

Math 129: Algebraic Number Theory

Lecture 8

William Stein

Tuesday, March 2, 2004

1 The Discriminant

Suppose w_1, \dots, w_n are a basis for a number field K , which we view as a \mathbf{Q} -vector space. Let $\sigma : K \hookrightarrow \mathbf{C}^n$ be the embedding $\sigma(a) = (\sigma_1(a), \dots, \sigma_n(a))$, where $\sigma_1, \dots, \sigma_n$ are the distinct embeddings of K into \mathbf{C} . Let A be the matrix whose rows are $\sigma(w_1), \dots, \sigma(w_n)$. The quantity $\text{Det}(A)$ depends on the ordering of the w_i , and need not be an integer.

If we consider $\text{Det}(A)^2$ instead, we obtain a number that is a well-defined integer which can be either positive or negative. Note that

$$\text{Det}(A)^2 = \text{Det}(AA) = \text{Det}(AA^t) = \text{Det}\left(\sum_{k=1, \dots, n} \sigma_k(w_i)\sigma_k(w_j)\right) = \text{Det}(\text{Tr}(w_i w_j)_{1 \leq i, j \leq n}),$$

so $\text{Det}(A)^2$ can be defined purely in terms of the trace without mentioning the embeddings σ_i . Also, changing the basis for \mathcal{O}_K is the same as left multiplying A by an integer matrix U of determinant ± 1 , which does not change the squared determinant, since $\text{Det}(UA)^2 = \text{Det}(U)^2 \text{Det}(A)^2 = \text{Det}(A)^2$. Thus $\text{Det}(A)^2$ is well defined, and does not depend on the choice of basis.

If we view K as a \mathbf{Q} -vector space, then $(x, y) \mapsto \text{Tr}(xy)$ defines a bilinear pairing $K \times K \rightarrow \mathbf{Q}$ on K , which we call the *trace pairing*. The following lemma asserts that this pairing is nondegenerate, so $\text{Det}(\text{Tr}(w_i w_j)) \neq 0$ hence $\text{Det}(A) \neq 0$.

Lemma 1.1. *The trace pairing is nondegenerate.*

Proof. If the trace pairing is degenerate, then there exists $a \in K$ such that for every $b \in K$ we have $\text{Tr}(ab) = 0$. In particular, taking $b = a^{-1}$ we see that $0 = \text{Tr}(aa^{-1}) = \text{Tr}(1) = [K : \mathbf{Q}] > 0$, which is absurd. \square

Definition 1.2 (Discriminant). Suppose a_1, \dots, a_n is any \mathbf{Q} -basis of K . The *discriminant* of a_1, \dots, a_n is

$$\text{Disc}(a_1, \dots, a_n) = \text{Det}(\text{Tr}(a_i a_j)_{1 \leq i, j \leq n}) \in \mathbf{Q}.$$

The *discriminant* $\text{Disc}(\mathcal{O})$ of an order \mathcal{O} in \mathcal{O}_K is the discriminant of any basis for \mathcal{O} . The *discriminant* $d_K = \text{Disc}(K)$ of the number field K is the discriminant of \mathcal{O}_K .

Note that the discriminants defined above are all nonzero by Lemma 1.1.

Warning: In MAGMA $\text{Disc}(K)$ is defined to be the discriminant of the polynomial you happened to use to define K , which is (in my opinion) a poor choice and goes against most of the literature.

The following proposition asserts that the discriminant of an order \mathcal{O} in \mathcal{O}_K is bigger than $\text{disc}(\mathcal{O}_K)$ by a factor of the square of the index.

Proposition 1.3. *Suppose \mathcal{O} is an order in \mathcal{O}_K . Then*

$$\text{Disc}(\mathcal{O}) = \text{Disc}(\mathcal{O}_K) \cdot [\mathcal{O}_K : \mathcal{O}]^2.$$

Proof. Let A be a matrix whose rows are the images via σ of a basis for \mathcal{O}_K , and let B be a matrix whose rows are the images via σ of a basis for \mathcal{O} . Since $\mathcal{O} \subset \mathcal{O}_K$ has finite index, there is an integer matrix C such that $CA = B$, and $|\text{Det}(C)| = [\mathcal{O}_K : \mathcal{O}]$. Then

$$\text{Disc}(\mathcal{O}) = \text{Det}(B)^2 = \text{Det}(CA)^2 = \text{Det}(C)^2 \text{Det}(A)^2 = [\mathcal{O}_K : \mathcal{O}]^2 \cdot \text{Disc}(\mathcal{O}_K).$$

