A FIRST CLASS IN ALGEBRAIC NUMBER THEORY

PÉTER MAGA

1. The fundamental theorem of arithmetic

The fundamental theorem of airthmetic states that every nonzero integer can be written – essentially uniquely – as a product of prime (irreducible) numbers. My formulation here is very misleading, since the heart of the theorem is the fact that primes and irreducibles are the same. Their coincidence is, however, far from being trivial. This will be our first goal.

Proposition 1 (Euclidean division)**.** *Given integers a, b, $b \neq 0$. Then there exist integers c, d satisfying $a = bc + d$ and $|d| < |b|$.*

Proof. Let $a \geq 0$, $b > 0$, the remaining cases are similar. Induct on a. For $a = 1$, the statement is trivial ($c = 1, d = 0$ if $b = 1$ and $c = 0, d = 1$ if $b > 1$). Now assume that the statement holds for any $0 \leq a' < a$. If $a < b$, then $c = 0, d = a$. If $a \geq b$, then $a - b = bc' + d'$ with $|d'| < |b|$ by induction, so $a = b(c' + 1) + d'$. □

Remark 2. A little more careful analysis shows that even $|d| \leq |b|/2$ can be required.

Proposition 3. *Assume $a, b \in \mathbf{Z}$. Then there exists an integer $\gcd(a,b)$ satisfying $\gcd(a,b) | a, b$ and also that whenever $d|a, b$, $d | \gcd(a,b)$.*

Proof. If $b = 0$, then $\gcd(a,b) = a$ does the job. Otherwise, consider the sequence $(a, b, d_1, \ldots, d_n, 0)$, where each d_i is defined via the Euclidean division $d_{i-2} = c_{i-1} d_{i-1} + d_i$ (with $d_{-1} = a, d_0 = b, d_{n+1} = 0$). It is clear that such a sequence of Euclidean divisions terminates, since the absolute value decreases in each step. Set $\gcd(a,b) = d_n$. It is clear that $d_n | d_{n+1}, d_n$, and then by induction, $d_n | d_i, d_{i-1}$ implies $d_n | d_{i-2}$. Also, if $d | d_{i-2}, d_{i-1}$ (which holds for $i = 1$), then $d | d_i$, yielding $d | d_n = \gcd(a,b)$. □

Proposition 4. *Assume $a, b \in \mathbf{Z}$. Then $\gcd(a,b) = au + bv$ for some $u, v \in \mathbf{Z}$.*

Proof. If $b = 0$, the statement is trivial. Otherwise, we can create the same sequence $(a, b, d_1, \ldots, d_n, 0)$ as in the proof above. Clearly $d_{-1} = a, d_0 = b$ are integral combinations of a and b. Also, if d_{i-2}, d_{i-1} are integral combinations, then so is d_i. □

Definition 5 (prime numbers)**.** A nonzero integer p is said to be prime, if $p \nmid 1$, and whenever $p | ab$, $p | a$ or $p | b$.

Definition 6 (irreducible number)**.** A nonzero integer p is said to be irreducible, if $p \nmid 1$, and whenever $p = ab$, $a | 1$ or $b | 1$.

Proposition 7. *An integer p is prime if and only if it is irreducible.*

Proof. Assume p is prime, and let $p = ab$. Then $a, b \neq 0$. If $a \nmid 1$ and $b \nmid 1$, then $1 < |a|, |b| < p$. Therefore $p \nmid a, b$, which is a contradiction.

Assume p is irreducible, and $p | ab$. If $p | a$, we are done. If $p \nmid a$, then $\gcd(a, p) = 1$, since p is irreducible. Then there exist integers u, v satisfying $au + pv = 1$. Multiplying by b, we obtain $abu + pbv = b$, the left-hand side is divisible by p, so is the right-hand side. □

Theorem 8 (fundamental theorem of arithmetic)**.** *Every nonzero integer can be written as a product of prime (irreducible) numbers. The decomposition is unique, apart from factors dividing* 1.

Proof. First we prove the existence by induction on $|n|$. For $|n| = 1$, it is trivial. Assume that the statement holds for any n' with $|n'| < |n|$. If n is irreducible, we are done. If not, we can write it as a product $n = ab$ with $|a|, |b| < |n|$. We are done by induction.

