An invitation to algebraic number theory

B.Sury

These are expanded notes of lectures delivered in the Advanced Foundational School II of the Indian Statistical Institute Bangalore during May 7 - June 2, 2007. We have already had many lectures on commutative algebra in this school. For the purpose of this school, methods from commutative algebra used to study number theoretic questions may be called algebraic number theory. We will not mention any analytic component of the subject like the Dedekind zeta function etc. The subject arose basically from the need to solve one particular problem in number theory - Fermat's last theorem. We recall the basic concepts and results and discuss some applications in traditional number theory along the way. One can use them often to solve Diophantine equations. The basic objects of study are certain rings similar to (but more general than) $\mathbf{Z}$ like :
$\{a+b\sqrt{d} : a, b \in \mathbf{Z}\}$ for some square-free integer d,
$\{a_0+a_1\zeta_p+\cdots+a_{p-2}\zeta_p^{p-2} : a_i \in \mathbf{Z}\}$ where p is a prime number and $\zeta_p = e^{2i\pi/p}$.
The basic departure from the usual $\mathbf{Z}$ that concerns us is that such rings may not have unique factorisation. It is worth recalling just a whiff of history. For this, we decant from an article by Israel Kleiner on the history of ring theory which appeared in Elemente der Mathematik in 1998. In a ring like the latter one mentioned above which figures in Fermat's equation, Kummer realized that unique factorisation may not hold (we will see below the example of $p = 23$) but he wanted to establish unique factorisation of some kind. As he wrote to Liouville, "unique factorisation can be saved by the introduction of new kind of complex numbers which I have called ideal numbers." Kummer's result was that every element is a unique product of "ideal primes." However, the notion of "ideal numbers" was implicit and unclear. Kummer said that they were like 'free radicals' in chemistry which can be discerned only by their effects. Moreover, Kummer's decomposition theory into 'ideal primes' was valid only for these above rings. It was Dedekind who devised a way to define such a notion for general rings of algebraic integers. To begin with, he showed that in a field $\mathbf{Q}(\alpha)$ where α is an algebraic integer, the ring $\mathbf{Z}[\alpha]$ which was hitherto studied was the wrong object ! For example, he showed that Kummer's theory did not apply correctly to the ring $\mathbf{Z}[\sqrt{-3}]$. At that time, the notion of algebraic numbers was well-understood but not that of algebraic integers. He showed that the right object in $\mathbf{Q}(\alpha)$ which had similar properties to $\mathbf{Z}$ was the ring of *all* algebraic integers there - the integral closure in modern language. He showed that instead of Kummer's ideal number, one needs to look at the whole set of algebraic integers 'divisible' by such an ideal number. This was the birth of the theory of ideals. Indeed, Dedekind defined for the first time rings axiomatically (although it was done only inside number fields and he called them 'orders'). To cap it, Dedekind defined perhaps the most important notion of commutative algebra - that of prime ideals. The important point about his definition of

'ideals' is their intrinsic definition. He defined an ideal I to divide another ideal J if I contains J. Later, he refined it by showing that this is the same as saying that $J = IK$ for some ideal K. Using the definition of prime ideals, he was able to prove that every non-zero ideal in the full ring of algebraic integers in a number field is a product of prime ideals in a unique manner. These are the notions and the notations prevalent in the modern day. Dedekind believed in focussing on intrinsic or conceptual properties rather than concrete representations or formulae. His work was the culmination of 70 years of work on unique factorisation by several mathematicians and led to the birth of algebraic number theory.

We first start with formal definitions. Some of these are already covered in the earlier lectures on commutative algebra but they bear repetition. Moreover, for the sake of completeness, these notes are far more detailed than what can actually be covered in this $\frac{1}{3}$-course.

Before beginning in earnest, let us whet our appetite by proving the following beautiful theorem due to Sophie Germain :

Theorem (Sophie Germain).
Assume $p > 2$ is a prime such that $2p + 1$ is also a prime. Then, if x, y, z are integers satisfying the equation $x^p + y^p + z^p = 0$, then $p|xyz$.

Proof.
Suppose p does not divide any of x, y, z. Then, we may assume that x, y, z are pairwise coprime. In that case, it is easy to see that $x + y$ and $\frac{x^p+y^p}{x+y}$ are coprime as well. Therefore $x + y = a^p, \frac{x^p+y^p}{x+y} = A^p$ for certain integers a, A. So, $x^p + y^p = (aA)^p$. Similarly, we have $y + z = b^p, z + x = c^p$ and $y^p + z^p = (bB)^p, z^p + x^p = (cC)^p$. Now, if $2p + 1$ does not divide x, then x^p *equiv* ± 1 mod $2p + 1$ by Fermat's little theorem. Thus, if $2p + 1$ does not divide xyz, then $x^p + y^p + z^p \equiv \pm 1 \pm 1 \pm 1$ mod $2p + 1$ and this can never be zero as $2p + 1 \neq 3$. Thus, $2p + 1$ divides one of them, say x. Therefore, it does not divide yz. Hence

$$b^p = y + z \equiv (x + y) + (x + z) = a^p + c^p \ \ mod \ \ 2p + 1.$$

This means once again as before that one of a, b, c must be divisible by $2p + 1$. As $2p+1$ divides x and not y and z, it is clear that $2p+1$ divides b. Thus, $y+z$ is a multiple of $2p+1$ which we write as $z \equiv -y$ mod $2p+1$. As $y^p+z^p = (bB)^p$, we get

$$B^p = y^{p-1} - y^{p-2}z + \cdots + z^{p-1} \equiv py^{p-1} \ \ mod \ \ 2p + 1.$$

As $2p + 1$ does not divide B, we get

$$py^{p-1} \equiv \pm 1 \ \ mod \ \ 2p + 1.$$

So, modulo $2p + 1$, $\pm y \equiv py^p \equiv \pm p$.
So, $p \equiv \pm y \equiv \pm(x + y) = \pm a^p \equiv \pm 1$ mod $2p + 1$, an impossibility.

Definition. A *Dedekind domain* (DD for short) is an integral domain A in which all ideals are finitely generated (i.e., A is Noetherian), elements of the quotient field of A which are not in A do not satisfy any monic polynomial over A (i.e., A is integrally closed) and all nonzero prime ideals are maximals (i.e., A is of dimension 1.)

Exercises.
(a) Any PID is a DD.
(b) $\mathbf{Z}[X]$ *is not a DD.*

A finite extension K of $\mathbf{Q}$ is usually called an algebraic number field and the integral closure of $\mathbf{Z}$ in K (that is, the set of $a \in K$ which satisfy a monic integral polynomial) is called its ring of integers. It is easy to see that K is the quotient field of its ring of integers $\mathcal{O}_K$. Also, each nonzero ideal I in $\mathcal{O}_K$ contains some natural number and if I is a prime ideal, it contains a unique prime number.
For, if $0 \neq a \in I$ satisfies $a^n + a_{n-1}a^{n-1} + \cdots + a_1 a + a_0 = 0$ with $a_i \in \mathbf{Z}$, then $a_0 \in I \cap \mathbf{Z}$. Moreover, if I is prime, some prime factor of a_0 must be in I; if there are two such primes p, q then the GCD 1 would also be in I.

Proposition. *The ring of integers in any algebraic number field is a DD. More generally, if A is a DD and L is a finite extension of the quotient field K of A, then the integral closure B of A in L is a DD.*
Proof when L/K is separable.
We already know that B must have dimension 1 and must be integrally closed. To show that B is Noetherian, we prove the stronger statement that B is an A-submodule of a free A-module of rank $n = [L : K]$. If this is proved, it would follow that B is a Noetherian A-module. Any ideal of B is, in particular, an A-submodule of B and, therefore, finitely generated as an A-module (and therefore as a B-module). Thus, it suffices to show that B is an A-submodule of a free A- module of rank n. To see this, let $e_1, \cdots, e_n$ be any K-basis of L lying in B (Why is it possible to choose such a basis?). Then, if $e_1^*, \cdots, e_n^*$ is its dual basis with respect to the trace form i.e., if $Tr_K^L(e_i e_j^*) = \delta_{ij}$, then any $x \in L$ is of the form $\sum_i Tr(xe_i)e_i^*$. If $x \in B$, then all the coefficients $Tr(xe_i) \in A$ as they are integral over A. Therefore, $B \subset \sum_i Ae_i$ which is a free A-module of rank n (as e_i's are linearly independent over K). Thus, the proof is complete.

Remarks
It can happen that $\mathcal{O}_L$ is not a free $\mathcal{O}_K$-module for number fields $K \subset L$.

Definition. If A is an integral domain and if K denotes its quotient field, one defines a *fractional ideal* to be a non-zero A-submodule I of K such that $I \subset d^{-1}A$ for some $d \neq 0$ in A. A *principal fractional ideal* is the A-module xA for some $x \in K$.

Remarks.
(a) Each finitely generated A-submodule of K is a fractional ideal.
(b) If A is Noetherian, then fractional ideals are none other than the finitely generated A-submodules of K.

(c) If I, J are fractional ideals, then so are $I \cap J, I + J, IJ$. Moreover, $IJ = JI$ and $I(JK) = (IJ)K$.

Warning :

Although fractional ideals have several properties similar to usual ideals, it is not true generally that $IJ \subset I \cap J$.

Proposition. *Let A be a DD and let P be a non-zero prime (= maximal) ideal. If K denotes the quotient field of A, then the set*

$$P' := \{x \in K : xP \subset A\}$$

is a fractional ideal of A and properly contains A. Further, P' is the unique fractional ideal such that $PP' = P'P = A$.

Proof.

Proof. It is trivial to see that P' is an A-module. Moreover, evidently $P' \subset d^{-1}A$ for any $d \neq 0$ in P. Thus, P' is a fractional ideal and clearly contains A. We shall show now that $A \neq P'$. For this, we make use of the following:

Claim: Every non-zero ideal of A contains a finite product of non-zero prime ideals.

