tau \in \operatorname{End}(\Lambda)$. Thus

$$\operatorname{End}(\Lambda) = \langle 1, 2\tau\rangle.$$

See Figure 4.

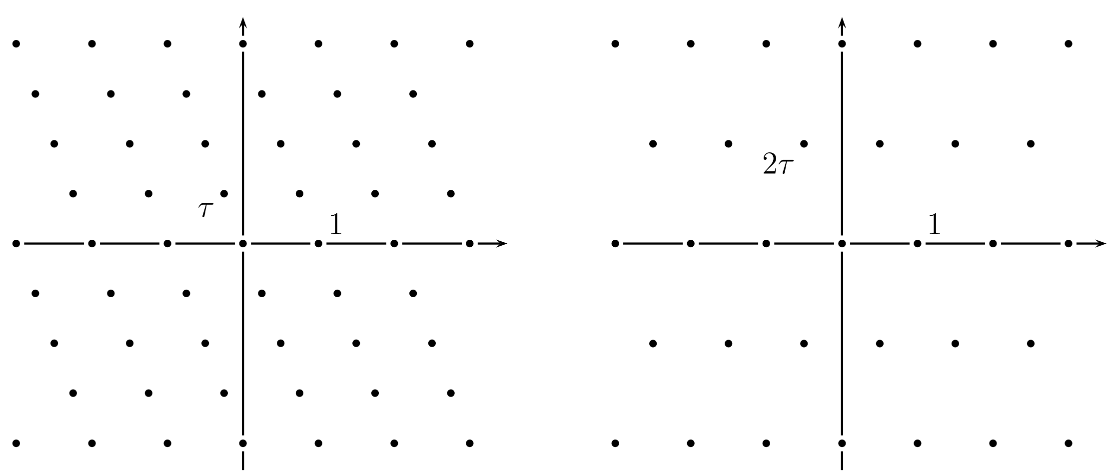

FIGURE 4. A lattice and its endomorphisms

8. March 18, 2008 (Tuesday)

To give a sense for where things may lead (though not in this course; but see for example [7]): In a more general sense, a **lattice** is a subgroup $\mathbb{Z}\alpha + \mathbb{Z}\beta$ or $\langle\alpha, \beta\rangle$ of $\mathbb{C}$ such that $\alpha \neq 0$ and $\beta/\alpha \notin \mathbb{R}$. Let Λ be such a lattice. Then we can form the quotient group $\mathbb{C}/\Lambda$. Geometrically, this is the parallelogram with vertices 0, α, β, and $\alpha+\beta$ (as in Figure 5), with opposite edges identified: thus it is a **torus.** There is a function $\wp$,

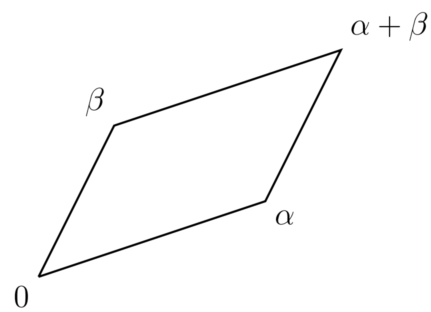

FIGURE 5. A fundamental parallelogram of a lattice

that is, $\wp_\Lambda$: the **Weierstraß function,** given by

$$\wp(z)=\frac{1}{z^2}+\sum_{\zeta\in\Lambda\smallsetminus\{0\}}\Bigl(\frac{1}{(z-\zeta)^2}-\frac{1}{\zeta^2}\Bigr).$$

This is **doubly periodic,** with period Λ: that is,

$$\zeta\in\Lambda\iff\wp(z+\zeta)=\wp(z)\text{ for all }z.$$

Hence $\wp$ is well-defined as a function on the torus $\mathbb{C}/\Lambda$. There are g_2 and g_3 (depending on Λ) in $\mathbb{C}$ such that $(\wp(z),\wp'(z))$ solves the equation

$$y^2=4x^3-g_2x-g_3.$$

This equation defines an **elliptic curve** (Figure 6). This curve can be made into a group

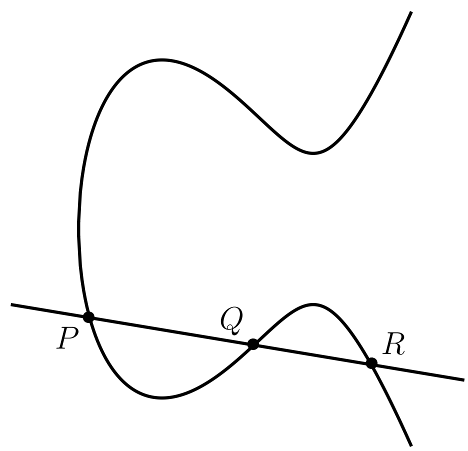

FIGURE 6. An elliptic curve

by the rule that, if a straight line meets the curve in P, Q, and R, then $P+Q+R=0$. (Also, a vertical line meets the curve at the 'point at infinity', which is defined to be the 0 of the group.) Then $(\wp,\wp')$ is an isomorphism from $\mathbb{C}/\Lambda$ to the elliptic curve.

An analytic endomorphism of $\mathbb{C}/\Lambda$ is a function $z\mapsto\alpha z$, where $\alpha\in\mathbb{C}$, such that $\alpha\Lambda\subseteq\Lambda$. The set of these α is what we are calling $\operatorname{End}(\Lambda)$. Always $\mathbb{Z}\subseteq\operatorname{End}(\Lambda)$. You can show that $\mathbb{Z}=\operatorname{End}(\Lambda)$ if and only if β/α is not quadratic—not the root of some x^2+bx+c, where $b,c\in\mathbb{Q}$.

* * * * *

We are interested in the quadratic case. Again suppose $K = \mathbb{Q}(\sqrt{d})$, where d is squarefree. Say $\alpha, \beta \in K$, and $\langle\alpha, \beta\rangle$ is a lattice Λ. In particular then, $\alpha \neq 0$ and $\beta/\alpha \notin \mathbb{Q}$. Every element $\alpha x + \beta y$ of Λ is a matrix product:

$$\alpha x + \beta y = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

Then $\langle\gamma, \delta\rangle \subseteq \langle\alpha, \beta\rangle$ if and only if

$$\begin{pmatrix} \gamma \\ \delta \end{pmatrix} = \begin{pmatrix} x & y \\ z & w \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

for some x, y, z, and w in $\mathbb{Z}$. Then $\langle\gamma, \delta\rangle = \langle\alpha, \beta\rangle$ if and only if

$$\begin{pmatrix} \gamma \\ \delta \end{pmatrix} = M \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

for some *invertible* matrix M over $\mathbb{Z}$: this means $\det M = \pm 1$.

Along with the sub-ring $\operatorname{End}(\Lambda)$ of K, we have the sub-ring $\mathfrak{O}_K$. What is the relation between the two rings?

Lemma 4. $\operatorname{End}(\Lambda) \subseteq \mathfrak{O}_K$.

Proof. Suppose $\gamma \in \operatorname{End}(\Lambda)$. Then there are x, y, z, and w in $\mathbb{Z}$ such that

$$\begin{pmatrix} x & y \\ z & w \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \gamma \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \gamma & 0 \\ 0 & \gamma \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \gamma - x & -y \\ -z & \gamma - w \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

Hence the last square matrix is not invertible over any field, so its determinant is 0: that is,

$$0 = (\gamma - x)(\gamma - w) - yz = \gamma^2 - (x + w)\gamma + xw - yz.$$

Since the coefficients here belong to $\mathbb{Z}$, we have that $\gamma \in \mathfrak{O}_K$. □

* * * * *

Problem 5. *Solve the Pell equation*

$$x^2 - 14y^2 = 1. \tag{17}$$

Solution. We first find the continued fraction expansion of $\sqrt{14}$ by our algorithm:

$$\begin{array}{rlll}
 & & a_0 = 3, & \xi_0 = \sqrt{14} - 3; \\
\dfrac{1}{\sqrt{14} - 3} & = \dfrac{\sqrt{14} + 3}{5}, & a_1 = 1, & \xi_1 = \dfrac{\sqrt{14} - 2}{5}; \\
\dfrac{5}{\sqrt{14} - 2} & = \dfrac{\sqrt{14} + 2}{2}, & a_2 = 2, & \xi_2 = \dfrac{\sqrt{14} - 2}{2}; \\
\dfrac{2}{\sqrt{14} - 2} & = \dfrac{\sqrt{14} + 2}{5}, & a_3 = 1, & \xi_3 = \dfrac{\sqrt{14} - 3}{5}; \\
\dfrac{5}{\sqrt{14} - 3} & = \sqrt{14} + 5, & a_4 = 6, & \xi_4 = \sqrt{14} - 3 = \xi_0;
\end{array}$$

therefore

$$\sqrt{14} = [3; \overline{1, 2, 1, 6}].$$

For the convergents p_n/q_n, we have

$$\frac{p_0}{q_0} = \frac{3}{1}, \qquad \frac{p_1}{q_1} = \frac{4}{1}, \qquad \frac{p_n}{q_n} = \frac{a_n p_{n-1} + p_{n-2}}{a_n q_{n-1} + q_{n-2}},$$

so the list is

$$\frac{3}{1}, \quad \frac{4}{1}, \quad \frac{11}{3}, \quad \frac{15}{4}, \quad \frac{101}{27}, \quad \dots$$