Now we prove the uniqueness. Assume n has two decompositions $p_1 \cdot \ldots \cdot p_k = q_1 \cdot \ldots \cdot q_l$. Here, p_1 divides the left-hand side, so it divides the right-hand side as well. Then, since it is a prime, it divides a factor of the right-hand

side, say, q_1. Then $p_1|q_1$, and also $q_1|p_1$, since q_1 is irreducible. Dividing by them, we can complete the proof by induction. □

2. Gaussian integers

The set $\mathbf{Q}(\sqrt{-1})$ of Gaussian numbers is the set of numbers of the form $a+b\sqrt{-1}$, where $a,b\in\mathbf{Q}$ and $\sqrt{-1}$ is a formal expression, a number, whose square is -1. Addition and subtraction are performed formally:

$$(a+b\sqrt{-1})\pm(a'+b'\sqrt{-1})=(a\pm a')+(b\pm b')\sqrt{-1}.$$

In the case of multiplication, we proceed again formally, and use $\sqrt{-1}^2=-1$, obtaining

$$(a+b\sqrt{-1})\cdot(a'+b'\sqrt{-1})=(aa'-bb')+(ab'+a'b)\sqrt{-1}.$$

Also, at least formally, we have

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

It is easy to check that the right-hand side is the reciprocal of $a+b\sqrt{-1}$, provided that at least one of a and b is nonzero. Of course, with reciprocals in hand, we can define division by any nonzero number. Altogether, in $\mathbf{Q}(\sqrt{-1})$, we can add, subtract, multiply and divide (by a nonzero number), just like in $\mathbf{Q}$. Moreover, $\mathbf{Q}\subset\mathbf{Q}(\sqrt{-1})$ via $a\mapsto a+0\sqrt{-1}$. The field $\mathbf{Q}$ embeds in $\mathbf{R}$, and this embedding gives rise to an embedding $\mathbf{Q}(\sqrt{-1})\hookrightarrow\mathbf{C}$ via $\sqrt{-1}\mapsto i$. In what follows, we shall often think of $\mathbf{Q}(\sqrt{-1})$ as *it is sitting* in $\mathbf{C}$.

From the number-theoretic point of view, the subset $\mathscr{O}(\sqrt{-1})=\{a+b\sqrt{-1}:a,b\in\mathbf{Z}\}$, called the ring of Gaussian integers, has a fundamental importance. In this ring, we can add, subtract and multiply. However, division (just like in $\mathbf{Z}$) is a subtlety: it may happen that $\alpha,0\neq\beta$ are Gaussian integers, but α/β is not. This leads to the notion of divisibility.

Definition 9. We say that $\alpha\in\mathscr{O}(\sqrt{-1})$ is divisible by $\beta\in\mathscr{O}(\sqrt{-1})$, if there exists $\gamma\in\mathscr{O}(\sqrt{-1})$ such that $\alpha=\beta\gamma$.

Definition 10 (units, associates)**.** We say that α is a unit, if $\alpha|1$. We say that $\alpha\neq0$ and $\beta\neq0$ are associates ($\alpha\sim\beta$), if $\alpha|\beta$ and $\beta|\alpha$.

Exercise 1. Prove that units are the associates of 1. Prove that $\alpha\sim\beta$ if and only if α/β is a unit.

Exercise 2. List the units of $\mathscr{O}(\sqrt{-1})$. (What are the units in $\mathbf{Z}$?)

Definition 11 (Gaussian norm)**.** Given a Gaussian number $\alpha=a+b\sqrt{-1}$. Then its norm is defined as $N(\alpha)=a^2+b^2$. In other words, $N(\alpha)=\alpha\overline{\alpha}$, where $\overline{\alpha}$ is defined as $a-b\sqrt{-1}$.

Exercise 3. Prove that $N(\alpha\beta)=N(\alpha)N(\beta)$.

A remarkable fact is that in $\mathscr{O}(\sqrt{-1})$, we have Euclidean division, with norm taking the place of absolute value. Then there is gcd, there are integral combinations, and then it follows that prime numbers and irreducibles (defined the same way as in $\mathbf{Z}$) are the same. Then the fundamental theorem of arithmetic holds.

Exercise 4. Work out the details.

To demonstrate the strength of the unique factorization in $\mathscr{O}(\sqrt{-1})$, we prove that rational prime numbers of the form $4k+1$ can be represented as the sum of two squares. Gaussian integers and their unique factorization handles the problem in a splendid manner.