The claim is proved as follows. If there are exceptions to the claim made above, consider the family of ideals which fail to contain a product as claimed. As A is Noetherian, there exists a maximal such ideal M. So, M itself cannot be prime. If $ab \in M$ with neither a nor b in M, then the ideals $M + (a)$ and $M + (b)$ contain products of prime ideals. As M contains their product, M contains a product of prime ideals, which contradicts our assumption. Therefore, the claim is indeed true. Now, let $a \neq 0$ be in P. Then, the ideal $(a) \supset P_1 P_2 \cdots P_n$ with n minimal possible and P_i's non-zero primes. So, $P \supset P_1 \cdots P_n$. As P is prime, we have $P \supset P_i$ for some i, say $P \supset P_1$. As P_i are maximal, we obtain $P = P_1$. Writing $I = P_2 \cdots P_n$ or A according as $n > 1$ or $n = 1$, we get $I \not\subset (a)$ by the minimality of n. Choose any $b \in I \setminus (a)$. Then, $ba^{-1} \notin A$. Now, $PI \subset (a) \Rightarrow Pb \subset (a)$ i.e., $ba^{-1} \in P'$. Hence, we have shown that $A \neq P'$. Further, we have $P = PA \subset PP' \subset A$ so that PP' is an (actual) ideal of A containing P. It must, therefore, be either equal to P or to the unit ideal A. If $x \in P' \setminus A$, we must have $xP \not\subset P$. The reason is that, otherwise, $A[x]P \subset P$ and $A[x]$ would be a finitely generated A-module (and so x is integral over A), a contradiction of the fact that $x \notin A$. This means that $xp \in P'P \setminus P$ for some $p \in P$. Thus, $PP' = A$. Finally, if P_0 is any fractional ideal such that $PP_0 = P_0P = A$, then $P' = AP' = (P_0P)P' = P_0(PP') = P_0A = P_0$ which proves uniqueness also.

Notation. One uses the notation P^{-n} instead of P'^n for any n. Then, (like ideals) one has $AP^{-n} = P^{-n}$.

Theorem. *Let A be a DD. Then, any fractional ideal $I \neq A$ can be uniquely written as $I = P_1^{n_1} \cdots P_k^{n_k}$ where n_i are non-zero integers and P_i are distinct prime ideals.*

In other words, the unique factorisation can be regained in the sense of ideals in the ring of integers of a number field.

Proof.
The uniqueness is easy to prove as follows.
If $P_1^{n_1} \cdots P_k^{n_k} = Q_1^{m_1} \cdots Q_r^{m_r}$, then one can shift all the negative powers on each side to the other side to obtain an equality where all powers are positive. Then, a simple induction on the sum of the exponents yields uniqueness.
We prove the existence of the prime ideal decomposition by contradiction. First, we assume that there is an (actual) ideal I which is not expressible as a product of prime ideals. By using the fact that A is Noetherien, we obtain an ideal I which is maximal with respect to this property. Of course, I is not a prime ideal. If $I \subset P$ with P maximal, then $I = AI \subset P^{-1}I \subset P^{-1}P = A$. Now, if $x \in P^{-1} \setminus A$, then $xI \not\subset I$ by the argument we saw earlier and, so $xi \subset P^{-1}I \setminus I$ for some $i \in I$. Hence $P^{-1}I$ is an (actual) ideal which contains I properly. By the choice of I, we obtain that $P^{-1}I$ must be a product of prime ideals. Therefore, clearly I itself is such a product, which manifestly contradicts the choice of I. Therefore, every ideal in A is, indeed, a product of prime ideals.
Finally, if J is any fractional ideal, there is some $d \neq 0$ in A such that dJ is an ideal of A. So, if $(d) = P_1^{a_1} \cdots P_r^{a_r}$ and $dJ = Q_1^{b_1} \cdots Q_s^{b_s}$, then $J = P_1^{-a_1} \cdots P_r^{-a_r} Q_1^{b_1} \cdots Q_s^{b_s}$. This proves the theorem.

Remarks. (a) Thus, the fractional ideals of an algebraic number field K form a group and the quotient group modulo the subgroup of principal fractional ideals is known as the ideal class group of the number field. We shall see shortly that this is a finite group and its order is called the class number of K. Thus, a finite group measures the deviation from unique factorisation into prime elements. More generally, for a non-zero fractional ideal I, one considers the part of the ideal class group which is generated by prime ideals not dividing I. The quotient of this by the *ray* $1+I$ is the *ray class group* $Cl^I(K)$. That the ray class groups are also finite is a consequence (exercise to be discussed in tutorial session) of the finiteness of $Cl(K)$ together with an application of the Chinese remainder theorem for commutative rings.
(b) In any DD, $P \supset P^2 \supset P^3 \cdots$ is a strictly decreasing chain (exercise).
(c) Every fractional ideal in a DD can be generated by two elements one of which can be taken to be any arbitrary element.
Idea: Enough to prove this for ideals I; in this case if $a \in I$ and if $(a) = P_1^{a_1} \cdots P_r^{a_r}$ and $I = P_1^{b_1} \cdots P_r^{b_r}$, then $a_i \geq b_i \geq 0$. We can use the Chinese remainder theorem to choose an appropriate element b in I so that $I = (a, b)$.
(d) A DD which has only finitely many prime ideals is a PID.
Idea : If $P_1, \cdots P_n$ are all the prime ideals, use the Chinese remainder theorem to choose $x_i \in P_i$, $x_i \notin P_i^2$ and $x_i \equiv 1 \bmod P_j$ for $i \neq j$. Then, $P_i = (x_i)$.
(d) $\mathbf{Z}[\sqrt{-5}] := \{a + b\sqrt{-5} : a, b \in \mathbf{Z}\}$ is not a PID as it is not a UFD. Therefore, by (c), it follows that there are infinitely many prime numbers (!)
Thus, the infinitude of prime numbers is a consequence of algebraic number theory !

An application :
Let us now use the uniqueness of decomposition into ideals to solve the equation $y^2 = x^3 - 2$ in integers.

It can be proved that $\mathbf{Z}[\sqrt{-2}]$ which is the ring of integers of $\mathbf{Q}(\sqrt{-2})$ is a PID. In fact, it is even a Euclidean domain as can be proved using the size function $N(a+b\sqrt{-2}) = a^2+2b^2$.
Now, for any possible solution, let us read the equation mod 8; it follows that both x and y must be odd.
Write $x^3 = y^2+2 = (y+\sqrt{-2})(y-\sqrt{-2})$. The ideals $(y+\sqrt{-2})$ and $(y-\sqrt{-2})$ must be coprime. Otherwise we can find a (usual) integer n dividing both these elements in $\mathbf{Z}[\sqrt{-2}]$ and, this implies that n divides $2\sqrt{-2}$; then the norm of n would be a factor of 8 whereas it divides y^2+2 which is odd.
Thus, by the uniqueness of factorisation into ideals, both $(y+\sqrt{-2})$ and $(y-\sqrt{-2})$ must be cubes of ideals. As they are principal, the elements $y+\sqrt{-2}$ and $y-\sqrt{-2}$ must themselves be cubes of elements up to units. Of course, the units in $\mathbf{Z}[\sqrt{-2}]$ are just 1 and -1 both of which are units. Writing out $y+\sqrt{-2} = (a+b\sqrt{-2})^3$, we get

$$y = a^3 - 6ab^2,$$
$$1 = b(3a^2 - 2b^2).$$

It trivially follows that $b = 1$ and $a = \pm 1$. Therefore, the solutions are $(x, y) = (3, \pm 5)$.

Prime decomposition in extension fields

Let K be an algebraic number field and $\mathcal{O}_K$ be its ring of integers. One has the following very interesting fact :
Proposition.
A DD can be a UFD only if it is a PID.
Proof.
Let A be a DD which is also a UFD. To show it is a PID, it suffices to show that prime ideals are principal (Exercise - why?). Let P be a non-zero prime ideal of A and let $0 \neq a \in P$. Write $a = p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r}$ where p_i are irreducible elements. Now $p_i \in P$ for some i as P is a prime ideal. So $(p_i) \subseteq P$. But evidently (p_i) is a nonzero prime (=maximal) ideal. Hence $P = (p_i)$.

Definition. For a field extension L/K of degree n, and an n-tuple of elements $v_1, \cdots, v_n$ of L, one defines *the discriminant of the n-tuple* $v_1, \cdots, v_n$ to be the element $D_K^L(v_1, \cdots, v_n) = \det(M)$ of K where $M_{ij} = Tr_K^L(v_i v_j)$. This is an important concept, and let us start with a few easy exercises to see its use.

Exercises. *Let L, K, v_i be as above.*
(a) Show that $D_K^L(v_1, \cdots, v_n) \neq 0$ if, and only if, $v_1, \cdots, v_n$ form a K-basis of L.
(b) If $K = \mathbf{Q}$ and v_i form a $\mathbf{Z}$-basis of the ring of integers (this always exists as we observed), then $D_K^L(v_1, \cdots, v_n)$ is an integer which is independent of the choice of the $\mathbf{Z}$-basis.
(c) If $\sigma_1, \cdots, \sigma_n$ are the K-embeddings of L in $\mathbf{C}$, then $D_K^L(v_1, \cdots, v_n) = \det(N)^2$ where $N_{ij} = \sigma_i(v_j)$.

Definition. The *discriminant D_K of an algebraic number field K* is the discriminant of any $\mathbf{Z}$-basis of its ring of integers. By the exercise (b) above, it is well-defined. Moreover, it is clear that if $\{v_1, \cdots, v_n\}$ are in $\mathcal{O}_K$ and satisfy $D_K = D^K_{\mathbf{Q}}(v_1, \cdots, v_n)$, then $\{v_i\}$ form an integral basis (Why?).

Exercise. *(a) For a square-free integer d, show that the discriminant of $\mathbf{Q}(\sqrt{d})$ is d or $4d$ according as whether $d \equiv 2, 3 \ mod\ 4$ or $d \equiv 1 \ mod\ 4$.*
(b) Let $K = \mathbf{Q}(\alpha)$ be an algebraic number field. Suppose the minimal (monic) polynomial of α is $f(X) = \prod_{i=1}^{n}(X - \alpha_i)$. Then, prove that

$$D^K_{\mathbf{Q}}(1, \alpha, \cdots, \alpha^{n-1}) = \prod_{i<j}(\alpha_i - \alpha_j)^2 = (-1)^{n(n-1)/2} N^K_{\mathbf{Q}} f'(\alpha)$$

where N denotes the norm map.
(c) Use (b) to show that for any n, and $K = \mathbf{Q}(\zeta_n)$ with ζ_n a primitive n-th root of unity, one has $D^K_{\mathbf{Q}}(1, \zeta, \cdots, \zeta^{\phi(n)-1})$ divides $n^{\phi(n)}$.
(d) Let K be an algebraic number field and let $\alpha_1, \cdots, \alpha_n$ be a $\mathbf{Q}$-basis of K contained in $\mathcal{O}_K$, the ring of integers of K. Then,

$$\mathcal{O}_K \subset \{\sum m_i \alpha_i / d : m_i \in \mathbf{Z}, d | m_i^2\}$$

Here d stands for $D^K_{\mathbf{Q}}(\alpha_1, \cdots, \alpha_n)$.
Hint: Write any $\alpha \in \mathcal{O}_K$ as $\sum_i t_i \alpha_i$ with $t_i \in \mathbf{Q}$. Apply the various embeddings of K to this equation and solve the system of linear equations by Cramer's rule.
(e) If K, L have degrees m, n over $\mathbf{Q}$ and if KL has degree mn, then $\mathcal{O}_{KL} \subset \frac{1}{d}\mathcal{O}_K \mathcal{O}_L$ where d is the GCD of D_K and D_L.
Hint: Use the fact (implied by the hypothesis $[KL : \mathbf{Q}] = mn$) that each embedding of K in $\mathbf{C}$ has a unique extension as an embedding of KL which restricts to the identity on L. Then, use the same idea as for (d).