Check for a solution to (17):

$$\begin{aligned} 3^2 - 14 \cdot 1^2 &= -5, \\ 4^2 - 14 \cdot 1^2 &= 2, \\ 11^2 - 14 \cdot 3^2 &= 121 - 126 = -5, \\ 15^2 - 14 \cdot 4^2 &= 225 - (15-1)(15+1) = 1. \end{aligned}$$

Then $15/4 = [3; 1, 2, 1]$, and $(15, 4)$ solves (17). This is the positive solution (a, b) for which $a + b\sqrt{14}$ is least: we shall prove this later, but meanwhile you can check it by trying all pairs (a, b) such that $0 < a < 15$ and $0 < b < 4$. Then every positive solution is

$$(a_n, b_n), \text{ where } a_n + b_n\sqrt{14} = (15 + 4\sqrt{14})^n.$$

Moreover, each of these solutions is (p_{4n+3}, q_{4n+3}), and

$$\frac{p_{4n+3}}{q_{4n+3}} = [3; \underbrace{\underbrace{1, 2, 1, 6}, \dots, \underbrace{1, 2, 1, 6}}_{n}, 1, 2, 1]$$

Indeed, if (k, ℓ) is a solution, then by the computation

$$(15 + 4\sqrt{14})(k + \ell\sqrt{14}) = 15k + 56\ell + (4k + 15\ell)\sqrt{14},$$

we have that $(15k + 56\ell, 4k + 15\ell)$ is a solution. But also, writing (p_{4n+3}, q_{4n+3}) as (a, b), we have

$$\frac{p_{4n+7}}{q_{4n+7}} = \left[3; 1, 2, 1, 3 + \frac{a}{b}\right] = 3 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{3 + \cfrac{a}{b}}}}} = 3 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{b}{a + 3b}}}}$$

$$= 3 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{a + 3b}{a + 4b}}} = 3 + \cfrac{1}{1 + \cfrac{a + 4b}{3a + 11b}} = 3 + \frac{3a + 11b}{4a + 15b} = \frac{15a + 56b}{4a + 15b}.$$

By induction, our claim is proved. □

The expansion $[3; \overline{1, 2, 1, 6}]$ of $\sqrt{14}$ has the period $(1, 2, 1, 6)$ of length 4, which is even. But

$$\sqrt{13} = [3; \overline{1, 1, 1, 1, 6}]$$

with a period of odd length 5. The convergents p_n/q_n of $\sqrt{d}$ are alternately above and below $\sqrt{d}$ (assuming this is irrational); in particular, the convergents p_{2n}/q_{2n} are below. Therefore $[3; 1, 1, 1, 1]$ cannot provide a solution to $x^2 - 13y^2 = 1$. But

$$[3; 1, 1, 1, 1, 6, 1, 1, 1, 1]$$

does provide the fundamental solution that generates the others: the solutions here are p_{10n+9}/q_{10n+9}.

9. March 25, 2008 (Tuesday)

A problem on last night's examination was to find solutions to the Diophantine equation

$$2x^2 - 3y^2 = 2. \tag{18}$$

Let us define

$$\begin{aligned} f(x, y) &= 2x^2 - 3y^2 \\ &= 2(x^2 - \frac{3}{2}y^2) \\ &= 2(x + \sqrt{\frac{3}{2}} \cdot y)(x - \sqrt{\frac{3}{2}} \cdot y) \\ &= \frac{1}{2}(2x + \sqrt{6} \cdot y)(2x - \sqrt{6} \cdot y). \end{aligned}$$

Working in $\mathbb{Q}(\sqrt{6})$, we have

$$f(x, y) = \frac{1}{2}\,\mathrm{N}(2x + \sqrt{6} \cdot y) = \frac{1}{2}\,\mathrm{N}(\alpha x + \beta y),$$

where $\alpha = 2$ and $\beta = \sqrt{6}$. We have a bijection $(x, y) \mapsto \alpha x + \beta y$ between:

(i) the solution-set of (18);
(ii) the set of ξ in $\langle \alpha, \beta \rangle$ such that $\mathrm{N}(\xi) = 4$.

In particular, $(5, 4)$ is a solution of (18), and $\mathrm{N}(5\alpha + 4\beta) = 4$. Then other solutions to $\mathrm{N}(\xi) = 4$ include $\varepsilon \cdot (5\alpha + 4\beta)$, provided:

(i) $\mathrm{N}(\varepsilon) = 1$;
(ii) $\varepsilon \cdot (5\alpha + 4\beta) \in \langle \alpha, \beta \rangle$,—and this is achieved if $\varepsilon\langle \alpha, \beta \rangle \subseteq \langle \alpha, \beta \rangle$, that is, $\varepsilon \in \mathrm{End}(\langle \alpha, \beta \rangle)$.

*   *   *   *   *

Let $f(x, y) = ax^2 + bxy + cy^2$ for some a, b, and c in $\mathbb{Q}$ (as in (16)). Again, the discriminant of f is given by

$$D = b^2 - 4ac = s^2 d,$$

where $s \in \mathbb{Q} \smallsetminus \{0\}$, $d \in \mathbb{Z}$, and d is square-free or 0. Let us assume $d \neq 0$ or 1: equivalently, $\sqrt{D} \notin \mathbb{Q}$. Then $a \neq 0$. By the quadratic formula,

$$\begin{aligned} f(x,y) &= a\Bigl(x - \frac{-b+\sqrt{D}}{2a}y\Bigr)\Bigl(x - \frac{-b-\sqrt{D}}{2a}y\Bigr) \\ &= \frac{1}{a}\Bigl(ax + \frac{b-\sqrt{D}}{2}y\Bigr)\Bigl(x + \frac{b+\sqrt{D}}{2}y\Bigr) \\ &= \frac{1}{a}(\alpha' x + \beta' y)(\alpha x + \beta y) \\ &= \frac{1}{a}\,\mathrm{N}(\alpha x + \beta y), \end{aligned}$$

where $\alpha = a$ and $\beta = (b + \sqrt{D})/2$, and the computations are in K, where $K = \mathbb{Q}(\sqrt{d})$. Since $\sqrt{D}$ is irrational, we have $\beta/\alpha \notin \mathbb{Q}$, that is, α and β are linearly independent over $\mathbb{Q}$; equivalently, $\{\alpha, \beta\}$ is a basis of K over $\mathbb{Q}$.

Now suppose conversely $\alpha, \beta \in K$. Let

$$f(x,y) = \mathrm{N}(\alpha x + \beta y) = (\alpha x + \beta y)(\alpha' x + \beta' y) = \mathrm{N}(\alpha)x + \mathrm{Tr}(\alpha\beta')xy + \mathrm{N}(\beta)y^2.$$

Then

$$D = \mathrm{Tr}(\alpha\beta')^2 - 4\,\mathrm{N}(\alpha\beta) = (\alpha\beta' + \alpha'\beta)^2 - 4\alpha\beta\alpha'\beta' = (\alpha\beta' - \alpha'\beta)^2 = \begin{vmatrix} \alpha & \alpha' \\ \beta & \beta' \end{vmatrix}^2. \tag{19}$$

Lemma 5. *Let K be a quadratic field $\mathbb{Q}(\sqrt{d})$, where d is a square-free rational integer (different from 1). Let $\alpha, \beta \in K$, and let D be the discriminant of the quadratic form $\mathrm{N}(\alpha x + \beta y)$. Then $D = s^2 d$ for some s in $\mathbb{Q}$. The following are equivalent:*

(i) *$D \neq 0$;*
(ii) *α and β are linearly independent over $\mathbb{Q}$;*
(iii) *$\sqrt{D}$ is irrational.*

Proof. If $\alpha = 0$, then (i), (ii), and (iii) all fail. Suppose $\alpha \neq 0$. Then we can write β/α as $r + t\sqrt{d}$ for some r and t in $\mathbb{Q}$. From (19), we have

$$D = \Bigl(\alpha\alpha'\Bigl(\frac{\beta'}{\alpha'} - \frac{\beta}{\alpha}\Bigr)\Bigr)^2 = \mathrm{N}(\alpha)^2\Bigl(\Bigl(\frac{\beta}{\alpha}\Bigr)' - \Bigl(\frac{\beta}{\alpha}\Bigr)\Bigr)^2 = \mathrm{N}(\alpha)^2 \cdot 4t^2 d = (2t\,\mathrm{N}(\alpha))^2 \cdot d.$$

Since $2t\,\mathrm{N}(\alpha) \in \mathbb{Q}$, we have

$$\sqrt{D} \in \mathbb{Q} \iff D = 0 \iff t = 0 \iff \beta/\alpha \in \mathbb{Q}.$$

Thus, (i), (ii), and (iii) are again equivalent. □

We have observed that two lattices $\langle \alpha, \beta \rangle$ and $\langle \gamma, \delta \rangle$ of K are the same lattice Λ if and only if

$$\begin{pmatrix} \gamma \\ \delta \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

for some a, b, c, and d in $\mathbb{Z}$ such that $ad - bc = \pm 1$. In this case,

$$\begin{pmatrix} \gamma & \gamma' \\ \delta & \delta' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \alpha & \alpha' \\ \beta & \beta' \end{pmatrix},$$

so that

$$\begin{vmatrix} \gamma & \gamma' \\ \delta & \delta' \end{vmatrix}^2 = \begin{vmatrix} \alpha & \alpha' \\ \beta & \beta' \end{vmatrix}^2.$$

Then this number is the **discriminant** of Λ, and we write

$$\Delta(\Lambda) = \Delta(\alpha, \beta) = \begin{vmatrix} \alpha & \alpha' \\ \beta & \beta' \end{vmatrix}^2.$$

So this is the discriminant of the quadratic forms $\mathrm{N}(\alpha x + \beta y)$ and $\mathrm{N}(\gamma x + \delta y)$.