Theorem 12. *Assume p is a prime number of the form $4k+1$. Then p can be represented as the sum of two squares.*

Proof. First we prove that there are integers a',b' such that they are not divisible by p and $a'^2+b'^2$ is divisible by p. Indeed, let $a'=1$, and $b'=((p-1)/2)!$. Since $(p-1)/2$ is even, $b'^2\equiv(p-1)!\equiv-1\pmod p$ by Wilson's theorem.

Now reducing modulo p and multiplying by -1 if necessary, we may assume that $a,b\in\mathbf{Z}$, $0<a,b<p/2$ such that $p|a^2+b^2$. Consider the Gaussian integer $\alpha=a+b\sqrt{-1}$. Then $N(\alpha)=\alpha\overline{\alpha}=a^2+b^2<2p^2/4<p^2$. Since $\alpha,\overline{\alpha}\in\mathscr{O}(\sqrt{-1})$, this shows that $p|\alpha\overline{\alpha}$. We claim that $p\nmid\alpha,\overline{\alpha}$. Indeed, $p|\alpha$ implies $p^2=N(p)|N(\alpha)<p^2/4$, which is a contradiction, since $N(\alpha)\neq0$ (with $\overline{\alpha}$, the proof is the same). Therefore, p is not a prime number in $\mathscr{O}(\sqrt{-1})$. It follows from the unique factorization that it is not irreducible. Take the proper factorization $p=\beta\gamma$. Here $N(\beta),N(\gamma)>1$, $N(\beta)N(\gamma)=p^2$, therefore $N(\beta)=p$. If $\beta=c+d\sqrt{-1}$, this means $c^2+d^2=p$. □

Problem 5. Classify the primitive Pythagorean triples (i.e.: solve $a^2+b^2=c^2$ under $\gcd(a,b,c)=1$).

Problem 6. Solve the Diophantine equation $x^2+1=y^3$.

Problem 7 (IMO 2008, P3)**.** Prove that there exist infinitely many positive integers n such that n^2+1 has a prime divisor which is greater than $2n+\sqrt{2n}$.

3. Field extensions and the ring of integers

Classical algebraic number theory revolves around the study of finite algebraic extensions of the rational field $\mathbf{Q}$. For simplicity, we will consider only extensions of degree two. Our aim is to put the Gaussian integers into a more general framework.

Take a degree two polynomial of rational coefficeints with no rational roots. By a linear change of variables, we may assume that it is of the form x^2-d, where d is not a square in $\mathbf{Q}$. This is the generality we want to work with: when $d=-1$, this gives back the case of Gaussians.

Definition 13 ($\mathbf{Q}(\sqrt{d})$)**.** The field $\mathbf{Q}(\sqrt{d})$ consists of formal expressions $a+b\sqrt{d}$ with operations defined as

$$(a+b\sqrt{d})\pm(a'+b'\sqrt{d})=(a\pm a')+(b\pm b')\sqrt{d},\qquad (a+b\sqrt{d})(a'+b'\sqrt{d})=(aa'+bb'd)+(ab'+ba')\sqrt{d}.$$

We regard $\mathbf{Q}$ as embedded in $\mathbf{Q}(\sqrt{d})$ via $a\mapsto a+0\sqrt{d}$.

Proposition 14. *$\mathbf{Q}(\sqrt{d})$ is a field.*

Proof. As all other statements are straight-forward calculations, we prove only that a nonzero element is invertible, that is, if $a\neq 0$ or $b\neq 0$, there is $a'+b'\sqrt{d}$ satisfying $(a+b\sqrt{d})(a'+b'\sqrt{d})=1$. Then, at least formally,

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

and the denominators are nonzero, since d is not a square in $\mathbf{Q}$. It is easy to check that the right-hand side is indeed a multiplicative inverse of $a+b\sqrt{d}$. □

In a field, it makes no sense to speak about divisibility, since everything divides everything (apart from 0, at least). However, there is a natural ring in this field, in which we can study number theory, namely, the ring of integers. The first guess could be $\mathbf{Z}+\mathbf{Z}\sqrt{d}$, which is indeed a subring. It turns out that this is not the right candidate. From this point on, we assume that d is a square-free integer: this does not lose generality, since $\mathbf{Q}(\sqrt{d})=\mathbf{Q}(\sqrt{q^2d})$ for any nonzero rational number q.

