Part one

Translator's introduction

Translator's introduction

0.1 General remarks

Dedekind's invention of ideals in the 1870s was a major turning point in the development of algebra. His aim was to apply ideals to number theory, but to do this he had to build the whole framework of commutative algebra: fields, rings, modules and vector spaces. These concepts, together with groups, were to form the core of the future abstract algebra. At the same time, he created algebraic number theory, which became the temporary home of algebra while its core concepts were growing up. Algebra finally became independent in the 1920s, when fields, rings and modules were generalised beyond the realm of numbers by Emmy Noether and Emil Artin. But even then, Emmy Noether used to say "Es steht schon bei Dedekind" ("It's already in Dedekind"), and urged her students to read all of Dedekind's works in ideal theory.

Today this is still worthwhile, but not so easy. Dedekind wrote for an audience that knew number theory – especially quadratic forms – but not the concepts of ring, field or module. Today's readers probably have the opposite qualifications, and of course most are not fluent in German and French. In an attempt to overcome these problems, I have translated the most accessible of Dedekind's works on ideal theory, *Sur la Théorie des Nombres Entiers Algébriques*, Dedekind (1877), which he wrote to explain his ideas to a general mathematical audience. This memoir shows the need for ideals in a very concrete case, the numbers $m + n\sqrt{-5}$ where $m, n \in \mathbb{Z}$, before going on to develop a general theory and to prove the theorem on unique factorisation into prime ideals.

The algebraic integers in Dedekind's title are a generalisation of the ordinary integers – created in response to certain limitations of classical number theory. The ordinary integers have been studied since ancient times, and their basic theory was laid down in Euclid's *Elements* (see Heath (1925)) around 300 BC. Yet even ancient number theory contains problems not solvable by Euclid's methods. Sometimes it is necessary to use irrational numbers, such as $\sqrt{2}$, to answer questions about the ordinary integers. A famous example is the so-called Pell equation

$$x^2 - cy^2 = 1$$

where c is a nonsquare integer and the solutions x, y are required to be integers. Solutions for certain values of c were known to the ancients, but the complete solution was not obtained until Lagrange (1768) related the equation to the continued fraction expansion of $\sqrt{c}$. He also showed that each solution is obtained from a certain "minimum" solution (x_0, y_0) by the formula

$$x_k + y_k\sqrt{c} = \pm(x_0 + y_0\sqrt{c})^k.$$

The irrational numbers $x_k + y_k\sqrt{c}$ in this formula are examples of *algebraic integers*, which are defined in general to be roots of equations of the form

$$\alpha^n + a_{n-1}\alpha^{n-1} + \cdots + a_1\alpha + a_0 = 0$$

where $a_{n-1}, \ldots, a_0$ are ordinary integers, that is, $a_{n-1}, \ldots, a_0 \in \mathbb{Z}$.

Algebraic integers are so called because they share some properties with the ordinary integers. In particular, they are closed under sum, difference and product, and the rational algebraic integers are just the ordinary integers (for more details, see 0.6.2 and §13 of Dedekind's memoir). Because of the second fact, the ordinary integers are also known as *rational integers*. The first fact implies that the algebraic integers form a ring. However, we are not interested in the ring of all algebraic integers so much as rings like

$$\mathbb{Z}[\sqrt{2}] = \{x + y\sqrt{2} : x, y \in \mathbb{Z}\}$$

and

$$\mathbb{Z}[i] = \{x + yi : x, y \in \mathbb{Z}\}.$$

In general we use the notation $\mathbb{Z}[\alpha]$ to denote the closure of the set $\mathbb{Z} \cup \{\alpha\}$ under +, − and ×. The reason for working in rings $\mathbb{Z}[\alpha]$ is that they more closely resemble $\mathbb{Z}$, and hence are more likely to yield information about $\mathbb{Z}$.