Lemma. *For any positive integer n, consider the field $K = \mathbf{Q}(\zeta)$ where $\zeta = e^{2i\pi/n}$. Then, $\mathcal{O}_K = \mathbf{Z}[\zeta]$.*

Proof. By the exercises (c) and (e) above, and the fact that Euler's phi-function is multiplicative, it suffices to prove the lemma when n is a power of a prime.

Let us use the notation $\mathrm{disc}(\alpha)$ when we talk about $D^K_{\mathbf{Q}}(1, \alpha, \cdots, \alpha^{m-1})$ for some number field $K = \mathbf{Q}(\alpha)$ of degree m. Let $n = p^r$ and ζ be a primitive n-th root of unity. From an earlier exercise, we have $\mathrm{disc}(\zeta) = \mathrm{disc}(1 - \zeta)$. Moreover, $p = \prod_{(k,p)=1}(1 - \zeta^k)$ as seen by evaluating the corresponding cyclotomic polynomial at 1. Here, the product is over k less than $\phi(p^r)$. Evidently, $1 - \zeta^k$ is an associate of $1 - \zeta$ for any k coprime to p. Therefore, p equals $(1 - \zeta)^{\phi(p^r)}$ upto a unit in $\mathbf{Z}[\zeta]$. Now, by an exercise above, every element of $\mathcal{O}_K$ is of the form

$$\sum_{i<\phi(p^r)} m_i (1 - \zeta)^{i-1}/d,$$

where $d = \mathrm{disc}(\zeta)$. Note that d is a power of p. If $\mathcal{O}_K \neq \mathbf{Z}[1 - \zeta]$, then there exists an element $x \in \mathcal{O}_K$ for which not all m_i are divisible by d. If all the

m_i's are divisible by p, we can divide them all by p and proceeding this way we finally arrive at an element in $\mathcal{O}_K$ of the form $x = \sum_{i \geq j} m_i(1-\zeta)^{i-1}/p$ with $j \geq 1$ and m_j not a multiple of p. Now, we noted in the beginning of the proof that p is an associate of $(1-\zeta)^{\phi(p^r)}$ in $\mathbf{Z}[\zeta]$. This means, in particular, that $p/(1-\zeta)^j \in \mathbf{Z}[\zeta] \subset \mathcal{O}_K$. Hence, we have $xp/(1-\zeta)^j \in \mathcal{O}_K$. Hence, we get from the expression for x that $m_j/(1-\zeta) \in \mathcal{O}_K$. So, $N_{\mathbf{Q}}^K(1-\zeta)$ divides $N_{\mathbf{Q}}^K(m_j) = m_j^{\phi(p^r)}$ i.e., $p|m_j$, which is a contradiction. This proves the lemma.

Some definitions. Let A be a DD, K its quotient field and L a finite, separable extension. Let B denote the integral closure of A in L. For any non-zero prime ideal P of A, as B is a DD, one can write $PB = P_1^{e_1} \cdots P_g^{e_g}$ where all $e_i > 0$. The integer e_i is called the *ramification index of* P_i and sometimes denoted by $e(P_i/P)$ to make its dependence clear. P is said to be *unramified in* B if each $e_i = 1$; otherwise it is said to be ramified. P is said to be *totally ramified* if $g = 1$ and $e_1 = [L : K]$. The primes P_i lie over P and these are all the primes lying over P. The degree f_i (denoted by $f(P_i/P)$) of the field extension $B/P_i \supset A/P$ is evidently at the most equal to the degree of L over K. The finite field A/P (*it is indeed finite*) is called the residue field of K at P. The field extension $B/P_i \supset A/P$ is called the residue field extension at P_i and f_i is called the *residue field degree of* P_i *over* P or *the relative degree of* P_i *over* P.

Proposition. *Let A be a DD, K its quotient field and L a finite separable extension. Let B denote the integral closure of A in L. For a non-zero prime ideal P of A, writing $PB = P_1^{e_1} \cdots P_g^{e_g}$ we have $\sum_{i=1}^g e_i f_i$ where $f_i = [B/P_i : A/P]$.*

Proof.
The trick is to localize at P i.e. consider $S^{-1}A$ and $S^{-1}B$ where $S = A \setminus P$. Now $S^{-1}B$ is the integral closure of $S^{-1}A$ in L, and $S^{-1}A/S^{-1}P \cong A/P$. Note also that $PS^{-1}B = Q_1^{e_1} \cdots Q_g^{e_g}$ where $Q_i = P_i S^{-1}B$ and that $S^{-1}B/Q_i \cong B/P_i$ Thus, to prove the proposition we may replace A, B by $S^{-1}A, S^{-1}B$. In this case, A, B are PIDs as they are DDs with only finitely many primes! Therefore, B which is a submodule of a free A-module is, itself, free of rank n (the rank is n as B contains a K-basis of L). Let $v_1, \cdots, v_n$ be an A-basis of B. If $\bar{v_i}$ denotes the image of v_i modulo PB, we have $B/PB = \sum_{i=1}^n (A/P)\bar{v_i}$. Moreover, if $\sum_{i=1}^n \bar{a_i}\bar{v_i} = 0$ in B/PB, then $\sum_{i=1}^n a_i v_i \in PB$. This forces each a_i to be in P since v_i's form a basis of B. Thus, $\bar{v_1}, \cdots, \bar{v_n}$ is a basis of the A/P-vector space B/PB. Thus, $\dim_{A/P} B/PB = n$. Let us count this same dimension in another way. By the Chinese remainder theorem, one has $B/PB = B/\prod P_i^{e_i} \cong \oplus B/P_i^{e_i}$ as rings as well as as A/P-vector spaces. We need to count the dimension of each $B/P_i^{e_i}$. Now, since $P \subset P_i$, we have $PP_i^r \subset P_i^{r+1}$ for all $r \geq 1$. Hence, P_i^r/P_i^{r+1} is an A/P-vector space. Thus, as A/P-vector spaces, we have

$$B/P_i^{e_i} \cong B/P_i \oplus P_i/P_i^2 \oplus \cdots \oplus P_i^{e_i-1}/P_i^{e_i}$$

Further, as B is a PID, one can write $P_i = (\pi_i)$. Then, for each r, the multiplication by π_i^r gives an A/P-isomorphism from B/P_i onto P_i^r/P_i^{r+1}. Hence, we have $\dim_{A/P} B/P_i^{e_i} = e_i f_i$ which gives that $n = \sum e_i f_i$.

Definition. A maximal ideal P of A is said to *split completely in* B if $e_i = 1 = f_i$; so PB is a product of n distinct prime ideals.

Remarks.
(a) The e's and the f's multiply in towers.
(b) Let p be a prime, $\zeta = e^{2i\pi/p}$ and $K = \mathbf{Q}(\zeta)$. Then, p is totally ramified in K.
Idea: Show that $p = \prod_{i=1}^{p-1}(1-\zeta^i)$ and that each $1-\zeta^i$ is a unit times $1-\zeta$.

Proposition.
If, in addition, L/K is a Galois extension, then all the e_i's are equal and all the f_i's are equal. Hence $n = efg$ for some positive integers e, f, g.
Proof.
The Galois group $\mathrm{Gal}(L/K)$ acts transitively on the set $\{P_1, \cdots, P_g\}$. For, if it does not, there exist $i \neq j$ such that $gP_i \neq P_j$ for all $g \in \mathrm{Gal}(L/K)$. Then, choose, by the Chinese remainder theorem, an element $b \in P_j$ such that $b \equiv 1$ mod gP_i for each $g \in G$. But then the element $a = N_K^L(b) = \prod_g g(b)$ is in A on the one hand, and is in P_j on the other. As $A \cap P_j = P$, this means that $\prod_g g(b) \in P \subset P_i$ i.e. some $g(b) \in P_i$, which contradicts the choice of b. Hence, it follows that the Galois group acts transitively. Then, if $gP_i = P_j$, the observation $PB = g(PB)$ along with the uniqueness of decomposition into prime ideals in B yields $e_i = e_j$. Therefore, all the e_i's are equal. Finally, if $g(P_i) = P_j$, then g induces an A/P-isomorphism from B/P_i to B/P_j and so $f_i = f_j$. The corollary is proved.

The *decomposition group* of P_i is the subgroup $D_{P_i} := \{g \in \mathrm{Gal}(L/K) : g(P_i) = P_i\}$. The Galois group induces a natural homomorphism θ_{P_i} from D_{P_i} to $\mathrm{Gal}((B/P_i)/(A/P))$. The kernel T_{P_i} is called the *inertia group of* P_i. If the inertia group T_{P_i} is trivial, one defines the *Frobenius element* Fr_{P_i} at P_i as the inverse image under the isomorphism θ_{P_i} of the Frobenius automorphism $t \mapsto t^{\#(A/P)}$ which generates $\mathrm{Gal}((B/P_i)/(A/P))$.

Remarks.
(a) The above homomorphism from D_{P_i} to $Gal((B/P_i)/(A/P))$ is surjective.
Idea: Use the Chinese remainder theorem.
(b) The D_{P_i}'s are mutually conjugate and $\#D_{P_i} = ef$, $\#T_{P_i} = e$ for all i.

For any algebraic number field K and a non-zero ideal I, the *norm* $N(I)$ *of* I is defined to be the cardinality of the finite ring $\mathcal{O}_K/I$.

Proposition. *Let K be an algebraic number field. Then,*
(a) = if I, J are non-zero ideals, $N(IJ) = N(I)N(J)$.
(b) if P is a maximal ideal, $N(P) = p^f$ where p is the prime number lying below P and $f = f(P/p)$.
(c) if L/K is an extension of degree n, then for any non-zero ideal I of $\mathcal{O}_K$, $N(I\mathcal{O}_K) = N(I)^n$.
(d) if $x \neq 0$ is in $\mathcal{O}_K$, $N((x)) = \mid N_{\mathbf{Q}}^K(x) \mid$.
Proof.
Exercise.

Lemma. *If $p \nmid n$, then p splits into $\phi(n)/f$ prime ideals in $K := \mathbf{Q}(\zeta_n)$ where f is the order of p mod n.*

Proof.