Definition 15. A number $\alpha\in\mathbf{Q}(\sqrt{d})$ is said to be an integer, if there exists a degree 2 polynomial p of integral coefficients and leading coefficient 1 such that $p(\alpha)=0$. The set of integers is denoted by $\mathcal{O}(\sqrt{d})$.

It is not obvious at all that the sum or product of two integers is an integer as well.

Theorem 16 (description of integers)**.** *If $d\equiv 2,3 \pmod 4$, then $\mathcal{O}(\sqrt{d})=\mathbf{Z}+\mathbf{Z}\sqrt{d}$. If $d\equiv 1 \pmod 4$, then $\mathcal{O}(\sqrt{d})=\mathbf{Z}+\mathbf{Z}(1+\sqrt{d})/2$.*

Proof. Assume $\alpha=a+b\sqrt{d}$ is an integer. Then by definition, for some *rational* integers A,B,

$$0=(a+b\sqrt{d})^2+A(a+b\sqrt{d})+B=a^2+b^2d+Aa+B+(2ab+Ab)\sqrt{d}.$$

If $b=0$, this means that $a^2+Aa+B=0$. Then a must be an integer: suppose not and p is a prime divisor of the denominator of a, this implies that the denominator of a^2+Aa+B is divisible by p^2 even in the simplest form, a contradiction. If $b\neq 0$, then $A=-2a$, from which we see that $a\in(\mathbf{Z}/2)$. If a is a rational integer, then b is also a rational integer, since $b^2=(a^2-B)/d$ and d is square-free. If a is not a rational integer, then $4B=(2a)^2-(2b)^2d$. Here, the left-hand side is divisible by 4, the first term on the right-hand side is 1 modulo 4. This leads to contradiction, if $d\equiv 2,3 \pmod 4$, and to $b\in(\mathbf{Z}+1/2)$, if $d\equiv 1 \pmod 4$.

Conversely, observe that $(a+b\sqrt{d})^2=2a(a+b\sqrt{d})-a^2+b^2d$ implies that $p(a+b\sqrt{d})=0$ with $p(x)=x^2-2ax+a^2-b^2d$. It is easy to check that the coefficients are integral in each case. □

4. Arithmetic in $\mathcal{O}(\sqrt{d})$

We can easily generalize some earlier notions.

Definition 17 (units, associates)**.** We say that α is a unit, if $\alpha|1$. We say that $\alpha \neq 0$ and $\beta \neq 0$ are associates ($\alpha \sim \beta$), if $\alpha|\beta$ and $\beta|\alpha$.

Exercise 8. Prove that units are the associates of 1. Prove that $\alpha \sim \beta$ if and only if α/β is a unit.

Definition 18 (norm)**.** Given a number $\mathbf{Q}(\sqrt{d}) \ni \alpha = a + b\sqrt{d}$. Then its norm is defined as $N(\alpha) = a^2 - b^2 d$. In other words, $N(\alpha) = \alpha\overline{\alpha}$, where $\overline{\alpha}$ is defined as $a - b\sqrt{d}$. It is easy to see that $N(\alpha) = 0$ if and only if $\alpha = 0$.

Exercise 9. Prove that $N(\alpha\beta) = N(\alpha)N(\beta)$.

The nature of the fields $\mathbf{Q}(\sqrt{d})$ with $d > 0$ (referred as *real quadratic fields*) differ very much from those with $d < 0$ (referred as *imaginary quadratic fields*). We study them separately.

4.1. **Imaginary quadratic fields ($d < 0$).** We start with the case $d < 0$. In this case, $\mathbf{Q}(\sqrt{d})$ can be regarded as embedded in $\mathbf{C}$ via $1 \mapsto 1$ and $\sqrt{d} \mapsto \sqrt{|d|}i$.

Exercise 10. Prove that $N(\alpha) = |\alpha|^2$, where $|\cdot|$ stands for the ordinary absolute value. In particular, N is nonnegative.

Exercise 11. List the units of $\mathcal{O}(\sqrt{d})$.

By a lattice in $\mathbf{C}$, we mean a set of the form $\mathbf{Z}\omega_1 + \mathbf{Z}\omega_2$, where $\omega_{1,2}$ are nonzero complex numbers such that ω_1/ω_2 is nonreal.