Any ring R of algebraic integers includes $\mathbb{Z}$, so theorems about $\mathbb{Z}$ may be obtainable as special cases of theorems about R (we shall see several examples later). However, useful theorems about R are provable only when R has all the basic properties of $\mathbb{Z}$, in particular, *unique prime factorisation.* This is not always the case. $\mathbb{Z}[\sqrt{-5}]$ is the simplest example where unique prime factorisation fails, and this is why Dedekind studies it in detail. His aim is to recapture unique prime factorisation by extending the concept of integer still further, to certain sets of algebraic integers he calls *ideals.* This works only if the size of R is limited in some way. The ring A of all algebraic integers is "too big" because it includes $\sqrt{\alpha}$ along with each algebraic integer α. This gives the factorisation $\alpha = \sqrt{\alpha}\sqrt{\alpha}$ and hence "primes" do not exist in A, let alone unique prime factorisation.

Dedekind found the appropriate "small" rings R in algebraic number fields of finite degree, each of which has the form $\mathbb{Q}(\alpha)$, where α is an algebraic integer. $\mathbb{Q}(\alpha)$ denotes the closure of $\mathbb{Q} \cup \{\alpha\}$ under $+,-,\times$ and $\div$ (by a nonzero number), and each $\mathbb{Q}(\alpha)$ has its own integers, which factorise into primes. In particular, $\mathbb{Z}[\sqrt{-5}]$ is the ring of integers of $\mathbb{Q}(\sqrt{-5})$, and $6 = 2 \times 3$ is a prime factorisation of 6. Not *the* prime factorisation, alas, because $6 = (1+\sqrt{-5})(1-\sqrt{-5})$ is also a factorisation into primes (see 0.4.5). However, unique prime factorisation is regained when one passes to the ideals of $\mathbb{Z}[\sqrt{-5}]$, and Dedekind generalises this to any $\mathbb{Q}(\alpha)$. The result is at last a theory of algebraic integers capable of yielding information about ordinary integers.

A lot of machinery is needed to build this theory, but Dedekind explains it well. Suffice to say that fields, rings and modules arise very naturally as sets of numbers closed under the basic operations of arithmetic. Fields are closed under $+$, $-$, $\times$ and $\div$, rings are closed under $+$, $-$ and $\times$, while modules are closed under $+$ and $-$. The term "ring" was actually introduced by Hilbert (1897); Dedekind calls them "domains" here, and I have thought it appropriate to retain this terminology, since these particular rings are prototypes of what are now called *Dedekind domains.* Dedekind presumably chose the name "module" because a module M is something for which "congruence modulo M" is meaningful. His name for field, Körper (which also means "body" in German), was chosen to describe "a system with a certain completeness, fullness and self-containedness; a naturally unified, organic whole", as he explained in his final exposition of ideal theory, Dedekind (1894), §160.

What Dedekind does not explain is where $\mathbb{Z}[\sqrt{-5}]$ comes from, and why it is important in number theory. This is understandable, because his first version of ideal theory was a supplement to Dirichlet's number theory lectures, *Vorlesungen über Zahlentheorie* (Dirichlet (1871)). In the present memoir he also refers to the *Vorlesungen* frequently, so his original audience was assumed to have a good background in number theory, and particularly the theory of quadratic forms. Such a background is less common today, but is easy and fun to acquire. Even experts may be surprised to learn how far back the story goes. The specific role of $\sqrt{-5}$ can be traced back to the anomalous behaviour of the quadratic form x^2+5y^2, first noticed by Fermat, and later explained in different ways by Lagrange, Gauss and Kummer. But the reason for Fermat's interest in x^2+5y^2 goes back much further, perhaps to the prehistory of mathematics in ancient Babylon. Let us begin there.