Let $\bar{p}$ denote the image of p in $\mathbf{Z}_n^*$ and, suppose it corresponds to $\sigma \in Gal(K/\mathbf{Q})$. Then, σ has order f. Now, let P be a prime lying over p; then the residue field extension O_K/P of $\mathbf{Z}_p$ is cyclic, of degree $f(P/p)$ which we must show to be equal to f. Let τ be the generator $x \mapsto x^p$ of this Galois extension of residue fields. We shall prove that $\sigma^a = Id$ if and only if $\tau^a = Id$. Now $\sigma^a = Id$ if and only if $p^a \equiv 1 \bmod n$. Now $\tau^a = Id$ if and only if $\zeta^a \equiv \zeta \bmod P$. This is so if and only if $1 - \zeta^{a-1} \in P$. Now, $n = \prod_{i=1}^{n-1}(1 - \zeta^i)$.
Write $p^a \equiv b \bmod n$; then $\zeta^b = \zeta^{p^a} \equiv \zeta \bmod P$; so, $1 - \zeta^{b-1} \in P$. If $b - 1 \neq 0$, then we would have $n \in P$ which is impossible as $p \in P$. Hence $b = 1$.

Remarks.

Let L/K be a Galois extension of number fields with Galois group G. Let $P \subset O_K$ be an unramified prime ideal and let $PO_L = (Q_1 \cdots Q_g)^e$. Now $fg = [L : K] = n$, say. Now, write $Q = Q_1$ and $D = G_Q$; it has order f. We have $G = \sqcup_{i \leq g} g_i D$ for some coset representatives $g_i \in G$. Now

$$\prod_{g \in G} g(Q) = (\prod_{i \leq g} g_i(Q))^{|D|} = (Q_1 \cdots Q_g)^f = (PO_L)^f = P^f O_L$$

One calls the ideal P^f of O_K to be the relative norm of Q over P from L to K. If Q happens to be principal, say $Q = (b)$, then clearly $\prod_{g \in G} g(Q) = N_{L/K}(b)O_L$. Thus, $P^f O_L = N_{L/K}(b)O_L$. Since $N_{L/K}(b) \in O_K$, therefore, $N_{L/K}(b)O_L \cap O_K = N_{L/K}(b)O_K$ as $O_L \cap K = O_K$. Hence P^f is principal.

Why is Fermat's last theorem not trivial to prove?

Let p be an odd prime and $\zeta = e^{2i\pi/p}$. The element $S = \sum_{i=1}^{p-1}(i/p)\zeta^i$ of $K = \mathbf{Q}(\zeta)$ satisfies $S^2 = (-1/p)p$. Hence, every quadratic extension of $\mathbf{Q}$ is contained in a cyclotomic extension.
Now, let $K = \mathbf{Q}(\sqrt{-23}), L = \mathbf{Q}(\zeta)$ where $\zeta = e^{2i\pi/23}$. Then $K \subset L$ by the above remark. Also, $2\mathcal{O}_K = P\bar{P}$ where $P = (2, \frac{1+\sqrt{-23}}{2})$ and $\bar{P} = (2, \frac{1-\sqrt{-23}}{2})$. If a prime Q in L lying over P is principal, then P^f is principal where $f = f(Q/P)$ by what we have proved above. As P is not principal and $P^3 = (\frac{-3+\sqrt{-23}}{2})$, P^f cannot be principal as f divides $[L : K]$. So $\mathcal{O}_L$ is not a PID.

Indeed, it turns out that for *every prime* ≥ 23, the ring of integers of the corresponding cyclotomic field fails to be a PID.

Remarks. When K is the quotient field of a DD A, and L is a finite, separable extension of K and B the integral closure of A in L, the following remarkable theorem of Kummer provides a way to read off the decomposition of a prime ideal in terms of the decomposition of the minimal polynomial of α modulo P. Here $L = K(\alpha)$ and $\alpha \in B$ and the theorem is valid under a mild assumption.

Theorem (Kummer). *Let $A, K, L = K(\alpha), B, P, f$ be as before. Assume,*

in addition, that $B = A[\alpha]$. *Write* $\bar{f} = \bar{p}_1^{e_1} \cdots \bar{p}_g^{e_g}$ *where* $\bar{p}_i$ *are irreducible polynomials in* $(A/P)[X]$ *and* $\bar{f}$ *denotes the image of* f *mod* P. *Then,*

$$PB = P_1^{e_1} \cdots P_g^{e_g}$$

where P_i*'s are prime ideals and* $f(P_i/P) = deg(\bar{p}_i)$. *Indeed,* $P_i = PB + (p_i(\alpha))$ *where* p_i*'s are arbitrary lifts of* $\bar{p}_i$*'s.*

Before proceeding to prove it, let us look at some applications to see really how powerful this is.

Applications of Kummer's theorem

I. Prime decomposition in quadratic fields

As we saw earlier, if $K = \mathbf{Q}(\sqrt{d})$ with d square-free, then $\mathcal{O}_K = \mathbf{Z}[\alpha]$ where $\alpha = \sqrt{d}$ or $\frac{1+\sqrt{d}}{2}$ according as $d \equiv 2, 3 \bmod 4$ or $d \equiv 1 \bmod 4$. The minimal polynomial f is $X^2 - d$ in the first case and $X^2 - X + \frac{1-d}{4}$ in the second. If $d \equiv 2$ or $3 \bmod 4$, $f(X) = X^2 - d$ is a square modulo any prime p dividing d and also modulo 2. Thus, 2 and primes dividing d are (totally) ramified. If an odd prime p does not divide d, then f modulo p is reducible or irreducible according as whether d is a square modulo p or not. Thus, these primes, respectively, split completely and remain inert. Similarly, one can argue for the case $X^2 - X + \frac{1-d}{4}$ corresponding to $d \equiv 1 \bmod 4$.
For any odd prime p, denote by (a/p) the Legendre symbol. To sum up :
(a) if $p|d$, p is (totally) ramified i.e. $p\mathcal{O}_K = P^2$ where the prime ideal $P = (p, \sqrt{d})$,
(b) if p is odd and coprime to d, it is unramified and splits completely or remains a prime according as whether $(d/p) = 1$ or not,
(b)$'$ if $d = q$ is a prime $\equiv 1 \bmod 4$, and p is an odd prime, then $(q/p) = 1 \Leftrightarrow$ the polynomial $X^2 - X + \frac{1-q}{4}$ has a solution mod $p \Leftrightarrow \mathbf{Q}(\sqrt{q})$ is fixed by the Frobenius $Fr_p \Leftrightarrow (p/q) = 1$.
(c) if d is odd, 2 is ramified if $d \equiv 3 \bmod 4$, splits completely if $d \equiv 1 \bmod 8$ and remains a prime if $d \equiv 5 \bmod 8$.
(d) One can prove the whole of quadratic reciprocity law by proving a corresponding version of (b)$'$ for primes $\equiv 3 \bmod 4$.

II. Dedekind's discriminant criterion for ramification

Theorem. *Suppose* $K = \mathbf{Q}(\alpha)$ *is an algebraic number field and assume that* $\mathcal{O}_K = \mathbf{Z}[\alpha]$ *for some* α. *Then, a prime* p *ramifies in* K *if, and only if,* p *divides* $Disc(K)$.
Proof.
Let $f(X) = \prod_i (X - \alpha_i)$ be the minimal polynomial of α. We have seen that $\mathrm{disc}(K) = \mathrm{disc}(f) = \pm\prod_{i \neq j}(\alpha_i - \alpha_j)$. By Kummer's theorem, a prime ramifies in K if, and only if, f has a multiple root modulo p. This is so if, and only if, $\mathrm{disc}(\bar{f}) \equiv 0 \bmod p$ i.e. if, and only if, p divides $\mathrm{disc}(f)$. Here $\bar{f}$ denotes the reduction of f modulo p.

Proof of Kummer's theorem.

Consider the ring homomorphisms

$$A[X] \to (A/P)[X] \to (A/P)[X]/(\bar{p}_i(X))$$

Call the composite map ϕ_i. Note that $(A/P)[X]/(\bar{p}_i(X)) \cong (A/P)[\alpha_i]$ for any root α_i of $\bar{p}_i$. Therefore, $\mathrm{Ker}(\phi_i)$ is a maximal ideal as ϕ_i is evidently surjective. Moreover, it is clear that $P \subset \mathrm{Ker}(\phi_i)$ and $p_i(X) \in \mathrm{Ker}(\phi_i)$ for any arbitrary $p_i \in A[X]$ which maps to $\bar{p}_i$. Further, it is clear from the definition of ϕ_i that $\mathrm{Ker}(\phi_i)$ is the ideal generated by P and p_i in $A[X]$. Now, by the hypothesis, $\bar{f} = \bar{p}_1^{e_1} \cdots \bar{p}_g^{e_g}$ which implies that $f \in (P, p_i) = \mathrm{Ker}(\phi_i)$. Therefore, ϕ_i factors through (f) to give a surjective homomorphism $\theta_i : A[X]/(f) \to (A/P)[X]/(\bar{p}_i(X))$. Note that we have assumed that $B = A[\alpha]$ which gives that $A[X]/(f) \cong B$ where X maps to α. So, we have obtained $\theta_i : B \to (A/P)[X]/(\bar{p}_i(X))$ which is surjective and has kernel $\mathrm{Ker}(\theta_i) = PB + p_i(\alpha)B$. Thus, $P_i :== PB + p_i(\alpha)B = \mathrm{Ker}(\theta_i)$ are maximal ideals in B. As they contain P, they lie over P. Note that $f(P_i/P) = [B/P_i : A/P] = \dim_{A/P}(A/P)[X]/(\bar{p}_i(X)) = \deg \bar{p}_i$. We shall prove now that P_i exhaust all the maximal ideals of B lying over P and have ramification indices equal to e_i.
Note first that the assumption $\bar{f} = \bar{p}_1^{e_1} \cdots \bar{p}_g^{e_g}$ gives, on comparing degrees that $\sum_i e_i f_i = \deg(f) = [L : K]$. The same thing also gives for arbitrary lifts p_i that $f - p_1^{e_1} \cdots p_g^{e_g} \in P[X]$ which, in turn gives, on evaluation at α, that $p_1(\alpha)^{e_1} \cdots p_g(\alpha)^{e_g} \in PA[\alpha] = PB$. So, if Q is any prime ideal of B lying over P, we have $p_1(\alpha)^{e_1} \cdots p_g(\alpha)^{e_g} \in PB \subset Q$. Then, $p_i(\alpha) \in Q$ for some i. But then, $P_i = PB + p_i(\alpha) \subset Q$ and, as both are maximal ideals, they must be equal.
Finally, let $PB = P_1^{d_1} \cdots P_g^{d_g}$. Then,

$$\begin{aligned} P_1^{e_1} \cdots P_g^{e_g} &= (P, p_1(\alpha))^{e_1} \cdots (P, p_g(\alpha))^{e_g} \\ &\subset PB + (p_1(\alpha)^{e_1} \cdots p_g(\alpha)^{e_g}) = PB = P_1^{d_1} \cdots P_g^{d_g}. \end{aligned}$$

Thus, $e_i \geq d_i$. As $\sum e_i f_i = [L : K] = \sum d_i f_i$, this forces $d_i = e_i$. The proof is complete.