Exercise 12. Prove that $\mathcal{O}(\sqrt{d})$ is a lattice in $\mathbf{C}$.

What about unique factorization in $\mathcal{O}(\sqrt{d})$? Our wishful thinking is crushed by simple examples.

Exercise 13. Prove that in $\mathcal{O}(\sqrt{-5})$, 6 decomposes in two essentially different ways.

Conjecture 19 (Gauss, 1801)**.** *Unique factorization holds only for finitely many negative d.*

This was proved by Heilbronn in 1934. Effectively, we have the following.

Fact 20 (Heegner, 1952)**.** *Unique factorization holds if and only if*

$$d \in \{-1, -2, -3, -7, -11, -19, -43, -67, -163\}.$$

4.2. **Real quadratic fields ($d > 0$).** In many aspects, real quadratic fields are much more complicated. First of all, the norm function N has negative values. Secondly, it is more difficult to discover $\mathcal{O}(\sqrt{d})$ as a lattice. However, this can be done as follows. Since $d > 0$, there is a real number $\sqrt{d}_{\mathbf{R}}$ satisfying $\sqrt{d}_{\mathbf{R}}^2 = d$ and $\sqrt{d}_{\mathbf{R}} > 0$. There are two ways to draw an image of $\mathbf{Q}(\sqrt{d})$ (note: this is an "abstract" field), namely

$$\sigma_1 : \sqrt{d} \mapsto \sqrt{d}_{\mathbf{R}}, 1 \mapsto 1_{\mathbf{R}}, \qquad \sigma_2 : \sqrt{d} \mapsto -\sqrt{d}_{\mathbf{R}}, 1 \mapsto 1_{\mathbf{R}}.$$

Since $\sqrt{d}_{\mathbf{R}}$ and $-\sqrt{d}_{\mathbf{R}}$ satisfy the same equations, algebraically σ_1 cannot be distinguished from σ_2. This is not the case analytically: there is a sequence of rational numbers $p_n/q_n = \sigma_1(p_n/q_n) = \sigma_2(p_n/q_n)$ tending to $\sqrt{d}_{\mathbf{R}} = \sigma_1(\sqrt{d}) = -\sigma_2(\sqrt{d}) \neq \sigma_2(\sqrt{d})$. The idea is to consider the embeddings $\sigma_{1,2} : \mathbf{Q}(\sqrt{d}) \hookrightarrow \mathbf{R}$ simultaneously. That is, for any $\alpha \in \mathbf{Q}(\sqrt{d})$, set

$$\sigma(\alpha) = (\sigma_1(\alpha), \sigma_2(\alpha)) \in \mathbf{R}^2.$$

As before, a lattice in $\mathbf{R}^2$ is a set of the form $\mathbf{Z}\omega_1 + \mathbf{Z}\omega_2$, where $\omega_{1,2}$ are nonzero vectors in $\mathbf{R}^2$ which are not collinear.

Exercise 14. Prove that $\sigma(\mathcal{O}(\sqrt{d}))$ is a lattice in $\mathbf{R}^2$.

Comparing to the imaginary case, we see that the norm function N does not reflect any notion of closeness in $\mathbf{R}^2$. This causes difficulties in listing the units: while in the imaginary case, we could simply consider the complex numbers close to 0 (in the complex absolute value), in the real case, a priori at least, there might be units very far from 0 (in fact, this will turn out to be the truth).

Let us denote by U the set of units. They form a group under multiplication, by which we mean that $\alpha, \beta \in U$ implies $\alpha\beta \in U$, $1/\alpha \in U$.

Exercise 15. Prove that $\alpha \in \mathcal{O}(\sqrt{d})$ is a unit if and only if $N(\alpha) = \pm 1$.

Consider the mappings, for any $0 \neq \alpha \in \mathbf{Q}(\sqrt{d})$

$$\alpha \mapsto \sigma(\alpha) = (\sigma_1(\alpha), \sigma_2(\alpha)) \mapsto |\sigma(\alpha)| = (|\sigma_1(\alpha)|, |\sigma_2(\alpha)|) \mapsto \log|\sigma(\alpha)| = (\log|\sigma_1(\alpha)|, \log|\sigma_2(\alpha)|).$$

Introduce the notation $L(\alpha) = (L_1(\alpha), L_2(\alpha)) = (\log|\sigma_1(\alpha)|, \log|\sigma_2(\alpha)|)$. Then $L : \mathbf{Q}(\sqrt{d}) \setminus \{0\} \to \mathbf{R}^2$ satisfies $L(\alpha\beta) = L(\alpha) + L(\beta)$, in other words, this is a logarithmic mapping.