Lemma 6. *Suppose $\alpha, \beta \in K$. Then*

(i) $\Delta(\alpha, \beta) \in \mathbb{Q}$;
(ii) $\alpha, \beta \in \mathfrak{O}_K \implies \Delta(\alpha, \beta) \in \mathbb{Z}$;
(iii) $\{\alpha, \beta\}$ *is a basis for K if and only if* $\Delta(\alpha, \beta) \neq 0$.

Proof. We have (i) and (iii) by Lemma 5. As for (ii), if $\alpha, \beta \in \mathfrak{O}_K$, then $\Delta(\alpha, \beta) \in \mathfrak{O}_K \cap \mathbb{Q} = \mathbb{Z}$ (exercise). □

* * * * *

Suppose

$$f(x, y) = 2x^2 + 6xy + 3y^2.$$

Then $D = 36 - 24 = 12 = 2^2 \cdot 3$. Also

$$\begin{aligned} f(x, y) &= 2\Big(x^2 + 3xy + \frac{3}{2}y^2\Big) = 2\Big(x - \frac{-3 + 2\sqrt{3}}{2}y\Big)\Big(x - \frac{-3 + 2\sqrt{3}}{2}y\Big) \\ &= \frac{1}{2}\big(2x + (3 + 2\sqrt{3})y\big)\big(2x + (3 - 2\sqrt{3})y\big). \end{aligned}$$

So we have a bijection $(x, y) \mapsto 2x + (3 + 2\sqrt{3})y$ between $\{(x, y) \in \mathbb{Z} \times \mathbb{Z} \colon f(x, y) = m\}$ and $\{\xi \in \langle 2, 3 + 2\sqrt{3}\rangle \colon \mathrm{N}(\xi) = 2m\}$, where the norm is computed in $\mathbb{Q}(\sqrt{3})$. We can write the form as a matrix product:

$$f(x, y) = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 3 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Then making a change of variable, as by

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix},$$

means forming a new product

$$\begin{pmatrix} u & v \end{pmatrix} \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 3 & 3 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}.$$

Such a change may be useful particularly if what we want to understand is the possible values of $f(x, y)$.

* * * * *

As usual, let d be square-free, and different from 1; and $K = \mathbb{Q}(\sqrt{d})$.

Lemma 7. *Let L be a subset of K. Then L is a lattice of K if and only if:*

(i) *L is an additive subgroup of K (that is, K contains 0 and is closed under addition and subtraction);*
(ii) *as a vector-space, K is spanned by L (over $\mathbb{Q}$);*

(iii) $nL \subseteq \mathfrak{O}_K$ *for some n in $\mathbb{Z} \smallsetminus \{0\}$.*

Proof. Suppose L is a lattice of K. Then (i) and (ii) hold by definition of lattice. Also $L = \langle\alpha,\beta\rangle$ for some α and β in K. But $\mathfrak{O}_K$ is a lattice $\langle 1,\omega\rangle$ for some ω. In particular, $(1,\omega)$ spans K. So $\alpha = k + \ell\omega$ and $\beta = r + s\omega$ for some k, ℓ, r, and s in $\mathbb{Q}$. Let n be a common multiple of their denominators. Then $n\alpha, n\beta \in \mathfrak{O}_K$, so $nL \subseteq \mathfrak{O}_K$.

Now suppose conversely that (i), (ii), and (iii) hold. Then L contains α and β such that $\{\alpha,\beta\}$ is a basis for K; and there is n in $\mathbb{Z} \smallsetminus \{0\}$ such that, for every such basis, $n\alpha, n\beta \in \mathfrak{O}_K$. By the last lemma, this means $\Delta(n\alpha, n\beta) \in \mathbb{Z}$. Also $\Delta(\alpha,\beta) \neq 0$. So we may suppose α and β have been chosen from L so as to minimize $|\Delta(n\alpha,n\beta)|$, which is $n^4|\Delta(\alpha,\beta)|$. We shall show $L = \langle\alpha,\beta\rangle$. Suppose $\gamma \in L$. Then $\gamma \in K$, so

$$\gamma = \alpha r + \beta s$$

for some r and s in $\mathbb{Q}$. We want to show $r, s \in \mathbb{Z}$. Since

$$\gamma - \alpha[r] = \alpha(r - [r]) + \beta s,$$

we have

$$\begin{aligned}\Delta(\gamma - \alpha[r], \beta) &= \begin{vmatrix} \gamma - \alpha[r] & \gamma' - \alpha'[r] \\ \beta & \beta' \end{vmatrix}^2 = \begin{vmatrix} \alpha(r-[r]) + \beta s & \alpha'(r-[r]) + \beta' s \\ \beta & \beta' \end{vmatrix}^2 \\ &= \begin{vmatrix} \alpha(r-[r]) & \alpha'(r-[r]) \\ \beta & \beta' \end{vmatrix}^2 = (r-[r])^2 \begin{vmatrix} \alpha & \alpha' \\ \beta & \beta' \end{vmatrix}^2 = (r-[r])^2\Delta(\alpha,\beta).\end{aligned}$$

By minimality of $|\Delta(\alpha,\beta)|$, we must have $r - [r] = 0$, so $r \in \mathbb{Z}$. By symmetry, $s \in \mathbb{Z}$. □

10. March 28, 2008 (Friday)

If Λ is a lattice of K, then the ring $\operatorname{End}(\Lambda)$ is also called the **order** of Λ and denoted by

$$\mathfrak{O}_\Lambda.$$

By Lemma 4, we know that this is a sub-ring of $\mathfrak{O}_K$.

Lemma 8. *Let Λ be a lattice of K. Then $\mathfrak{O}_\Lambda$ is also a lattice of K.*

Proof. By Lemma 7, it is enough to show that $\mathfrak{O}_\Lambda$ spans K over $\mathbb{Q}$. Write Λ as $\langle\alpha,\beta\rangle$. Let $\gamma \in K$. Then

$$\gamma\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \begin{pmatrix} r & s \\ t & u\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix}$$

for some rational numbers r, s, t, and u. Let n be a common multiple of their denominators. Then $n\gamma\Lambda \subseteq \Lambda$, that is, $n\gamma \in \mathfrak{O}_\Lambda$. But $\gamma = (1/n)n\gamma$. □

Theorem 9. $\mathfrak{O}_\Lambda = \langle 1, c\omega\rangle$ *for some positive integer c.*

Proof. We know $1 \in \mathfrak{O}_\Lambda$ and $\mathfrak{O}_\Lambda \subseteq \langle 1,\omega\rangle$. Since $\mathfrak{O}_\Lambda$ is a lattice, we must therefore have $m + n\omega \in \mathfrak{O}_\Lambda$ for some integers m and n, where $n \neq 0$. Hence $n\omega \in \mathfrak{O}_\Lambda$. Let c be the least positive integer such that $c\omega \in \mathfrak{O}_\Lambda$. Then $\langle 1, c\omega\rangle \subseteq \mathfrak{O}_\Lambda$. Conversely, suppose $m + n\omega \in \mathfrak{O}_\Lambda$. Then $n\omega \in \mathfrak{O}_\Lambda$, hence $\gcd(c,n)\omega \in \mathfrak{O}_\Lambda$. By minimality of c, we must have $\gcd(c,n) = c$, so $c \mid n$. Thus $\mathfrak{O}_\Lambda \subseteq \langle 1, c\omega\rangle$. □

The number c in the theorem is called the **conductor** of $\mathfrak{O}_\Lambda$.