The discriminant criterion was generalized by Dedekind to the situation when the integral closure B of A in a finite, separable extension L may not satisfy the condition $B = A[\alpha]$ for any α. The following example shows that the condition $B = A[\alpha]$ may not hold for any α.

Example. *Let K denote the unique subfield K of $L = \mathbf{Q}(\zeta_{31})$ of degree 6 over $\mathbf{Q}$. Then, $\mathcal{O}_K \neq \mathbf{Z}[\alpha]$ for any α.*
Idea:
In general, if E/F is a finite Galois extension, and D is the decomposition group at some prime Q of E, then, $P = Q \cap \mathcal{O}_F$ splits completely in E^D.
Returning to our situation, look at the prime 2 which is unramified. As the order of 2 modulo 31 is 5, 2 splits in $\mathcal{O}_L$ into $\phi(31)/5 = 6$ primes. Therefore, the decomposition group D at any prime of L lying over 2 has order 5. As $\mathrm{Gal}(L/\mathbf{Q})$ is cyclic, it has a unique subgroup of order 5 (indeed, of order any

divisor of 30). Thus the fixed field L^D is of degree 6 over $\mathbf{Q}$ and must be K. By the observation made in the beginning, it follows that 2 splits completely (into 6 primes) in K. Hence, if $\mathcal{O}_K$ were of the form $\mathbf{Z}[\alpha]$, it would follow by Kummer's theorem that the minimal polynomial of α would split modulo $Q \cap \mathbf{Z}$ into six distinct linear factors. However, over $\mathbf{Z}/2$, there are only two linear polynomials! This contradiction establishes the validity of the example.

Example.
For any n, let Φ_n denote the n-th cyclotomic polynomial (i.e. minimal polynomial of $e^{2i\pi/n}$ over $\mathbf{Q}$). Note that $X^n - 1 = \prod_{d|n} \Phi_d(X)$. Let p be a prime not dividing n and $a \in \mathbf{Z}$. Show that p divides $\Phi_n(a)$ if, and only if, a has order n in $(\mathbf{Z}/p)^*$. Moreover, this happens for some p, a if, and only if, $p \equiv 1$ mod n. Hence, there are infinitely many primes $p \equiv 1$ mod n.
For any n, and any prime $p \equiv 1$ mod n, p splits completely in the cyclotomic field $\mathbf{Q}(\zeta_n)$ into the prime ideals $P_i = (p, \zeta_n - i)$, where i has order n in $(\mathbf{Z}/p)^*$.

Remarks.
Let K be a number field, A its ring of integers, and suppose that L is a finite extension of K. Let B denote the ring of integers of L and let $P \subset A$ be a maximal ideal. If $PB = (P_1 \cdots P_g)^e$ in B, then there are fields E, F such that $K \subset F \subset E \subset L$ with $[L : E] = e$, $[E : F] = f$, $[F : K] = g$. Further, such E, F exist with the properties: (i) P splits completely in F into the product of the primes of F lying below $P_1, \cdots, P_g$,
(ii) each prime of F lying above P remains a prime in E,
and (iii) each prime of F lying above P totally ramifies in L.
Idea:
Look at the fixed fields under the decomposition group and the inertia group of any P_i.

Minkowski's bound and Dirichlet's unit theorem.

The class group (that is, the group of fractional ideals) of an algebraic number field is finite. Its order, called the *class number*, gives a measure of the deviation from the unique factorisation property. Although the finiteness is easy to establish, the easy proof gives a somewhat large bound. A much better bound was obtained by Minkowski using a geometric method.

Theorem. *For an algebraic number field K, the class group is finite.*
Proof. Fix an integral basis $\{v_1, \cdots, v_n\}$ of $\mathcal{O}_K$. Let $I \neq 0$ be any ideal and consider the subset S of $\mathcal{O}_K$ consisting of all $\sum_{i=1}^n m_i v_i$ with $0 \leq m_i \leq N(I)^{1/n}$. Evidently, $\#\ S > N(I) = \#\ (\mathcal{O}_K/I)$. Therefore, there exist $a \neq b \in S$ such that $a - b \in I$. Notice that $a - b = \sum_i m_i v_i$ for some integers m_i which satisfy $\mid m_i \mid \leq N(I)^{1/n}$. Let us compute its norm over $\mathbf{Q}$. We have $N_{K/\mathbf{Q}}(a - b) = \prod_i \sigma_i(\sum_j m_j v_j)$ where σ_i's are the embeddings of K in $\mathbf{C}$. Therefore,

$$\mid N_{K/\mathbf{Q}}(a-b) \mid \ = \prod_i \mid \sum_j m_j \sigma_i(v_j) \mid \ \leq \prod_i \sum_j \mid m_j \mid \ \mid \sigma_i(v_j) \mid \ \leq N(I)C,$$

where $C = \prod_i \sum_j \mid \sigma_i(v_j) \mid$ is a constant independent of the ideal I; it depends only on K. Now $a - b \in I \Rightarrow (a - b) = IJ$ for some non-zero ideal J. Thus $N_{K/\mathbf{Q}}(a - b) = N(I)N(J) \leq N(I)C$ and we get $N(J) \leq C$. As J is just the inverse of I in the class group, it runs through the class group when I does. Therefore, we have shown that any element of the class group has a representative ideal whose norm is at the most the constant C. As there are only finitely many ideals with the norm bounded by an absolute constant, the theorem follows.

Example. Let $K = \mathbf{Q}(\sqrt{2})$. Then, $\mathcal{O}_K = \mathbf{Z}[\sqrt{2}]$ has $\{1, \sqrt{2}\}$ as a $\mathbf{Z}$-basis. The constant C above is $C = (1 + \sqrt{2})^2 = 5.8....$ So, every ideal has a representative I with norm at the most 5. Thus, the prime ideals dividing I must have norm ≤ 5 which means that they are among those lying over $2, 3$ and 5. Now, $3, 5$ are unramified and must, therefore, be either inert or split. As 2 is not a square mod 3, 3 remains prime. So is the case with 5 also. Finally, 2 is the square of the prime ideal $(\sqrt{2})$. Thus, we have shown that every ideal class contains a representative ideal which is principal. Thus, the class group is trivial, i.e. $\mathcal{O}_K$ is a PID.

The bound given above is somewhat large. One can do rather better; proceeding as in the proof of the theorem, one can write out the matrix M of $a - b$ with respect to the basis $\{v_1, \cdots, v_n\}$. $M = \sum_i m_i M_i$ where M_i is the matrix of v_i with respect to the same ordered basis. Note that all the entries of M_i are integers whose absolute values are bounded by a constant depending only on the basis $\{v_i\}$ and not on the ideal I. Then, by definition, $\mid N_{K/\mathbf{Q}}(a - b) \mid \; = \; \mid \det(M) \mid \; \leq C_0 N(I)$. This constant C_0 is better than the constant C in the proof of the theorem. For example, when $K = \mathbf{Q}(\sqrt{-5})$, we have $C = 10,\ C_0 = 6$. But, in fact, a method due to Minkowski gives a much better bound. In this example, it will give a constant less than 3 which will enable us to conclude quite easily that the class number is 2.

Definitions. A *lattice* Λ in the Euclidean space $\mathbf{R}^n$ is the $\mathbf{Z}$-span of an $\mathbf{R}$-basis of $\mathbf{R}^n$. Clearly, the group $GL_n(\mathbf{R})$ of invertible $n \times n$ matrices acts transitively on the set of all lattices. Thus, any lattice can be identified with $g\mathbf{Z}^n$ for some $g \in GL_n(\mathbf{R})$. Given a lattice Λ, *a fundamental parallelotope* for it is the set of vectors $\{\sum_i t_i e_i : 0 \leq t_i < 1\}$ for any basis $\{e_i\}$ of Λ. As any two $\mathbf{Z}$-bases are transforms of each other under a matrix in $GL_n(\mathbf{Z}) = \{\gamma \in M_n(\mathbf{Z}) : det(\gamma) = \pm 1\}$, the *volume of the lattice* $\Lambda = g\mathbf{Z}^n$ is the well-defined non-zero real number $\mid det(g) \mid$. We write $\mathrm{Vol}(\mathbf{R}^n/\Lambda)$ for the volume of Λ.

Lemma. *Let K be an algebraic number field. Let $\sigma_1, \cdots, \sigma_r, \tau_1, \cdots, \tau_s, \bar{\tau}_1, \cdots, \bar{\tau}_s$ be the embeddings of K in $\mathbf{C}$. Here, the σ_i's take real values and the τ_j's take nonreal values. Then, the map $\theta : t \mapsto$*

$$(\sigma_1(t), \cdots, \sigma_r(t), Re(\tau_1(t)), \cdots, Re(\tau_s(t)), Im(\tau_1(t)), \cdots, Im(\tau_s(t)))$$

from K to $\mathbf{R}^n$ embeds $\mathcal{O}_K$ as a lattice. Its volume is $\sqrt{\mid \mathrm{disc}(K) \mid}/2^s$. In particular, K embeds densely in $\mathbf{R}^n$.

Proof.
Let $v_1, \cdots, v_n$ be a $\mathbf{Z}$-basis of $\mathcal{O}_K$. We show that $\theta(v_1), \cdots, \theta(v_n)$ are linearly independent. If we write $\theta = (\theta_1, \cdots, \theta_n)$ to mean the obvious, look at the matrix M with $m_{ij} = \theta_i(v_j)$. Elementary column operations transform M to the matrix whose i-th row is

$$(1/2i)^s(\sigma_1(v_i), \cdots, \sigma_r(v_i), \tau_1(v_1), \bar{\tau}_1(v_i), \cdots, \tau_s(v_i), \bar{\tau}_s(v_i))$$

This gives the result that the determinant of M is $(1/2i)^s\sqrt{\mathrm{disc}(K)}$;
so $\mathrm{Vol}(\mathbf{R}^n/\theta(\mathcal{O}_K)) = \sqrt{\mid \mathrm{disc}(K) \mid}/2^s$.

Definition and Remarks. Given a positive integer n and non-negative integers r, s such that $r+2s = n$, define a *norm on* $\mathbf{R}^n$ by $N_{r,s}(x) = x_1 \cdots x_r(x_{r+1}^2 + x_{r+2}^2) \cdots (x_{n-1}^2 + x_n^2)$. Thus, in the situation of a number field K of degree n over $\mathbf{Q}$ and r, s, θ as above, we have $N_{r,s}(\theta(t)) = N_{K/\mathbf{Q}}(t)$ for all $t \in \mathcal{O}_K$.