0.2 Squares

0.2.1 Pythagorean triples

Integers a, b, c such that

$$a^2+b^2=c^2$$

are one of the oldest treasures of mathematics. Such numbers occur as the sides of right-angled triangles, and they may even have been used to construct right angles in ancient times. They are called Pythagorean triples after Pythagoras, but they were actually discovered independently in several different cultures. The Babylonians were fascinated by them as early as 1800 BC, when they recorded fifteen of them on a tablet now known as Plimpton 322 (see Neugebauer and Sachs (1945)). Pythagorean triples other than the simplest ones (3,4,5), (5,12,13) or (8,15,17) are not easily found by trial and error, so the Babylonians probably knew a general formula such as

$$a=2uv, \quad b=u^2-v^2, \quad c=u^2+v^2$$

which yields an unlimited supply of Pythagorean triples by substituting different integers for u, v.

The general solution of $a^2+b^2=c^2$ is in fact

$$a=2uvw, \quad b=(u^2-v^2)w, \quad c=(u^2+v^2)w,$$

as may be found in Euclid's *Elements* Book X (lemma after Proposition 29). A key statement in Euclid's proof is: *if the product of relatively prime integers is a square, then the integers themselves are squares.*

Euclid first used the general formula for Pythagorean triples in his theory of irrational numbers, and it is in a different book from his theory of integers. The assumption that relatively prime integers are squares when their product is a square is justified by a long chain of propositions, stretching over several books of the *Elements*. However, a direct justification is possible from his theory of integer divisibility, which is in Book VII. This theory is fundamental to the theory of ordinary integers, and also the inspiration for Dedekind's theory of ideals, so we should recall its main features before going any further. Among other things, it identifies the important but elusive role of primes.

0.2.2 Divisors and prime factorisation

An integer m *divides* an integer n if $n = ml$ for some integer l. We also say that m is a *divisor* of n, or that n is a *multiple* of m. An integer p whose only divisors are ± 1 and $\pm p$ is called a *prime*, and any integer can be factorised into a finite number of primes by successively finding divisors unequal to ± 1 but of minimal absolute value. However, it is not obvious that each factorisation of an integer n involves the *same* set of primes. There is conceivably a factorisation of some integer

$$n = p_1 p_2 \cdots p_i = q_1 q_2 \cdots q_j$$

into primes $p_1, p_2, \ldots, p_i$ and $q_1, q_2, \ldots, q_j$ respectively, where one of the primes p is different from all the primes q.

Nonunique prime factorisation is ruled out by the following proposition of Euclid (*Elements*, Book VII, Proposition 30).

Prime divisor property. *If p is prime and p divides the product ab of integers a, b, then p divides a or p divides b.*

An interesting aspect of the proof is its reliance on the concept of *greatest common divisor* (gcd), particularly the fact that

$$\gcd(a, b) = ua + vb \quad \text{for some integers } u, v.$$

The set $\{ua + vb : u, v \in \mathbb{Z}\}$ is in fact an *ideal*, and unique prime factorisation is equivalent to the fact that this ideal consists of the multiples of one of its members, namely $\gcd(a, b)$.

It should be mentioned that Euclid proves only the prime divisor property, not unique prime factorisation. In fact its first explicit statement and proof are in Gauss (1801), the *Disquisitiones Arithmeticae*, article 16. As we shall see, this is possibly because Gauss was first to recognise generalisations of the integers for which unique prime factorisation is not valid.

0.2.3 Irrational numbers

As everybody knows, Pythagorean triples also have significance as the sides of right-angled triangles. In any right-angled triangle, the side lengths a, b, c satisfy

$$a^2 + b^2 = c^2$$

whether or not a, b and c are integers (Pythagoras' theorem). Hence it is tempting to try to interpret a right-angled triangle as a Pythagorean triple by choosing the unit of length so that a, b and c all become integer lengths. Pythagoras or one of his followers made the historic discovery that this is not always possible. The simplest counterexample is the triangle with sides 1, 1, $\sqrt{2}$. It is impossible to interpret this triangle as a Pythagorean triple because $\sqrt{2}$ is not a rational number.