Lemma 9. $\mathfrak{O}_{\gamma\Lambda} = \mathfrak{O}_\Lambda$ *for all non-zero γ in K.*

Proof. Since $\xi \mapsto \gamma\xi$ is a bijection from K to itself, we have $\xi\Lambda \subseteq \Lambda \iff \xi\gamma\Lambda \subseteq \gamma\Lambda$. $\square$

In looking for $\mathfrak{O}_\Lambda$, we may therefore assume that $\Lambda = \langle 1, \tau\rangle$ for some τ. Then

$$a\tau^2 + b\tau + c = 0$$

for some a, b, and c in $\mathbb{Z}$, where $\gcd(a, b, c) = 1$ and $a > 0$. Then

$$a\tau^2 = -b\tau - c,$$

which shows $\langle 1, a\tau\rangle \subseteq \mathfrak{O}_\Lambda$. That this inclusion is an equality can be seen in some examples. If $b = 0$ and $c = 1$, then we may assume $\tau = \mathrm{i}/\sqrt{a}$: see Figure 7. If $b = -1$

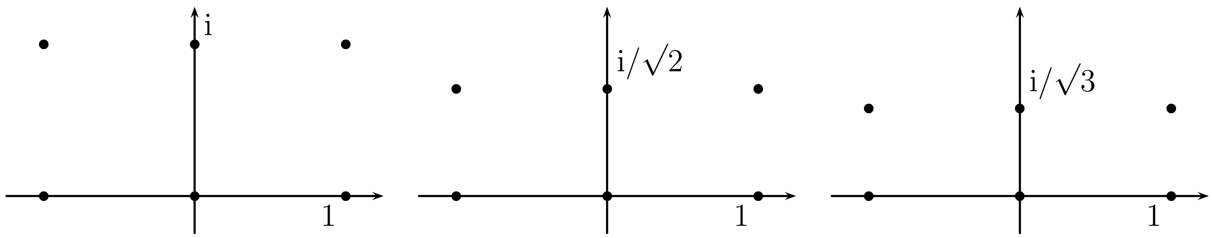

FIGURE 7. Lattices $\langle 1, \mathrm{i}/\sqrt{a}\rangle$

and $c = 1$, then $|\tau| = 1/\sqrt{a}$, and we may assume $\tau = (1 + \mathrm{i}\sqrt{4a-1})/2a$: see Figure 8.

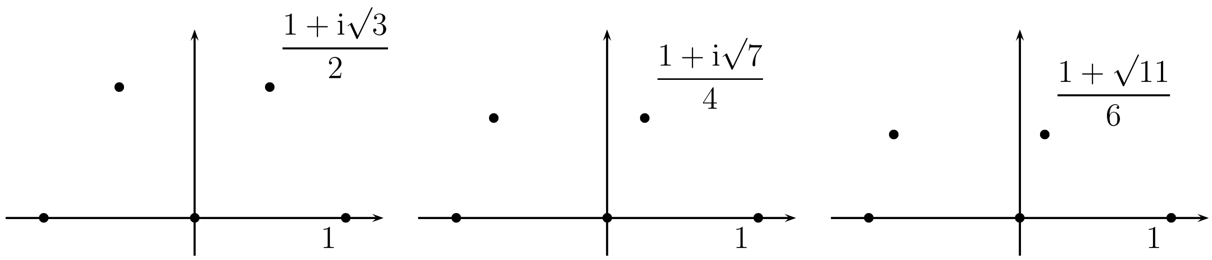

FIGURE 8. Lattices $\langle 1, (1 + \mathrm{i}\sqrt{4a-1})/2a\rangle$

Theorem 10. *Suppose* $\Lambda = \langle \alpha, \beta\rangle$. *Let* $\tau = \beta/\alpha$, *so that*

$$a\tau^2 + b\tau + c = 0$$

for some a, b, *and* c *in* $\mathbb{Z}$, *where* $\gcd(a, b, c) = 1$. *Then*

$$\mathfrak{O}_\Lambda = \langle 1, a\tau\rangle.$$

Proof. We have the following equivalences:

$$\begin{aligned}
\theta \in \mathfrak{O}_\Lambda &\iff \theta\langle 1,\tau\rangle \subseteq \langle 1,\tau\rangle\\
&\iff \theta \in \langle 1,\tau\rangle \;\&\; \theta\tau \in \langle 1,\tau\rangle\\
&\iff \theta = x + y\tau \;\&\; x\tau + y\tau^2 \in \langle 1,\tau\rangle \text{ for some } x \text{ and } y \text{ in } \mathbb{Z}\\
&\iff \theta = x + y\tau \;\&\; y\tau^2 \in \langle 1,\tau\rangle \text{ for some } x \text{ and } y \text{ in } \mathbb{Z}\\
&\iff \theta = x + y\tau \;\&\; \frac{yb}{a}\tau + \frac{yc}{a} \in \langle 1,\tau\rangle \text{ for some } x \text{ and } y \text{ in } \mathbb{Z}\\
&\iff \theta = x + y\tau \;\&\; a \mid yb \;\&\; a \mid yc \text{ for some } x \text{ and } y \text{ in } \mathbb{Z}\\
&\iff \theta = x + y\tau \;\&\; a \mid y \text{ for some } x \text{ and } y \text{ in } \mathbb{Z}.
\end{aligned}$$

In short, $\theta \in \mathfrak{O}_\Lambda \iff \theta \in \langle 1, a\tau\rangle$. $\square$

11. April 1, 2008 (Tuesday)

What then is the conductor of $\mathfrak{O}_\Lambda$? Since $\tau \in K$, we have

$$\tau = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{-b \pm s\sqrt{d}}{2a}$$

for some s in $\mathbb{Z}$. Hence

$$\mathfrak{O}_\Lambda = \Big\langle 1, \frac{-b \pm s\sqrt{d}}{2} \Big\rangle.$$

But we have

$$s^2 d \equiv b^2 \pmod 4.$$

If $d \equiv 2$ or 3, then (since squares are congruent to 0 or 1), we must have $s^2 \equiv 0$, so s is even, and also b is even, so that

$$\mathfrak{O}_\Lambda = \Big\langle 1, \frac{s}{2}\sqrt{d} \Big\rangle = \Big\langle 1, \frac{s}{2}\omega \Big\rangle.$$

If $d \equiv 1$, then $s^2 \equiv b^2$, so $b \pm s$ is even, and hence

$$\mathfrak{O}_\Lambda = \Big\langle 1, \frac{-b \mp s \pm s \pm s\sqrt{d}}{2} \Big\rangle = \Big\langle 1, \pm s \frac{1 \pm \sqrt{d}}{2} \Big\rangle;$$

this is either $\langle 1, s\omega \rangle$ immediately, or $\langle 1, -s\omega' \rangle$, which is $\langle 1, s\omega - s \rangle$, which is $\langle 1, s\omega \rangle$.

* * * * *

We now ask which elements of $\mathfrak{O}_\Lambda$ satisfy $\mathrm{N}(\xi) = 1$.

Lemma 10. *The units of $\mathfrak{O}_\Lambda$ are just those elements that satisfy* $\mathrm{N}(\xi) = \pm 1$.

Proof. We know $\mathfrak{O}_\Lambda \subseteq \mathfrak{O}_K$, so $\mathrm{N}(\alpha) \in \mathbb{Z}$ for all α in $\mathfrak{O}_\Lambda$. Suppose α is a unit of $\mathfrak{O}_\Lambda$. Then $\alpha \neq 0$, and $\alpha^{-1} \in \mathfrak{O}_\Lambda$. But $1 = \mathrm{N}(1) = \mathrm{N}(\alpha\alpha^{-1}) = \mathrm{N}(\alpha)\,\mathrm{N}(\alpha^{-1})$, and since these factors are in $\mathbb{Z}$, we have that $\mathrm{N}(\alpha)$ is a unit in $\mathbb{Z}$, that is, $\mathrm{N}(\alpha) = \pm 1$.

Suppose conversely $\alpha \in \mathfrak{O}_\Lambda$ and $\mathrm{N}(\alpha) = \pm 1$. This means $\alpha\alpha' = \pm 1$, so $\alpha^{-1} = \pm\alpha'$. But $\mathfrak{O}_\Lambda = \langle 1, c\omega \rangle$ for some c, so $\mathfrak{O}_\Lambda$ is closed under $\xi \mapsto \xi'$. Therefore $\alpha^{-1} \in \mathfrak{O}_\Lambda$, so α is a unit of $\mathfrak{O}_\Lambda$. □

Since $\mathfrak{O}_\Lambda = \langle 1, c\omega \rangle$, the units of $\mathfrak{O}_\Lambda$ are those elements $x + c\omega y$ such that $\mathrm{N}(x + c\omega y) = \pm 1$, that is,

$$\pm 1 = \begin{cases} x^2 - dc^2y^2, & \text{if } d \equiv 2 \text{ or } 3 \pmod 4; \\ (x + cy/2)^2 - dc^2y^2/4, & \text{if } d \equiv 1. \end{cases} \tag{20}$$

The easier case to consider is $d < 0$, when $\mathrm{N}(\xi) = |\xi|^2$. Then all units of $\mathfrak{O}_\Lambda$ lie on the unit circle: see Figure 9. If $d \equiv 2$ or 3, then (20) has the solutions

(i) $(\pm 1, 0)$, if $c > 1$ or $d < -1$;
(ii) $(\pm 1, 0)$ and $(0, \pm 1)$, if $c = 1$ and $d = -1$.

If $d \equiv 1$, then either $d = -3$, or else $d \leqslant -7$. In the latter case, the only solutions to (20) are $(\pm 1, 0)$. But if $d = -3$, so that (20) becomes

$$\Big(x + \frac{c}{2}y\Big)^2 + \frac{3}{4}c^2y^2 = \pm 1,$$

then the solutions are

(i) $(\pm 1, 0)$, if $c > 1$;

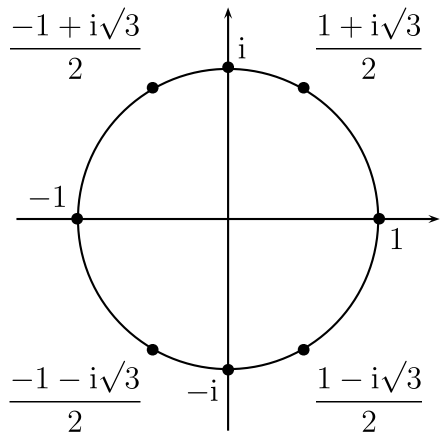

Figure 9. Units in imaginary quadratic fields

(ii) $(\pm1,0),(\pm1,\mp1),(0,\pm1)$, if $c=1$.