Theorem (Minkowski). *Every lattice Λ in $\mathbf{R}^n$ contains $x \neq 0$ with $N_{r,s}(x) \leq \frac{n!}{n^n}(\frac{8}{\pi})^s\mathrm{Vol}(\mathbf{R}^n/\Lambda)$.*

For the proof of Minkowski's theorem, one needs the following beautiful lemma on convex bodies which is of independent interest:

Minkowski's lemma. *Let Λ be a lattice in $\mathbf{R}^n$, E a convex, measurable, centrally symmetric subset of $\mathbf{R}^n$ such that $\mathrm{Vol}(E) > 2^n\mathrm{Vol}(\mathbf{R}^n/\Lambda)$. Then, E contains some non-zero point of Λ. Further, if E is also compact, then the strict inequality in the hypothesis can be weakened to $\geq$.*

Proof. Let F be a fundamental parallelotope for Λ. Then, we have $\mathbf{R}^n = \bigsqcup_{x\in\Lambda}(x + F)$. Now, $\frac{1}{2}E = \bigsqcup_{x\in\Lambda}(\frac{1}{2}E \cap (x + F))$. By the hypothesis,

$$\mathrm{Vol}(F) < \frac{\mathrm{Vol}(E)}{2^n} = \mathrm{Vol}(\frac{E}{2}) \quad = \quad \sum_{x\in\Lambda}\mathrm{Vol}(\frac{1}{2}E \cap (x + F)) = \sum_{x\in\Lambda}\mathrm{Vol}((\frac{1}{2}E - x) \cap F)$$

Therefore, as x runs over Λ, the sets $(\frac{1}{2}E - x) \cap F$ are not all disjoint. Thus, we get $x \neq y$ in Λ so that $\frac{1}{2}a - x = f = \frac{1}{2}b - y$ for some $a, b \in E, f \in F$. Clearly, then we get $0 \neq x - y = \frac{1}{2}a + \frac{1}{2}(-b) \in E \cap \Lambda$. This proves the main assertion. For the case when E is also compact, one may consider the sets $(1 + \frac{1}{n})E$ and obtain lattice points $x_n \neq 0$ as above. Evidently, then all the $x_n \in 2E \cap \Lambda$ which is a finite set. Thus, for some n_0, $x_{n_0} \in (1 + \frac{1}{n})E$ for infinitely many n i.e. $x_{n_0} \in \bar{E} = E$. The proof is complete.

Corollary. *Suppose that Ω is a compact, convex, centrally symmetric subset of $\mathbf{R}^n$ such that $Vol(\Omega) > 0$ and such that $\mid N_{r,s}(a) \mid \leq 1 \ \ \forall\ a \in \Omega$. Then, every lattice Λ contains a non-zero vector x with*

$$\mid N_{r,s}(x) \mid \leq 2^n\frac{Vol(\mathbf{R}^n/\Lambda)}{Vol(\Omega)}.$$

The proof is immediate from Minkowski's lemma applied to the set $E = t\Omega$ where $t^n = 2^n\frac{Vol(\mathbf{R}^n/\Lambda)}{Vol(\Omega)}$.

Proof of Minkowski's theorem. Let Ω be the subset of $\mathbf{R}^n$ defined by the inequality $\sum_{i=1}^{r} \mid x_i \mid +2\sqrt{(x_{r+1}^2 + x_{r+2}^2)} + \cdots + 2\sqrt{(x_{n-1}^2 + x_n^2)} \leq n$. We shall prove that Ω is convex, and that $\mid N_{r,s}(a) \mid \leq 1 \; \forall \; a \in \Omega$. Then, we shall compute its volume and apply the above corollary.
Step I: Ω is convex
¿From the definition of Ω, it is easy to see that if midpoints of any two points of Ω are in Ω, then Ω is convex. Let $(x_1, \cdots, x_n), (y_1, \cdots, y_n) \in \Omega$. Then, we have

$$\sum_{i=1}^{r} \mid x_i \mid +2\sqrt{(x_{r+1}^2 + x_{r+2}^2)} + \cdots + 2\sqrt{(x_{n-1}^2 + x_n^2)} \leq n,$$

$$\sum_{i=1}^{r} \mid y_i \mid +2\sqrt{(y_{r+1}^2 + y_{r+2}^2)} + \cdots + 2\sqrt{(y_{n-1}^2 + y_n^2)} \leq n.$$

Adding and using the triangle inequality

$$\sqrt{(a^2 + b^2)} + \sqrt{(c^2 + d^2)} \geq \sqrt{((a+c)^2 + (b+d)^2)}$$

one concludes that $(\frac{x_1+y_1}{2}, \cdots, \frac{x_n+y_n}{2}) \in \Omega$.
Step II: $\mid N_{r,s}(a) \mid \leq 1 \; \forall \; a$.
This is clear from the usual inequality $A.M \geq G.M$.
Step III: $Vol(\Omega) = \frac{(2n)^n}{n!}(\frac{\pi}{8})^s$.
Let $V_{r,s}(t)$ denote the volume of the set Ω_t defined in a similar fashion to Ω but with n replaced by the real number $t > 0$. It is easy to see from the definition that $V_{r,s}(t) = V_{r,s}(1)t^{r+2s}$. Now, if $r > 0$, then

$$\begin{aligned} V_{r,s}(1) &= 2\int_0^1 V_{r-1,s}(1-x)dx \\ &= 2V_{r-1,s}(1)\int_0^1 (1-x)^{r-1+2s}dx = \frac{2}{r+2s}V_{r-1,s}(1). \end{aligned}$$

Proceeding inductively, one obtains finally that $V_{r,s}(1) = \frac{2^r}{(r+2s)\cdots(2s+1)}V_{o,s}(1)$. Similarly, if $s > 0$, then

$$\begin{aligned} V_{0,s}(1) &= \int\int_{x^2+y^2 \leq 1/4} V_{0,s-1}\left(1 - 2\sqrt{(x^2+y^2)}\right)dxdy \\ &= \int_0^{2\pi}\int_0^{1/2} V_{0,s-1}(1-2\rho)\rho d\rho d\theta. \end{aligned}$$

Once again, iterating inductively, one finally obtains $V_{0,s}(1) = (\frac{\pi}{2})^s \frac{1}{(2s)!}$. Then, $Vol(\Omega_t) = t^n V_{r,s}(1) = t^n 2^{r-s}\pi^s \frac{1}{n!}$ which gives that $Vol(\Omega = \Omega_n) = n^n \frac{2^n}{2^{3s}}\pi^s \frac{1}{n!} = \frac{(2n)^n}{n!}(\frac{\pi}{8})^s$. The proof of Step III and, along with it, that of Minkowski's theorem, is complete.

Corollary. *Let $[K : \mathbf{Q}] = n$ and r, s have the usual meaning. Then,*
(a) Every non-zero ideal I contains $x \neq 0$ with

$$\mid N(x) \mid \leq \frac{n!}{n^n}\left(\frac{4}{\pi}\right)^s \sqrt{\mid \mathrm{disc}(K) \mid}\; N(I).$$

(b) Every ideal class contains an ideal I with

$$| N(I) | \leq \frac{n!}{n^n} \left(\frac{4}{\pi}\right)^s \sqrt{| \operatorname{disc}(K) |}.$$

(c) $\operatorname{disc}(K) > 1$ *if* $K \neq \mathbf{Q}$.
(d) If $K \neq \mathbf{Q}$, *then some prime number* p *ramifies in* K.
Proof.
Using the lemma above, $\mathcal{O}_K$ can be viewed as a lattice in $\mathbf{R}^n$ whose volume has also been computed. Therefore, both (a) and (b) are direct consequences of Minkowski's theorem. To prove (c), just observe that the number $\frac{n^n}{n!}(\frac{\pi}{4})^s > 2^{n-1}(\frac{\pi}{4})^n = \frac{1}{2}(\pi/2)^n > 1$ for $n > 1$. Finally, (d) follows from Dedekind's theorem which showed that prime numbers which divide the discriminant of K must ramify in K.

Examples. (I) Let $K = \mathbf{Q}(\sqrt{-11})$.
The discriminant is -11 and the Minkowski bound is $\frac{2}{\pi}\sqrt{11} = 2.11....$
Look at ideals of norm 2. As 2 remains prime, there are no ideals of norm 2. Hence the class number is 1.
(II) Let $K = \mathbf{Q}(\sqrt{-5})$.
Then, the Minkowski bound shows that each ideal class contains a representative ideal I of norm $N(I) \leq \frac{4\sqrt{5}}{\pi} < 3$. So, one need only consider the ideals lying above 2 viz., $(2, 1 \pm \sqrt{-5})$. It is easy to see that these are not principal and thus it follows that K has class number 2.

Application:
Let us use the fact that $K = \mathbf{Q}(\sqrt{-5})$ has class number 2 to show that the equation $y^2 = x^3 - 5$ has no integral solutions.
Reading a solution mod 4 tells us that x must be odd. Also, if x, y had a common prime factor p, then $p = 5$. But then the powers of 5 dividing y^2 and $x^3 - 5$ are unequal. Thus, $(x, y) = 1$. Write $(y + \sqrt{-5})(y - \sqrt{-5}) = x^3$, and look at a prime ideal P dividing both the principal ideals $(y + \sqrt{-5})$ and $(y - \sqrt{-5})$ in $\mathbf{Z}[\sqrt{-5}]$. Then P divides $(2y)$. As P divides (x^3) and x is odd, P does not divide (2) (compare norms again). Thus, P divides (y). But then P divides (y) as well as $(x^3) = (x)^3$ which means P divides (x), a contradiction. Thus, the ideals $(y + \sqrt{-5})$ and $(y - \sqrt{-5})$ are coprime. Since their product is a cube, both the elements $y + \sqrt{-5}$ and $y - \sqrt{-5}$) are cubes in $\mathbf{Z}[\sqrt{-5}]$ since the only units are 1 and -1 both of which are cubes.
Writing $y + \sqrt{-5} = (a + b\sqrt{-5})^3$ in $\mathbf{Z}[\sqrt{-5}]$, we have

$$1 = b(3a^2 - 5b^2).$$

This is impossible in integers a, b. Thus, the equation has no integral solutions.

Dirichlet's unit theorem

Now, we use Minkowski's method to find the structure of the units in any algebraic number field K. Recall that we embedded $\mathcal{O}_K$ as a lattice Λ_0 in $\mathbf{R}^n$

by means of $\theta : a \mapsto (\sigma_1(a), \cdots, \sigma_r(a), Re\tau_1(a), Im\tau_1(a), \cdots, Re\tau_s(a), Im\tau_s(a))$. Here $n = [K : \mathbf{Q}]$ and $\sigma_1, \cdots, \sigma_r, \tau_1, \bar{\tau}_1, \cdots, \tau_s, \bar{\tau}_s$ are the distinct embeddings of K in $\mathbf{C}$. Clearly, if a is a unit in $\mathcal{O}_K$, then both u and u^{-1} map to vectors which are linearly dependent. Thus, one needs to go to a subspace of $\mathbf{R}^n$ to be sensitive to the units.