□

This result is enough to give an algorithm for computing \mathcal{O}_K , albeit a potentially slow one. Given K , find some order $\mathcal{O} \subset K$, and compute $d = \text{Disc}(\mathcal{O})$. Factor d , and use the factorization to write $d = s \cdot f^2$, where f^2 is the largest square that divides d . Then the index of \mathcal{O} in \mathcal{O}_K is a divisor of f , and we (tediously) can enumerate all rings R with $\mathcal{O} \subset R \subset K$ and $[R : \mathcal{O}] \mid f$, until we find the largest one all of whose elements are integral.

Example 1.4. Consider the ring $\mathcal{O}_K = \mathbf{Z}[(1 + \sqrt{5})/2]$ of integers of $K = \mathbf{Q}(\sqrt{5})$. The discriminant of the basis $1, a = (1 + \sqrt{5})/2$ is

$$\text{Disc}(\mathcal{O}_K) = \left| \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} \right| = 5.$$

Let $\mathcal{O} = \mathbf{Z}[\sqrt{5}]$ be the order generated by $\sqrt{5}$. Then \mathcal{O} has basis $1, \sqrt{5}$, so

$$\text{Disc}(\mathcal{O}) = \left| \begin{pmatrix} 2 & 0 \\ 0 & 10 \end{pmatrix} \right| = 20 = [\mathcal{O}_K : \mathcal{O}]^2 \cdot 5.$$

2 Norms of Ideals

In this section we extend the notion of norm to ideals. This will be helpful in proving of class groups in the next section. For example, we will prove that the group of fractional ideals modulo principal fractional ideals of a number field is finite by showing that every ideal is equivalent to an ideal with norm at most some a priori bound.

Definition 2.1 (Lattice Index). If L and M are two lattices in vector space V , then the *lattice index* $[L : M]$ is by definition the absolute value of the determinant of any linear automorphism A of V such that $A(L) = M$.

The lattice index has the following properties:

- If $M \subset L$, then $[L : M] = \#(L/M)$.
- If M, L, N are lattices then $[L : N] = [L : M] \cdot [M : N]$.

Definition 2.2 (Norm of Fractional Ideal). Suppose I is a fractional ideal of \mathcal{O}_K . The *norm* of I is the lattice index

$$\text{Norm}(I) = [\mathcal{O}_K : I] \in \mathbf{Q}_{\geq 0},$$

or 0 if $I = 0$.

Note that if I is an integral ideal, then $\text{Norm}(I) = \#(\mathcal{O}_K/I)$.

Lemma 2.3. *Suppose $a \in K$ and I is an integral ideal. Then*

$$\text{Norm}(aI) = |\text{Norm}_{K/\mathbf{Q}}(a)| \text{Norm}(I).$$

Proof. By properties of the lattice index mentioned above we have

$$[\mathcal{O}_K : aI] = [\mathcal{O}_K : I] \cdot [I : aI] = \text{Norm}(I) \cdot |\text{Norm}_{K/\mathbf{Q}}(a)|.$$

Here we have used that $[I : aI] = |\text{Norm}_{K/\mathbf{Q}}(a)|$, which is because left multiplication ℓ_a is an automorphism of K that sends I onto aI , so $[I : aI] = |\text{Det}(\ell_a)| = |\text{Norm}_{K/\mathbf{Q}}(a)|$. \square

Proposition 2.4. *If I and J are fractional ideals, then*

$$\text{Norm}(IJ) = \text{Norm}(I) \cdot \text{Norm}(J).$$

Proof. By Lemma 2.3, it suffices to prove this when I and J are integral ideals. If I and J are coprime, then Theorem ?? (Chinese Remainder Theorem) implies that $\text{Norm}(IJ) = \text{Norm}(I) \cdot \text{Norm}(J)$. Thus we reduce to the case when $I = \mathfrak{p}^m$ and $J = \mathfrak{p}^k$ for some prime ideal \mathfrak{p} and integers m, k . By Proposition ?? (consequence of CRT that $\mathcal{O}_K/\mathfrak{p} \cong \mathfrak{p}^n/\mathfrak{p}^{n+1}$), the filtration of $\mathcal{O}_K/\mathfrak{p}^n$ given by powers of \mathfrak{p} has successive quotients isomorphic to $\mathcal{O}_K/\mathfrak{p}$, so we see that $\#(\mathcal{O}_K/\mathfrak{p}^n) = \#(\mathcal{O}_K/\mathfrak{p})^n$, which proves that $\text{Norm}(\mathfrak{p}^n) = \text{Norm}(\mathfrak{p})^n$. \square