Thus we have shown:

Theorem 11. *When $d<0$, then the units of $\langle 1, c\omega\rangle$ are:*

(i) ±1, $\pm\omega$, *when* $c=1$ *and* $d=-1$*;*
(ii) ±1, $\pm\omega'$, $\pm\omega$, *when* $c=1$ *and* $d=-3$*;*
(iii) ±1, *in all other cases.*

Problem 6. *Solve the quadratic Diophantine equation*

$$x^2+xy+y^2=3. \tag{21}$$

Solution. Evidently $(1,1)$ is a solution. What are the others? We have

$$\begin{aligned} x^2+xy+y^2 &= x^2+xy+\frac{1}{4}y^2+\frac{3}{4}y^2 \\ &= \Big(x+\frac{1}{2}y\Big)^2+\Big(\frac{\sqrt{3}}{2}y\Big)^2 \\ &= \Big(x+\frac{1}{2}y+\frac{i\sqrt{3}}{2}\Big)\Big(x+\frac{1}{2}y-\frac{i\sqrt{3}}{2}\Big) \\ &= (x+\omega y)(x+\omega' y) \\ &= \mathrm{N}(x+\omega y), \end{aligned}$$

where we work in $\mathbb{Q}(\sqrt{-3})$. Let $\Lambda=\langle 1,\omega\rangle$, so that $\mathfrak{O}_\Lambda=\Lambda=\mathfrak{O}_K$, which has the six units ±1, $\pm\omega$, and $\pm\omega'$, all of norm 1. Since $1+\omega$ is a solution of

$$\mathrm{N}(\xi)=3$$

from Λ, so are $\pm(1+\omega)$, $\pm\omega(1+\omega)$, and $\pm\omega'(1+\omega)$. Since $\omega^2-\omega+1=0$, and $\omega+\omega'=1$, these solutions are $\pm(1+\omega)$, $\pm(2\omega-1)$, and $\pm(2-\omega)$, as in Figure 10. The corresponding 6 solutions of (21) are

$$(\pm1,\pm1),\quad(\mp1,\pm2),\quad(\pm2,\mp1),$$

as in Figure 11. It is easy to see from Figure 10 that there are no other solutions. Also, we can rewrite (21) as

$$\frac{(x+y/2)^2}{3}+\frac{y^2}{4}=1,$$

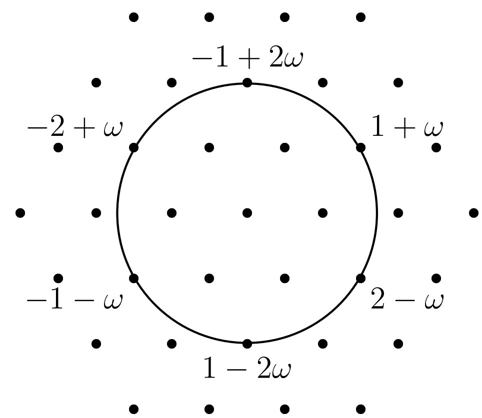

FIGURE 10. Solutions of $\mathrm{N}(x+\omega y)=3$ in $\mathbb{Q}(\sqrt{-3})$

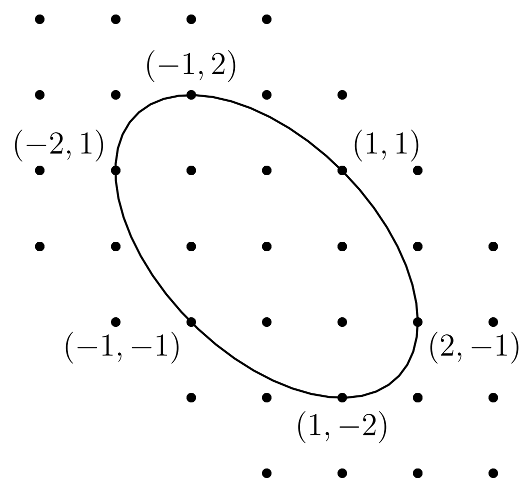

FIGURE 11. Solutions of $x^2+xy+y^2=3$

which defines the ellipse in Figure 11; then we just look for the integer points on the ellipse—there are only finitely many. However, it is not see easy to tell at a glance which integer points *are* on the ellipse. □

Problem 7. *Solve*

$$4x^2+2xy+y^2=7. \tag{22}$$

Solution. Again, one solution is $(1,1)$. We can try to factorize:

$$\begin{aligned} 4x^2+2xy+y^2 &= 3x^2+(x+y)^2 \\ &= (\sqrt{3}x+\mathrm{i}(x+y))(\sqrt{3}x-\mathrm{i}(x+y)) \\ &= ((\sqrt{3}+\mathrm{i})x+\mathrm{i}y)((\sqrt{3}-\mathrm{i})x-\mathrm{i}y), \end{aligned} \tag{23}$$

but this is not over a quadratic field. Indeed, a field that contains $\sqrt{3}+\mathrm{i}$ and i contains also $\sqrt{3}$. But $[\mathbb{Q}(\sqrt{3},\mathrm{i}):\mathbb{Q}]=4$ (see Figure 12). We can fix this problem by multiplying each factor in (23) by the appropriate unit, such as $-\mathrm{i}$ and i. What amounts to the same

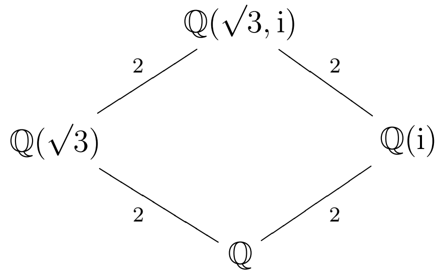

FIGURE 12. Subfields of $\mathbb{Q}(\sqrt{3}, \mathrm{i})$

thing is to compute as follows. We have

$$\begin{aligned} 3x^2 + (x+y)^2 &= (x+y)^2 + 3x^2 \\ &= (x + y + \mathrm{i}\sqrt{3}x)(x + y - \mathrm{i}\sqrt{3}x) \\ &= (2\omega x + y)(2\omega' x + y) \\ &= \mathrm{N}(2\omega x + y), \end{aligned}$$

again in $\mathbb{Q}(\sqrt{-3})$. Let $\Lambda = \langle 2\omega, 1\rangle = \langle 1, 2\omega\rangle$. We want to find the solutions of

$$\mathrm{N}(\xi) = 7 \tag{24}$$

in Λ. We know one solution, namely $1 + 2\omega$. Since $(2\omega)^2 - 2(2\omega) + 4 = 0$, we have $\mathfrak{O}_\Lambda = \langle 1, 2\omega\rangle = \Lambda$. The only units of $\mathfrak{O}_K$ in this are ± 1. Hence we have the solutions $\pm(1 + 2\omega)$ of (24). To find any others, again we can draw a picture, Figure 13. So (24)

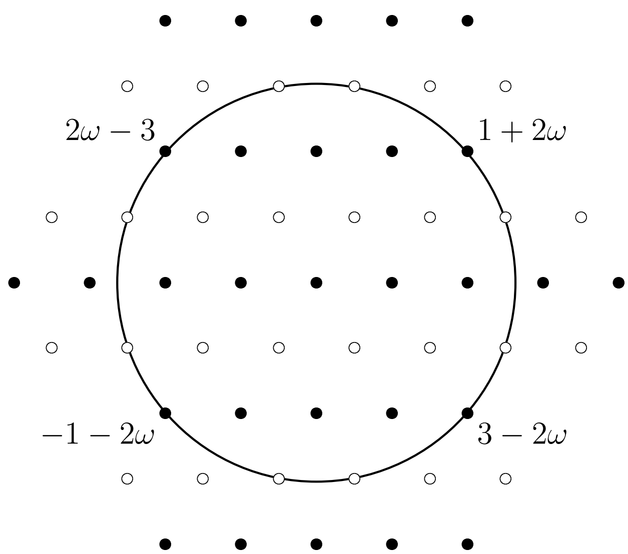

FIGURE 13. Solutions of $\mathrm{N}(\xi) = 7$ from $\langle 1, 2\omega\rangle$ in $\mathbb{Q}(\sqrt{-3})$

has the solutions $\pm(1 + 2\omega)$ and $\pm(3 - 2\omega)$, and no others. The solutions of (22) are therefore $(\pm 1, \pm 1)$ and $(\mp 1, \pm 3)$. These appear on the graph of (22) in Figure 14. □

In the same way, we can solve any quadratic Diophantine equation

$$ax^2 + bxy + cy^2 = m,$$

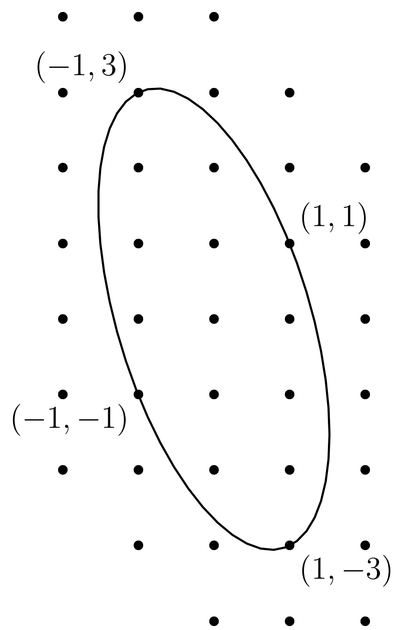

FIGURE 14. Solutions to $4x^2 + 2xy + 1 = 7$

provided $b^2 - 4ac < 0$. For in this case, the equation defines an ellipse, which is bounded, so that there are only finitely many possible solutions to check.