Lemma. *Consider the composite map L in*

$$\mathcal{O}_K^* \subset \mathcal{O}_K \setminus 0 \xrightarrow{\theta} \Lambda_0 \setminus 0 \to \mathbf{R}^{r+s}$$

where the last map is $(x_1, \cdots, x_n) \mapsto$
$(log(\mid x_1 \mid), \cdots, log(\mid x_r \mid), log(x_{r+1}^2 + x_{r+2}^2), \cdots, log(x_{n-1}^2 + x_n^2)\ )$. *Then,*
(i) the image of $L : \mathcal{O}_K^* \to \mathbf{R}^{r+s}$ is contained in the hyperplane H of vectors $(x_1, \cdots, x_{r+s})$ such that $\sum_{i=1}^{r+s} x_i = 0$.
(ii) L is a homomorphism.
(iii) $Im(L) \cong \mathbf{Z}^d$ for some $d \leq r + s - 1$.
(iv) $\mathrm{Ker}(L) \cong \mu(K)$*, the group of roots of unity in K and $\mathcal{O}_K^* \cong \mu(K) \times \mathbf{Z}^d$ for some $d \leq r + s - 1$.*

Proof. (i) follows since units must have norm ± 1. (ii) is obvious. To see that (iii) holds, let R be any bounded region in $H \subset \mathbf{R}^{r+s}$ and let $L(u) \in R$. Then, all the conjugates of u have absolute values bounded by a constant depending on R. As the coefficients of the minimal polynomial of u are symmetric functions of the various conjugates of u, this means that there are only finitely many polynomials satisfied by units whose images under L lie in the bounded region R. In other words, $R \cap \mathrm{Im}(L)$ is finite i.e. $\mathrm{Im}(L)$ is discrete in H. Now, (iii) follows by the easy exercise below. The first assertion of (iv) is trivial and the second one follows because one can check easily that units $u_1, \cdots, u_d$ mapping under L to a basis of $\mathrm{Im}(L)$ have to generate a free abelian group.

Exercise. *Show by induction on n that a discrete subgroup of $\mathbf{R}^m$ is isomorphic to $\mathbf{Z}^d$ for some $d \leq m$.*

Dirichlet's unit theorem. *$\mathcal{O}_K^* = \mu(K) \times V$ where $V \cong \mathbf{Z}^{r+s-1}$.*

In other words, the image of $\mathcal{O}_K^*$ under L is actually a lattice in H. This will be seen by actually showing the existence of $r + s - 1$ units whose images under L are linearly independent.

Lemma. *Fix any $k \leq r + s$. Then, $\forall\ \alpha \neq 0$ in $\mathcal{O}_K$, there exists $\beta \in \mathcal{O}_K$ with $\mid N(\beta) \mid \leq (\frac{2}{\pi})^s \sqrt{\mid \mathrm{disc}(K) \mid}$ and satisfies $\beta_i < \alpha_i\ \forall\ i \neq k$. Here α_i, β_i denote the co-ordinates of their images under L.*

Proof. Let c_i be constants such that $0 < c_i < e^{\alpha_i}\ \forall\ i \neq k$ and $c_k = (\frac{2}{\pi})^s \sqrt{\mid \mathrm{disc}(K) \mid} / \prod_{i \neq k} c_i$. Consider the set $\Omega \subset \mathbf{R}^n$ defined by

$$\mid x_i \mid \leq c_i,\ \forall\ i \leq r \quad , \quad x_{r+1}^2 + x_{r+2}^2 \leq c_{r+1}, \cdots, x_{n-1}^2 + x_n^2 \leq c_{r+s}.$$

Then,

$$\mathrm{Vol}(\Omega) = (2c_1) \cdots (2c_r)(\pi c_{r+1}) \cdots (\pi c_{r+s}) = 2^n \mathrm{Vol}(\mathbf{R}^n / \Lambda_0).$$

Applying Minkowski's lemma, one gets some $t \neq 0$ in $\Omega \cap \Lambda_0$. Then, choose $\beta \in \mathcal{O}_K$ corresponding to t.

Lemma. *Fix any $k \leq r+s$. Then, $\exists\ u \in \mathcal{O}_K^*$ such that $L(u) = (u_1, \cdots, u_{r+s})$ satisfies $u_i < 0\ \forall\ i \neq k$.*

Proof. Start with any $\alpha_1 \neq 0$ in $\mathcal{O}_K$ and apply the previous lemma to get some β as above; call that α_2. Repetitively, one gets a sequence $\{\alpha_n\}$ in $\mathcal{O}_K$ such that for all $i \neq k$, the i-th co-ordinate of $L(\alpha_{n+1})$ is less than that of $L(\alpha_n)$. By the lemma, $\mid N(\alpha_n) \mid$ are bounded above as $n \to \infty$. Therefore, the principal ideals (α_n) are only finitely many. Taking any $n < m$ so that $(\alpha_n) = (\alpha_m)$, we have $\alpha_m = \alpha_n u$ for some unit u. Evidently, u does the job.

The proof of Dirichlet's unit theorem is completed as follows. Observe that the units $u_k, k \leq r+s$, obtained by the previous lemma have the property that the $(r+s) \times (r+s)$ matrix $A = (a_{ij})$ whose k-th row is $L(u_k)$ satisfies $a_{ij} < 0$ for all $i \neq j$ and each row sums to 0. It is an easy elementary exercise to see that the rank of A must be $r+s-1$.

Remark
Indeed, Dirichlet's theorem has a more general version. If S is any finite set of places (equivalence class of valuations) of an algebraic number field which contains all the archimedean places, then the group of units of the ring $\mathcal{O}_S := \{a \in K : v(a) \geq 0 \forall\ v \notin S\}$ is finitely generated and has rank $|S| - 1$.

If K is a real quadratic field or a cubic field with a unique real embedding, the unit group has rank 1. Then, the unique unit > 1 of infinite order is called *the fundamental unit.*
If $K = \mathbf{Q}(\sqrt{d})$ is real, quadratic, then the fundamental unit is obtained as follows:
For $d \equiv 2$ or $3 \bmod 4$, look at the smallest positive integer b such that $db^2 \pm 1$ is a square (say, a^2 with $a > 0$); then $a + b\sqrt{d}$ is the fundamental unit (because if $a + b\sqrt{d} = (u + v\sqrt{d})^k$, then $0 < u \leq a, 0 < v \leq b$ and $dv^2 \pm 1$ is a square u^2).
For $d \equiv 1 \bmod 4$, consider the smallest positive integer b such that $db^2 \pm 4$ is a square (say, a^2 with $a > 0$); then $(a + b\sqrt{d})/2$ is the fundamental unit (because if $(a + b\sqrt{d})/2 = \left((u + v\sqrt{d})/2\right)^k$, then $0 < u \leq a, 0 < v \leq b$ and $dv^2 \pm 4$ is a square u^2).

Application to sums of 4 squares
As an application of Minkowski's lemma, let us prove that every positive integer is a sum of 4 squares of integers. It suffices to prove this for any odd prime p. Again, considering the sets $\{a^2\}$ of squares mod p and $\{-b^2 - 1\}$ mod p (each of which is $(p+1)/2$ in number), the sets must intersect. Hence, there exist integers a, b such that $a^2 + b^2 + 1 \equiv 0 \bmod p$.

Consider the lattice $\Gamma \subset \mathbf{R}^4$ generated by the vectors

$$(p,0,0,0),(0,p,0,0),(a,b,1,0),(b,-a,0,1).$$

Clearly,

$$\text{Vol}(\mathbf{R}^4/\Gamma) = det \begin{pmatrix} p & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ a & b & 1 & 0 \\ b & -a & 0 & 1 \end{pmatrix} = p^2.$$

Now, the open ball B of radius $\sqrt{2p}$ has volume $2p^2\pi^2$ (note that volume of the unit ball in $\mathbf{R}^n$ is $\pi^{n/2}/\Gamma(1+\frac{n}{2})$).
Thus Vol $(B) > 2^4p^2 = 2^4$ Vol $(\mathbf{R}^4/\Gamma)$.
By Minkowski's lemma, there exists $0 \neq \gamma \in \Gamma \cap B$.
As $||\gamma||^2 < 2p$ whereas $||x||^2 \equiv 0 \bmod p$ for all $x \in \Gamma$, we have $||\gamma||^2 = p$.

The Brahmagupta equation and continued fractions

For a natural number d, the Brahmagupta equations $x^2 - dy^2 = \pm 1$ (also called Pell's equation erroneously - considering that Pell came 1000 years later and had nothing to do with this equation!) are solved by finding the fundamental unit of $\mathbf{Q}(\sqrt{d})$. The latter problem has been successfully tackled by using the notion of continued fractions. First, we recall that a continued fraction is an expression

$$a_0 + \frac{1}{a_1+}\frac{1}{a_2+}\cdots\cdots\cdots$$

which is to be interpreted as the limit of the rational numbers

$$a_0 + \frac{1}{a_1+}\frac{1}{a_2+}\cdots\frac{1}{a_n}$$

as $n \to \infty$. The rational number

$$a_0 + \frac{1}{a_1+}\frac{1}{a_2+}\cdots\frac{1}{a_n}$$

is called the n-th convergent of the original continued fraction - the latter is generally denoted as $[a_0, a_1, \cdots]$ and the n-th convergent C_n is denoted by $[a_0, a_1, \cdots, a_n]$. It is easy to see that when a_i are positive integers, the sequence $\{C_n\}$ of n-th convergents does converge. It is also easy to see using the greedy algorithm that every positive irrational (real) number has a unique continued fraction expansion with all a_i's positive integers. The crucial property relevant to us is that a continued fraction of positive integers is eventually perdiodic if and only if, it converges to an element in a real quadratic field. We will talk henceforth only of continued fractions with a_i's positive integers. The following properties are easy exercises :
(i) Writing p_n and q_n for the numerator and denominator of the n-th convergent

C_n, one has $p_n q_{n-1} - q_n p_{n-1} = (-1)^{n-1}$.
(ii) The integers p_n and q_n (coprime by (i)) satisfy:

$$\begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix}$$

The basic result we need is :

Let d be a square-free positive integer. Let $\sqrt{d} = [a_0, a_1, \cdots, \cdots]$ be the C.F. and suppose it has period n (that is, for some $N \geq 0$, $a_k = a_{k+n}$ for all $k \geq N$ and n is the least such). Then,
all integer solutions of $x^2 - dy^2 = \pm 1$ are given by $x + y\sqrt{d} = \pm(p_{n-1} + q_{n-1}\sqrt{d})^l$ for some integer l. Further, if $d \not\equiv 1 \ mod\ 4$, then $p_{n-1} + q_{n-1}\sqrt{d}$ is the fundamental unit of $\mathbf{Q}(\sqrt{d})$.