Lemma 2.5. *Fix a number field K . Let B be a positive integer. There are only finitely many integral ideals I of \mathcal{O}_K with norm at most B .*

Proof. An integral ideal I is a subgroup of \mathcal{O}_K of index equal to the norm of I . If G is any finitely generated abelian group, then there are only finitely many subgroups of G of index at most B , since the subgroups of index dividing an integer n are all subgroups of G that contain nG , and the group G/nG is finite. This proves the lemma. \square

3 Finiteness of the Class Group via Geometry of Numbers

We have seen examples in which \mathcal{O}_K is not a unique factorization domain. If \mathcal{O}_K is a principal ideal domain, then it is a unique factorization domain, so it is of interest to understand how badly \mathcal{O}_K fails to be a principal ideal domain. The class group of \mathcal{O}_K measures this failure. As one sees in a course on Class Field Theory, the class group and its generalizations also yield deep insight into the possible abelian Galois extensions of K .

Definition 3.1 (Class Group). Let \mathcal{O}_K be the ring of integers of a number field K . The *class group* C_K of K is the group of nonzero fractional ideals modulo the subgroup of principal fractional ideals (a) , for $a \in K$.

Note that if we let $\text{Div}(K)$ denote the group of nonzero fractional ideals, then there is an exact sequence

$$0 \rightarrow \mathcal{O}_K^* \rightarrow K^* \rightarrow \text{Div}(K) \rightarrow C_K \rightarrow 0.$$

A basic theorem in algebraic number theory is that the class group C_K is finite, which follows from the first part of the following theorem and the fact that there are only finitely many ideals of norm less than a given integer.

Theorem 3.2 (Finiteness of the Class Group). *Let K be a number field. There is a constant $C_{r,s}$ that depends only on the number r , s of real and pairs of complex conjugate embeddings of K such that every ideal class of \mathcal{O}_K contains an integral ideal of norm at most $C_{r,s}\sqrt{|d_K|}$, where $d_K = \text{Disc}(\mathcal{O}_K)$. Thus by Lemma 2.5 the class group C_K of K is finite. One can choose $C_{r,s}$ such that every ideal class in C_K contains an integral ideal of norm at most*

$$\sqrt{|d_K|} \cdot \left(\frac{4}{\pi}\right)^s \frac{n!}{n^n}.$$

The explicit bound in the theorem is called the Minkowski bound, and I think it is the best known unconditional general bound (though there are better bounds in certain special cases).

Before proving Theorem 3.2, we prove a few lemmas. The strategy of the proof will be to start with any nonzero ideal I , and prove that there is some nonzero $a \in K$, with very small norm, such that aI is an integral ideal. Then $\text{Norm}(aI) = \text{Norm}_{K/\mathbf{Q}}(a)\text{Norm}(I)$ will be small, since $\text{Norm}_{K/\mathbf{Q}}(a)$ is small. The trick is to determine precisely how small an a we can choose subject to the condition that aI be an integral ideal, i.e., that $a \in I^{-1}$.

Let S be a subset of $V = \mathbf{R}^n$. Then S is *convex* if whenever $x, y \in S$ then the line connecting x and y lies entirely in S . We say that S is *symmetric about the origin* if whenever $x \in S$ then $-x \in S$ also. If L is a lattice in V , then the *volume* of V/L is the volume of the compact real manifold V/L , which is the same thing as the absolute value of the determinant of any matrix whose rows form a basis for L .