Exercise 16. Prove that $L(\alpha) = 0$ if and only if $\alpha = \pm 1$.

Exercise 17. Prove that if α is a unit in $\mathcal{O}(\sqrt{d})$, then $L(\alpha)$ satisfies the equation $L_1(\alpha) + L_2(\alpha) = 0$. In other words, units are mapped to the line $x + y = 0$ in $\mathbf{R}^2 = \{(x, y) : x, y \in \mathbf{R}\}$.

Assume temporarily that $L(U)$ has a nontrivial element, in other words, there is a unit differing from ± 1.

Exercise 18. Prove that $L(U)$ has an element a closest to 0 and then $L(U) = \{\mathbf{Z}a\}$.

Therefore, there exists $\alpha \in U$ satisfying $L(\alpha) = a$, a closest element to 0 on the line defined by $x + y = 0$. We claim that each element of U is of the form $\pm\alpha^m$ for some $m \in \mathbf{Z}$. Indeed, let $\beta \in U$. Then $L(\beta) = b = ma = mL(\alpha)$ for some integer m. This means that $|\sigma_1(\beta)| = |\sigma_1(\alpha)|^m$ and $|\sigma_2(\beta)| = |\sigma_2(\alpha)|^m$. Thus $\sigma_1(\beta/\alpha^m) = \pm 1$ and $\sigma_2(\beta/\alpha^m) = \pm 1$. Moreover, they must admit the same sign, since ± 1 are rational numbers.

All we have to prove is that there exists a nontrivial unit. This is equivalent to show that there is a solution to Pell's equation

$$x^2 - dy^2 = 1, \qquad x, y \in \mathbf{Z}, y \neq 0.$$

Theorem 21. *Pell's equation has nontrivial solutions.*

Proof. By Dirichlet's approximation theorem, there are infinitely many rational numbers x_n/y_n ($x_n, y_n \to \infty$) satisfying

$$\left| \frac{x_n}{y_n} - \sqrt{d}_{\mathbf{R}} \right| < 1/y_n^2.$$

Fix such an infinite sequence of x_n/y_n. Then $|x_n - \sqrt{d}_{\mathbf{R}} y_n| < 1/y_n$. Moreover, $|x_n| < |\sqrt{d}_{\mathbf{R}} y_n| + 1$, which implies that $|x_n + \sqrt{d}_{\mathbf{R}} y_n| < T y_n$ for some real number T (depending only on d). Therefore,

$$|x_n^2 - dy_n^2| < T$$

for all n. Since $|x_n^2 - dy_n^2|$ is an integer, there exists $t \in \mathbf{Z}$ such that $x_n^2 - dy_n^2 = t$ for infinitely many n. In an appropriate subsequent, we may assume that $x_n^2 - dy_n^2 = t$ for all n and also that (x_n) and (y_n) are both constant modulo t. Then in $\mathbf{Q}(\sqrt{d})$,

$$\frac{x_m + \sqrt{d}y_m}{x_n + \sqrt{d}y_n} = \frac{x_m + \sqrt{d}y_m}{x_n + \sqrt{d}y_n} \cdot \frac{x_n - \sqrt{d}y_n}{x_n - \sqrt{d}y_n} = \frac{x_m x_n - d y_m y_n + \sqrt{d}(-x_m y_n + x_n y_m)}{x_n^2 - dy_n^2} = X_{m,n} + \sqrt{d} Y_{m,n},$$

where $X_{m,n}, Y_{m,n} \in \mathbf{Z}$, since the denominator is t, and in the numerator, both $x_m x_n - d y_m y_n$ and $-x_m y_n + x_n y_m$ are 0 modulo t. Now

$$X_{m,n}^2 - dY_{m,n}^2 = N(X_{m,n} + \sqrt{d} Y_{m,n}) = \frac{N(x_m + \sqrt{d}y_m)}{N(x_n + \sqrt{d}y_n)} = 1.$$