What is to follow - a peek !

Now, we give a sketch of what is to come later. It can be read after this school.

Abelian class field theory

This is a powerful theory which describes the abelian extensions of an algebraic number field K in terms of invariants of K itself. We do not discuss it here but describe some consequences of it. Let us begin with the notion of a reciprocity law of which the quadratic reciprocity law is an example.
We discussed Kummer's theorem on the decomposition of prime ideals and saw that the quadratic reciprocity law could be proved using the ramification of a prime in quadratic extensions. For instance, the set of odd primes modulo which the polynomial $X^2 + 1$ has roots, consists precisely of all primes in the arithmetic progression $4n + 1$. This is equivalent to

$$p \equiv 1 \ mod\ 4 \Leftrightarrow (\frac{-1}{p}) = 1.$$

Further, we can prove statements like :

$$p \equiv 1 \ or\ 3 \ mod\ 8 \Leftrightarrow (\frac{-2}{p}) = 1.$$

$$p \equiv \pm 1 \ or\ \pm 3 \ or\ \pm 9 \ mod\ 28 \Leftrightarrow (\frac{7}{p}) = 1.$$

Thus, we have a nice criterion to decide when a prime splits completely in a quadratic extension. The criterion is in terms of some congruences. One of the principal aims of ramification theory is to give a 'nice' criterion for a prime to split completely in a given extension; one calls such a criterion to be a reciprocity law. The reason that one is interested in a criterion to decide which primes split

completely is that given K, the set of primes of K which split in L determine L uniquely. The last fact mentioned is deep in general and the proof requires class field theory. We describe one particular case.

A Cyclotomic reciprocity law. *Let n be a positive integer and p be a prime not dividing n. Denote by ζ a primitive n-th root of unity. Then, p is unramified in $K = \mathbf{Q}(\zeta)$ and splits into $\phi(n)/f$ primes where f is the order of p in the unit group of $\mathbf{Z}/n$ and ϕ is Euler's phi function. In particular, p splits completely in K if, and only if, $p \equiv 1$ mod n.*

Primes expressible as $a^2 + 27b^2$
Like the quadratic reciprocity law, one has higher power reciprocity laws also. For instance, from the cubic reciprocity law, we can show that 2 is a perfect cube modulo a prime $p \equiv 1 \bmod 3$ if, and only if, $p = a^2 + 27b^2$ for some integers a, b. Using this, it follows for instance that the equation $x^3 - 2y^3 = 23z^n$ has no solutions in coprime integers x, y, z for any n.

Let f be a monic integral polynomial of degree n. Suppose that f has distinct roots $\alpha_1, \cdots, \alpha_n \in \mathbf{C}$; equivalently, the discriminant $disc(f) \neq 0$. Let $K = \mathbf{Q}(\alpha_1, \cdots, \alpha_n)$, the subfield of $\mathbf{C}$ generated by the roots. We look at the Galois group of f, denoted by $\mathrm{Gal}(f)$. For instance, if $f(X) = X^2 - a$ for some nonsquare integer a, then $K = \mathbf{Q}(\sqrt{a})$ where $\sqrt{a}$ denotes a square root of a in $\mathbf{C}$ and G has two elements I, σ where σ interchanges $\sqrt{a}$ and $-\sqrt{a}$. In general, although G is a subgroup of S_n, the permutations which belong to G are rather restricted; for example if f is irreducible over $\mathbf{Q}$, then a permutation in G is necessarily transitive on the α_i's. If $p \not| disc(f)$, then the decomposition type of f modulo p gives a partition of n. On the other hand, each element of G has a cycle decomposition as an element of S_n and, thus defines a partition of n as well. Frobenius's wonderful idea is to relate the numbers of such partitions for a particular type. This will be expressed in terms of a notion of density of a set of prime numbers.

A set S of primes is said to have *density* δ if $\dfrac{\sum_{p \in S} 1/p}{\sum_{all\ p} 1/p} \to \delta$ as $s \to 1^+$. Here 1^+ means the limit when s tends to 1 from the right. For instance, any finite set of primes has density 0. Using this notion of density, we have the :

Frobenius density theorem
The set of primes p modulo which a monic integral, irreducible polynomial f has a given decomposition type $n_1, n_2, \cdots, n_r$, has density equal to $N/O(Gal(f))$ where $N = |\{\sigma \in Gal(f) :\ \sigma$ has a cycle pattern $n_1, n_2, \cdots, n_r\}|$.

An application.
Let us show using this that *if f is an irreducible integral polynomial which does not have roots modulo infinitely many primes, then it is linear.*
Suppose not; then the theorem shows that each σ has a cycle pattern of the form $1, n_2, \cdots$ This means that each element of $\mathrm{Gal}(f)$ fixes a root. Since the roots

of f are transitively moved around by $\mathrm{Gal}(f)$, this group would be the union of the conjugates of its subgroup H consisting of elements which fix a root of f, say α_1. However, it is an elementary exercise that a finite group cannot be the union of conjugates of a proper subgroup. Thus, in our case $H = \mathrm{Gal}(f)$. This means that $\mathrm{Gal}(f)$ fixes each α_i and is therefore trivial. That is, f is linear.

There is a stronger result due to Chebotarev. To state Chebotarev's theorem, we use the Frobenius map. For any prime number p, the p-th power map $Frob_p$ is an automorphism of the field $\bar{\mathbf{F}}_p$ which is identity on $\mathbf{F}_p$. Therefore, $Frob_p$ permutes the roots of any polynomial over $\mathbf{F}_p$. Indeed, *the Galois theory of finite fields amounts to the statement that if g is a polynomial over $\mathbf{F}_p$ with simple roots, then the cycle pattern of $Frob_p$ viewed as a permutation of the roots of g coincides with the decompoasition type of g over $\mathbf{F}_p$*. In our case, we start with an integral polynomial f and look at it modulo p for various primes p. The above basic theory of algebraic numbers shows that whenever $p \not| disc(f)$, the automorphism $Frob_p$ gives rise to a *conjugacy class* in $\mathrm{Gal}(f)$, called the Frobenius conjugacy class modulo p.
In Frobenius's density theorem, one cannot distinguish between two primes p, q defining different conjugacy classes $C(x)$ and $C(y)$ but some powers of x and y are conjugate. For instance, for the polynomial $X^{10} - 1$, the decomposition type modulo primes congruent to $1, 3, 7, 9$ mod 10 are, respectively, $1,1,1,1,1,1,1,1,1,1$; $1,1,4,4$; $1,1,4,4$; $1,1,2,2,2,2$.
Frobenius's theorem cannot distinguish between primes which are 3 mod 10 and those which are 7 mod 10; they define different conjugacy classes in $\mathrm{Gal}(X^{10}-1)$. Thus, it would imply that the number of primes $\equiv 3$ or 7 mod 10 is infinite but doesn't say whether each congruence class contains infinitely many primes. This is what Chebotarev's theorem asserts.

Chebotarev's density theorem.
Let f be monic integral and assume that disc (f) does not vanish. Let C be a conjugacy class of Gal(f). Then, the set of primes p not dividing disc (f) for which $\sigma_p \in C$, has a well-defined density which equals $\frac{|C|}{|G|}$.

Illustrations of Chebotarev's theorem

Illustration I.

Look at the imaginary quadratic field $\mathbf{Q}(i)$ where i is a square root of -1. Therefore, for every prime p, the decomposition group D_p is either trivial or the full Galois group according as $Frob_p$ is trivial or not. Now, for a prime $p \neq 2$, $\mathbf{Q}(i) \otimes_{\mathbf{Q}} \mathbf{Q}_p$ is a field precisely when -1 is not a square in $\mathbf{Q}_p$. Moreover, in this case, D_p can be identified with the Galois group $\mathrm{Gal}(\mathbf{Q}(i)/\mathbf{Q}) \cong \mathbf{Z}/2$. Hence, $Frob_p$ is trvial or not according as -1 is a square mod p or not. The latter is equivalent respectively to whether $p \equiv 1$ mod 4 or $p \equiv 3$ mod 4. Therefore, the density theorem asserts in this case that *there are as many odd primes of the form $4k+1$ as are of the form $4k+3$.*

Illustration II.

Look at the quadratic extension $\mathbf{Q}(\sqrt{p})$ of $\mathbf{Q}$ for an odd prime p. Now, Gal $(\mathbf{Q}(\sqrt{p})/\mathbf{Q}) = \{1, \sigma\} \cong \mathbf{Z}/2$ where $\sigma : \sqrt{p} \mapsto -\sqrt{p}$. An odd prime $l \neq 2$ either splits completely or remains a prime accordingly as to whether p is a square mod l or not. Once again, therefore, for every prime $l \neq p$, the decomposition group D_l is either trivial or the full Galois group according as $Frob_l$ is trivial or not.
If l is such that p is a square mod l, then $\mathbf{Q}(\sqrt{p}) \otimes_{\mathbf{Q}} \mathbf{Q}_l$ is not a field; so the decomposition group D_l is trivial.
If, on the other hand, p is *not* a square mod l, then $\mathbf{Q}(\sqrt{p}) \otimes_{\mathbf{Q}} \mathbf{Q}_l$ is a quadratic extension of $\mathbf{Q}_l$ and the corresponding Galois group can be naturally identified with $D_l = \{1, Frob_l\}$.
In other words, $Frob_l$ is trivial or not, according as p is, or is not, a square mod l. The statement of Chebotarev's theorem in this case is simply that *a given prime p is a square modulo exactly half the proportion of primes.*

Illustration III.

For $n \in \mathbf{N}$ and a primitive n-th root of unity ζ_n, consider the cyclotomic field extension $\mathbf{Q}(\zeta_n)$ of $\mathbf{Q}$. Chebotarev's theorem, in this case, asserts that for each a coprime to n, primes in the congruence class a mod n have density $\frac{1}{\phi(n)}$ where $\phi(n)$ is the order of $\mathrm{Gal}(\mathbf{Q}(\zeta_n)/\mathbf{Q})$. It is not very difficult to show that the last statement implies Dirichlet's theorem on prime numbers in arithmetic progressions.

We state here without proof two results which can be proved with the aid of Chebotarev's density theorem. These concrete applications are :
(I) The set of primes which are expressible in the form $3x^2+xy+4y^2$ for integers x, y, has density $1/5$.
(II) The set of primes p for which the decimal expansion of $1/p$ has odd period, has density $1/3$.