Lemma 3.3. *Let L be a lattice in $V = \mathbf{R}^n$, and let S be a bounded closed convex subset of V that is symmetric about the origin. Assume that $\text{Vol}(S) \geq 2^n \text{Vol}(V/L)$. Then S contains a nonzero element of L .*

Proof. First assume that $\text{Vol}(S) > 2^n \cdot \text{Vol}(V/L)$. If the map $\pi : \frac{1}{2}S \rightarrow V/L$ is injective, then

$$\frac{1}{2^n} \text{Vol}(S) = \text{Vol}\left(\frac{1}{2}S\right) \leq \text{Vol}(V/L),$$

a contradiction. Thus π is not injective, so there exist $P_1 \neq P_2 \in \frac{1}{2}S$ such that $P_1 - P_2 \in L$. By symmetry $-P_2 \in \frac{1}{2}S$. By convexity, the average $\frac{1}{2}(P_1 - P_2)$ of P_1 and $-P_2$ is also in $\frac{1}{2}S$. Thus $0 \neq P_1 - P_2 \in S \cap L$, as claimed.

Next assume that $\text{Vol}(S) = 2^n \cdot \text{Vol}(V/L)$. Then for all $\varepsilon > 0$ there is $0 \neq Q_\varepsilon \in L \cap (1 + \varepsilon)S$, since $\text{Vol}((1 + \varepsilon)S) > \text{Vol}(S) = 2^n \cdot \text{Vol}(V/L)$. If $\varepsilon < 1$ then the Q_ε are all in $L \cap 2S$, which is finite since $2S$ is bounded and L is discrete. Hence there exists $Q = Q_\varepsilon \in L \cap (1 + \varepsilon)S$ for arbitrarily small ε . Since S is closed, $Q \in L \cap S$. \square

Lemma 3.4. *If L_1 and L_2 are lattices in V , then*

$$\text{Vol}(V/L_2) = \text{Vol}(V/L_1) \cdot [L_1 : L_2].$$

Proof. Let A be an automorphism of V such that $A(L_1) = L_2$. Then A defines an isomorphism of real manifolds $V/L_1 \rightarrow V/L_2$ that changes volume by a factor of $|\text{Det}(A)| = [L_1 : L_2]$. The claimed formula then follows. \square

Fix a number field K with ring of integers \mathcal{O}_K . Let $\sigma : K \rightarrow V = \mathbf{R}^n$ be the embedding

$$\begin{aligned} \sigma(x) = & (\sigma_1(x), \sigma_2(x), \dots, \sigma_r(x), \\ & \text{Re}(\sigma_{r+1}(x)), \dots, \text{Re}(\sigma_{r+s}(x)), \text{Im}(\sigma_{r+1}(x)), \dots, \text{Im}(\sigma_{r+s}(x))), \end{aligned}$$

where $\sigma_1, \dots, \sigma_r$ are the real embeddings of K and $\sigma_{r+1}, \dots, \sigma_{r+s}$ are half the complex embeddings of K , with one representative of each pair of complex conjugate embeddings. Note that this σ is *not* exactly the same as the one at the beginning of Section 1.

Lemma 3.5.

$$\text{Vol}(V/\sigma(\mathcal{O}_K)) = 2^{-s} \sqrt{|d_K|}.$$

Proof. Let $L = \sigma(\mathcal{O}_K)$. From a basis w_1, \dots, w_n for \mathcal{O}_K we obtain a matrix A whose i th row is

$$(\sigma_1(w_i), \dots, \sigma_r(w_i), \text{Re}(\sigma_{r+1}(w_i)), \dots, \text{Re}(\sigma_{r+s}(w_i)), \text{Im}(\sigma_{r+1}(w_i)), \dots, \text{Im}(\sigma_{r+s}(w_i)))$$

and whose determinant has absolute value equal to the volume of V/L . By doing the following three column operations, we obtain a matrix whose rows are exactly the images of the w_i under *all* embeddings of K into \mathbf{C} , which is the matrix that came up when we defined d_K .

1. Add $i = \sqrt{-1}$ times each column with entries $\text{Im}(\sigma_{r+j}(w_i))$ to the column with entries $\text{Re}(\sigma_{r+j}(w_i))$.
2. Multiply all columns $\text{Im}(\sigma_{r+j}(w_i))$ by $-2i$, thus changing the determinant by $(-2i)^s$.
3. Add each columns with entries $\text{Re}(\sigma_{r+j}(w_i))$ to the the column with entries $-2i\text{Im}(\sigma_{r+j}(w_i))$.