A proof of this fact, which also proves the irrationality of $\sqrt{3}$, $\sqrt{5}$, $\sqrt{6}$ and so on, uses unique prime factorisation to see that each prime appears to an even power in a square. Then the equation

$$2n^2 = m^2$$

is impossible because the prime 2 occurs an odd number of times in the prime factorisation of $2n^2$, and an even number of times in the prime factorisation of m^2.

The irrationality of $\sqrt{2}$ led the Greeks to study the so-called *Pell equation*

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

They found it could used to approach $\sqrt{2}$ rationally, via increasingly large integer solutions, x_n, y_n. Since $x_n^2 - 2y_n^2 = 1$, the quotient x_n/y_n necessarily tends to $\sqrt{2}$. The general Pell equation

$$x^2 - cy^2 = 1, \quad \text{where } c \text{ is a nonsquare integer,}$$

can similarly be used to approach the irrational number $\sqrt{c}$. This equation later proved fruitful in many other ways; Dedekind even used it to prove the irrationality of $\sqrt{c}$ (Dedekind (1872), Section IV).

0.2.4 Diophantus

The equations $a^2+b^2 = c^2$ and $x^2-2y^2 = 1$ are examples of what we now call *Diophantine equations*, after Diophantus of Alexandria. Diophantus lived sometime between 150 AD and 350 AD and wrote a collection of books on number theory known as the *Arithmetica* (Heath (1910)). They consist entirely of equations and ingenious particular solutions. The term "Diophantine" refers to the type of solution sought: either rational or integer. For Diophantus it is usually a rational solution, but for some equations, such as the Pell equation, the integer solutions are of more interest. The Pell equation was actually not studied by Diophantus, but he mentioned an integer solution to another remarkable equation: the solution $x = 5$, $y = 3$ of $y^3 = x^2 + 2$. (See 0.4.1 for the astonishing sequel to this solution.)

Although all Diophantus' solutions are special cases, they usually seem chosen to illustrate general methods. Euler (1756) went so far as to say

> Nevertheless, the actual methods that he uses for solving any of his problems are as general as those in use today ... there is hardly any method yet invented in this kind of analysis not already traceable to Diophantus. (Euler *Opera Omnia* I,2, p. 429–430.)

And if anyone would know, Euler would. The first mathematician to understand Diophantus properly was Fermat (1601–1665), but his comments were as cryptic as the *Arithmetica* itself. Euler spent about 40 years, off and on, reading between the lines of Fermat and Diophantus, until he could reconstruct most their methods and prove their theorems. We shall study the connection between Diophantus, Fermat and Euler more thoroughly later, but one example is worth mentioning here. It shows how much theory can be latent in a single numerical fact.

In the *Arithmetica*, Book III, Problem 19, Diophantus remarks

> 65 is naturally divided into two squares in two ways, namely into $7^2 + 4^2$ and $8^2 + 1^2$, which is due to the fact that 65 is the product of 13 and 5, each of which is the sum of two squares.

It appears from this that Diophantus is aware of the identity

$$(a^2 + b^2)(c^2 + d^2) = (ac \pm bd)^2 + (ad \mp bc)^2$$

though he makes no such general statement. However, Fermat saw much deeper than this. Noticing, with Diophantus, that the identity reduces the representations of a number as a sum of two squares to the representations of its prime factors, his comment on the problem is:

> A prime number of the form $4n+1$ is the hypotenuse of a right-angled triangle [that is, a sum of squares] in one way only . . . If a prime number which is the sum of two squares be multiplied by another prime number which is also the sum of two squares, the product will be the sum of two squares in two ways. (Heath (1910), p. 268.)