* * * * *

Now we move to the case where $d > 0$, so $K \subseteq \mathbb{R}$. We have

$$\langle 1, c\sqrt{d}\rangle \subseteq \langle 1, c\omega\rangle = \mathfrak{O}_\Lambda.$$

A unit of $\mathfrak{O}_\Lambda$ of the form $x + cy\sqrt{d}$ thus corresponds to a solution of

$$x^2 - dc^2y^2 = \pm 1.$$

The Pell equation $x^2 - dc^2y^2 = \pm 1$ has infinitely many solutions, and therefore $\mathfrak{O}_\Lambda$ has infinitely many units. We want to find them.

Suppose ε is a unit of $\mathfrak{O}_\Lambda$. Since there are infinitely many units, there are units other than ± 1. So we may assume $\varepsilon \neq \pm 1$. If $\varepsilon < 0$, then $-\varepsilon$ is a unit greater than 0. So we may assume $\varepsilon > 0$. If $0 < \varepsilon < 1$, then ε^{-1} is a unit greater than 1. So we may assume $\varepsilon > 1$. Also $\varepsilon < n$ for some n. But

$$\varepsilon^2 - (\varepsilon + \varepsilon')\varepsilon + \varepsilon\varepsilon' = 0,$$

that is, $\varepsilon^2 - \mathrm{Tr}(\varepsilon)\varepsilon + \mathrm{N}(\varepsilon) = 0$. Since $\pm 1 = \mathrm{N}(\varepsilon) = \varepsilon\varepsilon'$, we have $|\varepsilon'| = \varepsilon^{-1}$. Hence

$$|\mathrm{Tr}(\varepsilon)| = |\varepsilon + \varepsilon'| \leqslant \varepsilon + \varepsilon^{-1} < n + 1.$$

This shows that there are only finitely many possibilities for the equation $x^2 - \mathrm{Tr}(\varepsilon)x + \mathrm{N}(\varepsilon) = 0$. Hence there are only finitely many units of $\mathfrak{O}_\Lambda$ between 1 and n. Therefore there is a least such unit, the **fundamental unit,** which we may denote by

$$\varepsilon_\Lambda.$$

Then $(\varepsilon_\Lambda^n : n \in \mathbb{Z})$ is an increasing sequence, $\lim_{n\to\infty} \varepsilon_\Lambda^n = \infty$, and $\lim_{n\to-\infty} \varepsilon_\Lambda^n = 0$. Suppose ζ is a positive unit of $\mathfrak{O}_\Lambda$. Then

$$\varepsilon_\Lambda{}^n \leqslant \zeta < \varepsilon_\Lambda{}^{n+1}$$

for some n. Hence $1 \leqslant \varepsilon_\Lambda{}^{-n}\zeta < \varepsilon_\Lambda$. But $\varepsilon_\Lambda{}^{-n}\zeta$ is a unit too. By minimality of ε_Λ, we conclude that $\zeta = \varepsilon_\Lambda{}^n$. We have proved:

Theorem 12. *When $d > 0$, then the units of $\mathfrak{O}_\Lambda$ compose the multiplicative group generated by ε_Λ and -1. In particular, every unit is $\pm\varepsilon_\Lambda{}^n$ for some n in $\mathbb{Z}$. If $\mathrm{N}(\varepsilon_\Lambda) = 1$, then every unit has norm 1. If $\mathrm{N}(\varepsilon_\Lambda) = -1$, then the units of norm 1 are $\pm\varepsilon_\Lambda{}^{2n}$.*

How do we find ε_Λ?

Lemma 11. *Assuming $d > 0$, let ε be a unit $x + \omega y$ of $\mathfrak{O}_K$ such that $\varepsilon > 1$. Then either $x, y > 0$, or else $d = 5$ and $\varepsilon = \omega = (1 + \sqrt{5})/2$.*

Proof. We have

$$(\omega - \omega')y = \varepsilon - \varepsilon' \geqslant \varepsilon - |\varepsilon^{-1}| > 0,$$

and $\omega > \omega'$, so $y > 0$. Also

$$1 > |\varepsilon'| = |x + \omega' y|;$$

so since $\omega' < 0$, and hence $\omega' y < 0$, we must have $x \geqslant 0$, since $x \in \mathbb{Z}$. If $x > 0$, we are done. Suppose $x = 0$. Then

$$\pm 1 = \mathrm{N}(\omega y) = \begin{cases} -dy^2, & \text{if } d \equiv 2 \text{ or } 3 \pmod 4; \\ \dfrac{1-d}{4} y^2, & \text{if } d \equiv 1. \end{cases}$$

The only way this can happen is if $d = 5$ and $y = 1$ (since $y > 0$). □

12. April 4, 2008 (Friday)

When $d = 5$, then $\omega = \phi$, the so-called **Golden Ratio:**

$$\phi = \frac{1 + \sqrt{5}}{2}.$$

This has an intimate connexion with the sequence $(\mathrm{F}_n : n \in \omega)$ of **Fibonacci numbers,** given by

$$\mathrm{F}_0 = 0, \quad \mathrm{F}_1 = 1, \quad \mathrm{F}_{n+2} = \mathrm{F}_n + \mathrm{F}_{n+1}.$$

We can continue the sequence backwards, so that, if $n < 0$, then

$$\mathrm{F}_n = \mathrm{F}_{n+2} - \mathrm{F}_{n+1}.$$

Then the bi-directional sequence is

$$\dots,\quad 13,\quad -8,\quad 5,\quad -3,\quad 2,\quad -1,\quad 1,\quad 0,\quad 1,\quad 1,\quad 2,\quad 3,\quad 5,\quad 8,\quad 13,\quad \dots$$

Theorem 13. *The units of the ring of integers of $\mathbb{Q}(\sqrt{5})$ are $\pm\phi^n$; and*

$$\phi^n = \mathrm{F}_{n-1} + \mathrm{F}_n\,\phi. \tag{25}$$

Proof. Let $K = \mathbb{Q}(\sqrt{5})$. By Lemma 11, ϕ is the least unit of $\mathfrak{O}_K$ that is greater than 1. Then every unit is $\pm\phi^n$ for some n in $\mathbb{Z}$, by Theorem 12. Trivially (25) holds when $n = 1$. Also, ϕ is a root of

$$x^2 - x - 1 = 0,$$

so $\phi^2 = 1 + \phi$, which means

$$(x + y\phi)\phi = x\phi + y\phi^2 = y + (x + y)\phi. \tag{26}$$

Hence, if (25) holds when $n = k$, then

$$\phi^{k+1} = (\mathrm{F}_{k-1} + \mathrm{F}_k\,\phi)\phi = \mathrm{F}_k + (\mathrm{F}_{k-1} + \mathrm{F}_k)\phi = \mathrm{F}_k + \mathrm{F}_{k+1}\,\phi,$$

so it holds when $n = k + 1$. Therefore (25) holds for all positive n. But from (26) we have

$$x + y\phi = (y + (x + y)\phi)\phi^{-1}.$$

By letting $y = u$ and $x = v - u$, we get

$$v - u + u\phi = (u + v\phi)\phi^{-1}.$$

Thus, if (25) holds for some k, then

$$\phi^{k-1} = (\mathrm{F}_{k-1} + \mathrm{F}_k\,\phi)\phi^{-1} = \mathrm{F}_k - \mathrm{F}_{k-1} + \mathrm{F}_{k-1}\,\phi = \mathrm{F}_{k-2} + \mathrm{F}_{k-1}\,\phi,$$

so (25) holds when $n = k - 1$. Thus (25) holds for all n in $\mathbb{Z}$. □

13. April 8, 2008 (Tuesday)

Problem 8. *Solve the quadratic Diophantine equation*

$$4x^2 + 2xy - y^2 = 4. \tag{27}$$

Solution. We have

$$\begin{aligned} 4x^2 + 2xy - y^2 &= 4x^2 + 2xy + \frac{1}{4}y^2 - \frac{5}{4}y^2 \\ &= \Bigl(2x + \frac{1}{2}y\Bigr)^2 - \frac{5}{4}y^2 \\ &= (2x + \phi y)(2x + y\phi') \\ &= \mathrm{N}(2x + \phi y) \end{aligned}$$