We are left to guarantee that $Y_{m,n} \neq 0$, in other words, the solution is nontrivial. Use the embedding σ_1, this means $\sigma_1(\sqrt{d}) > 0$. Fixing n and tending to ∞ with m, we see that $\sigma_1(x_m + \sqrt{d}y_m) \to \infty$. Then $\sigma_1(X_{m,n} + \sqrt{d} Y_{m,n}) \to \infty$. We see that this cannot happen if $Y_{m,n} = 0$, since it would imply $X_{m,n} = \pm 1$. □

We see now that even the description of the units needed a careful analysis. Gauss conjectured something in this case as well.

Conjecture 22 (Gauss, 1801)**.** *Unique factorization holds for infinitely many positive d.*

This is still open.

5. The nature of primes

In this section, we assume that unique factorization holds in $\mathcal{O}(\sqrt{d})$. (Although, according to the current state of art, we cannot be sure that there are infinitely many such d.)

Take any prime number $p \in \mathbf{Z}$. We are interested in the behavior of p in $\mathcal{O}(\sqrt{d})$. By unique factorization in $\mathcal{O}(\sqrt{d})$, p decomposes into a product of primes, say, $p = \text{unit} \cdot \pi_1^{k_1} \cdot \ldots \cdot \pi_r^{k_r}$. Since $N(p) = p^2$, $1 < |N(\pi_j)|$ and $N(p) = N(\pi_1)^{k_1} \cdot \ldots \cdot N(\pi_r)^{k_r}$, there are three possibilities:

(1) $r = 1$, $p = \text{unit} \cdot \pi^2$: in this case, p is called a *ramified* prime (in the extension $[\mathbf{Q}(\sqrt{d}) : \mathbf{Q}]$);
(2) $r = 1$, $p = \text{unit} \cdot \pi$: in this case, p remains a prime in the extension, and is called an *inert* prime;
(3) $r = 2$, $p = \text{unit} \cdot \pi_1 \pi_2$: in this case, p is called a *split* prime.

Problem 19. What are the ramified, split, inert primes for $d = -1, -3$?

Problem 20. Prove that if $p \nmid 2d$, then p cannot be a ramified prime. Conclude that there are finitely many ramified primes.

What about the number of inert and splitting primes?

Fact 23 (Chebotarev's density theorem)**.** *The density of split and inert primes are the same, i.e.*

$$\lim_{N\to\infty} \frac{\#\{p < N : p \text{ is split}\}}{\#\{p < N\}} = \lim_{N\to\infty} \frac{\#\{p < N : p \text{ is inert}\}}{\#\{p < N\}} = \frac{1}{2}.$$

This can be considered as a generalization of Dirichlet's theorem about primes in arithmetic progressions, at least when it is formulated for field extensions more general than quadratic ones. In fact, from the current formulation, Dirichlet's theorem follows for the arithmetic progressions $4k \pm 1$, $3k \pm 1$. For Dirichlet's theorem in its full strength, we have to apply Chebotarev's theorem for the so-called cyclotomic extensions.

6. Some more problems

Problem 21. Prove that unique factorization holds in $\mathcal{O}(\sqrt{d})$ for $d \in \{-2, -3, -7\}$ and for $d \in \{2, 3\}$.

Problem 22. Solve the Diophantine equation $x^2 - 1 = y^3$.

Problem 23. Solve the Diophantine equation $x^2 + 2 = y^3$.

If $d > 0$, the signature of an element $0 \neq \alpha \in \mathbf{Q}(\sqrt{d})$ is defined as $(\text{sign}(\sigma_1(\alpha)), \text{sign}(\sigma_1(\alpha)))$. Therefore, the signature of an element is in the set $\{(+,+),(+,-),(-,-),(-,+)\}$. An element is said to be totally positive (resp. totally negative), if its signature is $(+,+)$ (resp. $(-,-)$).

Exercise 24. Let $d > 0$. Show that in $\mathbf{Q}(\sqrt{d})$ and even more in $\mathcal{O}(\sqrt{d})$, there are elements of any given signature.

Exercise 25. Show that in $\mathcal{O}(\sqrt{2})$, there are units of any given signature. Show that in $\mathcal{O}(\sqrt{3})$, each unit is either totally positive or totally negative.