The restriction to primes of the form $4n+1$ is understandable because a prime $p \neq 2$ cannot be the sum of two squares unless it is of the form $4n+1$ (by a congruence mod 4 argument). But Fermat's claim that any prime $p = 4n+1$ is a sum of two squares comes right out of the blue. No one knows how he proved it and the first known proof is due to Euler (1756). As we shall see later, Lagrange, Gauss and Dedekind all used this theorem of Fermat to test the strength of new methods in number theory.

0.3 Quadratic forms

0.3.1 Fermat

Unlike Euclid or Diophantus, Fermat never wrote a book. His reputation rests on a short manuscript containing his discovery of coordinate geometry (independent of Descartes), his letters, and his marginal notes on Diophantus. He took up number theory only in his late 30s, and left only one reasonably complete proof, in the posthumously published Fermat (1670). However, it is a beautiful piece of work, and fully establishes his credentials as both an innovator and a student of the ancients. It also has a place in our story, as an application of Pythagorean triples, and as the first proven instance of Fermat's last theorem. Fermat's proof shows that there are no positive integers x, y, z such that $x^4+y^4=z^4$, by showing that there are not even positive integers x, y, z such that $x^4+y^4=z^2$. It turns out to be the only instance of Fermat's last theorem with a really elementary proof, involving just Euclid's theory of divisibility.

The argument is by contradiction, and the gist of it is as follows (omitting mainly routine checks that certain integers are relatively prime).

Suppose that there are positive integers x, y, z such that $x^4+y^4=z^2$, or in other words, $(x^2)^2+(y^2)^2=z^2$. This says that x^2, y^2, z is a Pythagorean triple, which we can take to be primitive, hence there are integers u, v such that

$$x^2 = 2uv, \quad y^2 = u^2 - v^2, \quad z = u^2 + v^2,$$

by Euclid's formulas (0.2.1). The middle equation says that v, y, u is also a Pythagorean triple, and it is also primitive, hence there are integers s, t such that

$$v = 2st, \quad y = s^2 - t^2, \quad u = s^2 + t^2.$$

This gives

$$x^2 = 2uv = 4st(s^2 + t^2),$$

so the relatively prime integers s, t and $s^2 + t^2$ have product equal to the square $(x/2)^2$. It follows that each is itself a square, say

$$s = x_1^2, \quad t = y_1^2, \quad s^2 + t^2 = z_1^2,$$

and hence

$$x_1^4 + y_1^4 = z_1^2.$$

Thus we have found another sum of two fourth powers equal to a square, and by retracing the argument we find that the new square z_1^2 is *smaller* than the old, z^2, but still nonzero. By repeating the process we can therefore obtain an infinite descending sequence of positive integers, which is a contradiction. □

Fermat called the method used in this proof *infinite descent*, and used it for many of his other theorems. He claimed, for example, to have proved that any prime of the form $4n + 1$ is a sum of two squares by supposing $p = 4n + 1$ to be a prime not the sum of two squares, and finding a smaller prime with the same property. However, it is very hard to see how to make the descent in this case. Euler (1749) found a proof only after several years of effort. In 0.3.4 we shall see an easier proof of the two squares theorem due to Lagrange. Lagrange's proof does use another famous theorem of Fermat, but it is the easy one known as Fermat's "little" theorem: *for any prime number p, and any integer $a \not\equiv 0 \pmod p$, we have $a^{p-1} \equiv 1 \pmod p$* (Fermat (1640b)).

The proof of Fermat's little theorem most likely used by Fermat uses induction on a and the fact that a prime p divides each of the binomial coefficients

$$\binom{p}{i} = \frac{p(p-1)(p-2)\cdots(p-i+1)}{i!} \quad \text{for} \quad 1 \le i \le p-1,$$

as is clear from the fact that p is a factor of the numerator but not of the denominator. This proof implicitly contains the "mod p binomial theorem",

$$(a+b)^p \equiv a^p + b^p \pmod p,$$

which has its uses elsewhere (see for example Gauss's proof of quadratic reciprocity in 0.5.4).

The proof more often seen today is based on that of Euler (1761),