in $\mathbb{Q}(\sqrt{5})$. Let $\Lambda = \langle 2, \phi\rangle$. Then $\mathfrak{O}_\Lambda = \mathrm{End}(\langle 2, \phi\rangle) = \mathrm{End}(\langle 1, \phi/2\rangle)$ by Lemma 9. Since

$$4\Bigl(\frac{\phi}{2}\Bigr)^2 - 2\cdot\frac{\phi}{2} - 1 = 0,$$

we have by Theorem 10 that $\mathfrak{O}_\Lambda = \langle 1, 2\phi\rangle$. Since $\mathrm{N}(\phi) = -1$, the positive elements of $\mathfrak{O}_\Lambda$ of norm 1 are the powers of the least power ϕ^{2n} (where $n > 0$) that belongs to $\langle 1, 2\phi\rangle$. By the previous theorem, we have

n	0	1	2	3	4	5	6
ϕ^n	1	ϕ	$1+\phi$	$1+2\phi$	$2+3\phi$	$3+5\phi$	$5+8\phi$

So every element of $\mathfrak{O}_\Lambda$ of norm 1 is $\pm(5 + 8\phi)^n$ for some n in $\mathbb{Z}$. This means, if γ is a solution of

$$\mathrm{N}(\xi) = 4$$

from Λ, then so is $\pm(5 + 8\phi)^n\gamma$. But we can choose n so that

$$1 \leqslant (5 + 8\phi)^n|\gamma| < 5 + 8\phi.$$

Let $(5 + 8\phi)^n|\gamma| = 2k + \ell\phi$. Then (k, ℓ) is a point on the graph of

$$1 \leqslant 2x + y\phi < 5 + 8\phi;$$

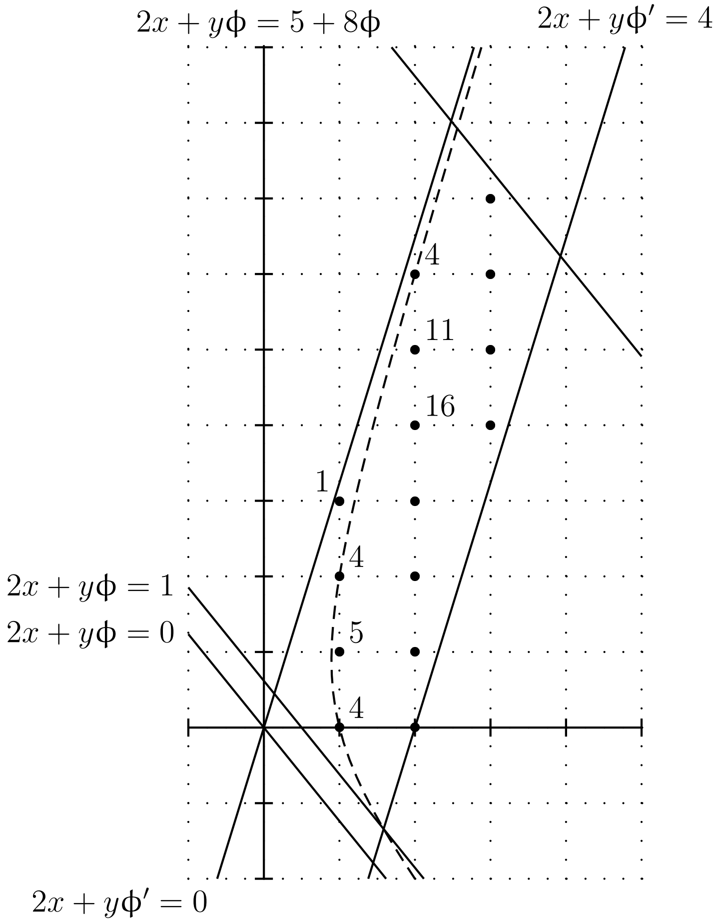

FIGURE 15. Solutions of $4x^2 + 2xy - y^2 = 4$

that is, (k, ℓ) lies between the straight lines given by

$$2x + y\phi = 1; \qquad 2x + y\phi = 5 + 8\phi. \tag{28}$$

(See Figure 15.) But also, (k, ℓ) lies on the hyperbola given by

$$(2x + y\phi)(2x + y\phi') = 4, \tag{29}$$

whose asymptotes are given by

$$(2x + y\phi)(2x + y\phi') = 0.$$

One of the asymptotes, given by $2x + y\phi = 0$, is parallel to the bounding lines given by (28). Directly from (29), the hyperbola itself meets the bounding line given by $2x + y\phi = 1$ at this line's intersection with the line given by $2x + y\phi' = 4$, parallel to the other asymptote. This means (k, ℓ) lies within the parallelogram in Figure 15. There are finitely many integer points in that parallelogram; for every such point (x, y), we compute $\mathrm{N}(2x+y\phi)$. In fact, once we have computed the norms indicated in the figure, we can see that the only points for which the corresponding norm is 4 are $(1, 0)$, $(1, 2)$, and $(2, 6)$. Therefore the solutions to (27) are those (x, y) such that $2x + y\phi = \pm(5 + 8\phi)^n\gamma$, where $n \in \mathbb{Z}$ and $\gamma \in \{2, 2 + 2\phi, 4 + 6\phi\}$. $\square$

14. April 11, 2008 (Friday)

Theorem 13 can be understood in terms of matrices. Multiplication in $\langle 1, \phi\rangle$ by ϕ corresponds to a matrix multiplication:

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ x+y \end{pmatrix}$$

Inverting the matrix, we have

$$\begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y-x \\ x \end{pmatrix}.$$

corresponding to multiplication by ϕ^{-1}.

We have

$$(x+y\phi)(5+8\phi) = 5x + (8x+5y)\phi + 8y\phi^2 = 5x+8y+(8x+13y)\phi,$$

and

$$\begin{pmatrix} 5 & 8 \\ 8 & 13 \end{pmatrix}^{-1} = \begin{pmatrix} 13 & 8 \\ 8 & 5 \end{pmatrix}.$$

We also have the correspondence $(x, y) \mapsto 2x+y\phi$ between solutions to (27) and elements of $\langle 2, \phi\rangle$ of norm 4. If (a, b) is a solution, we compute

$$\begin{pmatrix} 5 & 8 \\ 8 & 13 \end{pmatrix}\begin{pmatrix} 2a \\ b \end{pmatrix} = \begin{pmatrix} 10a+8b \\ 16a+13b \end{pmatrix}, \qquad \begin{pmatrix} 13 & 8 \\ 8 & 5 \end{pmatrix}\begin{pmatrix} 2a \\ b \end{pmatrix} = \begin{pmatrix} 26a+8b \\ 16a+5b \end{pmatrix},$$

so that $(5a+4b, 16a+13b)$ and $(13a+4b, 16a+5b)$ are also solutions. Hence the three bi-directional sequences of solutions (along the branch of the hyperbola depicted in Figure 15) can be written thus:

$$\begin{array}{l} \dots,(4181,-5168),(233,-288),(13,-16),(1,0),\ (5,16),\quad (89,288),\quad (1597,5168),\ \dots \\ \dots,(1597,-1974),\ (89,-110),\ \ (5,-6),\ \ (1,2),\ (13,42),\ (233,754),\ (4181,13530),\dots \\ \dots,\ \ (610,-754),\quad (34,-42),\quad (2,-2),\ \ (2,6),(34,110),(610,1974),(10946,35422),\dots \end{array}$$

We may note that each entry (except 0) appears more than once. And we can combine these solutions into one sequence, thus:

$$\dots,(34,-42),(13,-16),(5,-6),(2,-2),(1,0),(1,2),(2,6),(5,16),(13,42),(34,110),\dots$$

Dividing the second coordinates by 2 leaves

$$\dots,(34,-21),(13,-8),(5,-3),(2,-1),(1,0),(1,1),(2,3),(5,8),(13,21),(34,55),\dots$$

Here we see all of the Fibonacci numbers. We can obtain all solutions of (27) from $(1, 0)$ by the composition of operations

$$(x, 2y) \mapsto (x, y) \mapsto (x+y, x+2y) \mapsto (x+y, 2x+4y),$$

along with the inverse of this composition. The middle operation in this composition corresponds to multiplication by $1+\phi$:

$$(x+y\phi)(1+\phi) = x+(x+y)\phi + y\phi^2 = x+y+(x+2y)\phi.$$

Thus every solution of (27) is (x, y), where $2x+y\phi = \pm 2(1+\phi)^n$ for some n in $\mathbb{Z}$. Note however that $1+\phi \notin \langle 1, 2\phi\rangle$, that is, $1+\phi \notin \mathfrak{O}_\Lambda$ when $\Lambda = \langle 2, \phi\rangle$.

15. April 15, 2008 (Tuesday)

If we convert a quadratic Diophantine equation to the form $\mathrm{N}(x\alpha + y\beta) = m$, where $\alpha, \beta \in K$, then we can solve as in the examples above, provided we can find the units of $\mathfrak{O}_K$. The case where $d > 0$ is the challenging case. What is the fundamental unit ε (such that every unit of $\mathfrak{O}_K$ is $\pm\varepsilon^n$ for some n)?

We have $\mathfrak{O}_K = \langle 1, \omega\rangle$, and

$$\mathrm{N}(x + y\omega) = \begin{cases} x^2 - dy^2, & \text{if } d \equiv 4 \text{ or } 3 \pmod 4; \\ (x + y/2)^2 - dy^2/4, & \text{if } d \equiv 1. \end{cases}$$

We know $\varepsilon = a + b\omega$, where $a, b > 0$, unless $d = 5$. Assuming $d \neq 5$, we shall show that a/b is a convergent of $\sqrt{d}$, if $d \equiv 2$ or 3; otherwise, $(2a + b)/b$ is a convergent of $\sqrt{d}$.

Lemma 12. *Assuming $\sqrt{d} = [a_0; a_1, a_2, \dots]$, let p_n/q_n be the nth convergent, that is, $p_n/q_n = [a_0; a_1, \dots, a_n]$. Suppose $a, b \in \mathbb{Z}$ and $1 \leqslant b < p_{n+1}$. Then*

$$|p_n - q_n\sqrt{d}| \leqslant a - b\sqrt{d},$$

so that

$$q_n\left|\frac{p_n}{q_n} - \sqrt{d}\right| \leqslant b\left|\frac{a}{b} - \sqrt{d}\right|.$$

Proof. By Theorem 2, we have

$$(-1)^n = p_{n+1}q_n - p_nq_{n+1} = \begin{vmatrix} p_{n+1} & p_n \\ q_{n+1} & q_n \end{vmatrix}.$$

So there are s and t in $\mathbb{Z}$ such that

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} p_{n+1} & p_n \\ q_{n+1} & q_n \end{pmatrix}\begin{pmatrix} s \\ t \end{pmatrix} = \begin{pmatrix} sp_{n+1} + tp_n \\ sq_{n+1} + tq_n \end{pmatrix}$$

Then

$$a - b\sqrt{d} = sp_{n+1} + tp_n - sq_{n+1}\sqrt{d} - tq_n\sqrt{d} = s(p_{n+1} - q_{n+1}\sqrt{d}) + t(p_n - q_n\sqrt{d}).$$

So it is enough to show that $t \neq 0$ and the two terms here, $s(p_{n+1} - q_{n+1}\sqrt{d})$ and $t(p_n - q_n\sqrt{d})$ have the same sign. But the factors $p_{n+1} - q_{n+1}\sqrt{d}$ and $p_n - q_n\sqrt{d}$, have opposite sign. So it is enough to show $t \neq 0$ and $st \leqslant 0$.

To show $t \neq 0$, we note

$$\begin{pmatrix} s \\ t \end{pmatrix} = (-1)^n\begin{pmatrix} q_n & -p_n \\ -q_{n+1} & p_{n+1} \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix},$$

so

$$t = (-1)^n(-aq_{n+1} + bp_{n+1}).$$

If $t = 0$, then $aq_{n+1} = bp_{n+1}$; but $\gcd(p_{n+1}, q_{n+1}) = 1$, so $q_{n+1} \mid b$, hence $q_{n+1} \leqslant b$.

To show $st \leqslant 0$, suppose $s \neq 0$. We have

$$b = sq_{n+1} + tq_n.$$

If $s < 0$ and $1 \leqslant b$, then $t > 0$; if $s > 0$ and $b < q_{n+1}$, then $t < 0$. □

The lemma uses only that $\sqrt{d}$ *has* convergents up to p_{n+1}/q_{n+1}. The following theorem requires only that all convergents of $\sqrt{d}$ exist, that is, $\sqrt{d}$ must be irrational.

Theorem 14. *If a and b are positive rational integers, and*
$$\left|\frac{a}{b}-\sqrt{d}\right|<\frac{1}{2b^2},$$
then a/b is a convergent of $\sqrt{d}$.

Proof. Since $(q_n\colon n\in\omega)$ increases to ∞, we can find n such that
$$q_n\leqslant b<q_{n+1}.$$
By the lemma, we have
$$\begin{gathered}q_n\left|\frac{p_n}{q_n}-\sqrt{d}\right|\leqslant b\left|\frac{a}{b}-\sqrt{d}\right|<\frac{1}{2b},\\ \left|\frac{p_n}{q_n}-\sqrt{d}\right|<\frac{1}{2bq_n}.\end{gathered}$$
Then
$$\begin{aligned}\frac{1}{bq_n}|aq_n-bp_n|&=\left|\frac{a}{b}-\frac{p_n}{q_n}\right|\leqslant\left|\frac{a}{b}-\sqrt{d}\right|+\left|\sqrt{d}-\frac{p_n}{q_n}\right|<\frac{1}{2b^2}+\frac{1}{2bq_n}\leqslant\frac{1}{bq_n},\\ |aq_n-bp_n|&<1,\end{aligned}$$
so $aq_n=bp_n$ and $a/b=p_n/q_n$. □

Theorem 15. *Assuming $d>0$, let $a+b\omega$ be a unit of $\mathfrak{O}_K$, where $a,b>0$.*

(i) *If $d\equiv 2$ or $3 \pmod 4$, then a/b is a convergent of $\sqrt{d}$.*
(ii) *If $d\equiv 1$, then $(2a+b)/b$ is a convergent of $\sqrt{d}$, provided either $d\geqslant 17$, or else $d=13$ and $a+b\omega$ is the fundamental unit of $\mathfrak{O}_K$.*

Also, a is the nearest integer to $-b\omega'$.

Proof. Suppose first $d\equiv 2$ or 3, so that
$$a^2-db^2=\pm 1.$$
By the last theorem, it is enough to show
$$\left|\frac{a}{b}-\sqrt{d}\right|<\frac{1}{2b^2},$$
that is,
$$|a-b\sqrt{d}|<\frac{1}{2b},$$
that is (multiplying by $a+b\sqrt{d}$),
$$1<\frac{a+b\sqrt{d}}{2b}=\frac{1}{2}\left(\frac{a}{b}+\sqrt{d}\right).$$
But we have
$$\begin{gathered}a^2-db^2\geqslant -1,\\ \left(\frac{a}{b}\right)^2\geqslant d-\frac{1}{b^2}\geqslant d-1,\\ \frac{a}{b}\geqslant\sqrt{d-1},\\ \frac{1}{2}\left(\frac{a}{b}+\sqrt{d}\right)\geqslant\frac{1}{2}(\sqrt{d-1}+\sqrt{d})>1\end{gathered}$$
since $d\geqslant 2$.

In case $d \equiv 1$, we try to proceed as before. We have

$$(2a+b)^2 - db^2 = \pm 4,$$

so that

$$\left(\frac{2a+b}{b}\right)^2 \geqslant d - \frac{4}{b^2} \geqslant d - 4. \tag{30}$$

We should like to show

$$4 < \frac{1}{2}\left(\frac{2a+b}{b} + \sqrt{d}\right).$$

It is enough if we can show

$$4 < \frac{1}{2}(\sqrt{d-4} + \sqrt{d}).$$

We have this if $d \geqslant 21$. It remains to consider the cases when d is 13 or 17. We can do this with the second part of the theorem.

Indeed, since $a, b > 1$, we have $a + b\omega > 2$. Since

$$1 = (a + b\omega)|a + b\omega'|,$$

we conclude

$$|a + b\omega'| < \frac{1}{2},$$

so a is the nearest integer to $-b\omega'$.

In case $d = 13$, we have $-\omega' \approx 1.3$, to which 1 is the nearest integer; and $1+\omega$ is indeed a unit (of norm -1) and is the least possible unit greater than 1, so it is the fundamental unit of $\mathfrak{O}_K$. But $(2 \cdot 1 + 1)/1 = 3$, which is the first convergent of $\sqrt{13}$.

When $d = 17$, we have $-\omega' \approx 1.56$, to which 2 is nearest; but $\mathrm{N}(2+\omega) = 2$. So $b > 1$. Then instead of (30) we have

$$\left(\frac{2a+b}{b}\right)^2 \geqslant d - \frac{4}{b^2} \geqslant d - 1.$$

So it is enough if we have

$$4 < \frac{1}{2}(\sqrt{d-1} + \sqrt{d});$$

but we do have this. $\square$

Index

References

[1] William W. Adams and Larry Joel Goldstein. *Introduction to number theory*. Prentice-Hall Inc., Englewood Cliffs, N.J., 1976.

[2] David M. Burton. *Elementary Number Theory*. McGraw-Hill, Boston, sixth edition, 2007.

[3] Graham Everest and Thomas Ward. *An introduction to number theory*, volume 232 of *Graduate Texts in Mathematics*. Springer-Verlag London Ltd., London, 2005.

[4] G. H. Hardy and E. M. Wright. *An introduction to the theory of numbers*. The Clarendon Press Oxford University Press, New York, fifth edition, 1979.

[5] D. Hilbert and S. Cohn-Vossen. *Geometry and the imagination*. Chelsea Publishing Company, New York, N. Y., 1952. Translated by P. Neményi.

[6] Edmund Landau. *Elementary number theory*. Chelsea Publishing Co., New York, N.Y., 1958. Translated by J. E. Goodman.

[7] Serge Lang. *Elliptic functions*, volume 112 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1987. With an appendix by J. Tate.

Mathematics Dept, Middle East Technical University, Ankara 06531, Turkey
E-mail address: `dpierce@metu.edu.tr`
URL: `http://www.math.metu.edu.tr/~dpierce/`