Recalling the definition of discriminant, we see that if B is the matrix constructed by the above three operations, then $\text{Det}(B)^2 = d_K$. Thus

$$\text{Vol}(V/L) = |\text{Det}(A)| = |(-2i)^{-s} \cdot \text{Det}(B)| = 2^{-s} \sqrt{|d_K|}.$$

□

Lemma 3.6. *If I is a nonzero fractional ideal for \mathcal{O}_K , then $\sigma(I)$ is a lattice in V , and*

$$\text{Vol}(V/\sigma(I)) = 2^{-s} \sqrt{|d_K|} \cdot \text{Norm}(I).$$

Proof. We know that $[\mathcal{O}_K : I] = \text{Norm}(I)$ is a nonzero rational number. Lemma 3.5 implies that $\sigma(\mathcal{O}_K)$ is a lattice in V , since $\sigma(\mathcal{O}_K)$ has rank n as abelian group and spans V , so $\sigma(I)$ is also a lattice in V . For the volume formula, combine Lemmas 3.4–3.5 to get

$$\text{Vol}(V/\sigma(I)) = \text{Vol}(V/\sigma(\mathcal{O}_K)) \cdot [\mathcal{O}_K : I] = 2^{-s} \sqrt{|d_K|} \text{Norm}(I).$$

□

Proof of Theorem 3.2. Let K be a number field with ring of integers \mathcal{O}_K , let $\sigma : K \hookrightarrow V \cong \mathbf{R}^n$ be as above, and let $f : V \rightarrow \mathbf{R}$ be the function defined by

$$f(x_1, \dots, x_n) = |x_1 \cdots x_r \cdot (x_{r+1}^2 + x_{(r+1)+s}^2) \cdots (x_{r+s}^2 + x_n^2)|.$$

Notice that if $x \in K$ then $f(\sigma(x)) = |\text{Norm}_{K/\mathbf{Q}}(x)|$.

Let $S \subset V$ be any closed, bounded, convex, subset that is symmetric with respect to the origin and has positive volume. Since S is closed and bounded,

$$M = \max\{f(x) : x \in S\}$$

exists.

Suppose I is any nonzero fractional ideal of \mathcal{O}_K . Our goal is to prove there is an integral ideal aI with small norm. We will do this by finding an appropriate $a \in I^{-1}$. By Lemma 3.6,

$$c = \text{Vol}(V/I^{-1}) = \frac{2^{-s} \sqrt{|d_K|}}{\text{Norm}(I)}.$$

Let $\lambda = 2 \cdot \left(\frac{c}{v}\right)^{1/n}$, where $v = \text{Vol}(S)$. Then

$$\text{Vol}(\lambda S) = \lambda^n \text{Vol}(S) = 2^n \frac{c}{v} \cdot v = 2^n \cdot c = 2^n \text{Vol}(V/I^{-1}),$$

so by Lemma 3.3 there exists $0 \neq a \in I^{-1} \cap \lambda S$. Since M is the largest norm of an element of S , the largest norm of an element of $I^{-1} \cap \lambda S$ is at most $\lambda^n M$, so

$$|\text{Norm}_{K/\mathbf{Q}}(a)| \leq \lambda^n M.$$

Since $a \in I^{-1}$, we have $aI \subset \mathcal{O}_K$, so aI is an integral ideal of \mathcal{O}_K that is equivalent to I , and

$$\begin{aligned} \text{Norm}(aI) &= |\text{Norm}_{K/\mathbf{Q}}(a)| \cdot \text{Norm}(I) \\ &\leq \lambda^n M \cdot \text{Norm}(I) \\ &\leq 2^n \frac{c}{v} M \cdot \text{Norm}(I) \\ &\leq 2^n \cdot 2^{-s} \sqrt{|d_K|} \cdot M \cdot v^{-1} \\ &= 2^{r+s} \sqrt{|d_K|} \cdot M \cdot v^{-1}. \end{aligned}$$

Notice that the right hand side is independent of I . It depends only on r , s , $|d_K|$, and our choice of S . This completes the proof of the theorem, except for the assertion that S can be chosen to give the claim at the end of the theorem, which we leave as an exercise. \square

Corollary 3.7. *Suppose that $K \neq \mathbf{Q}$ is a number field. Then $|d_K| > 1$.*

Proof. Applying Theorem 3.2 to the unit ideal, we get the bound

$$1 \leq \sqrt{|d_K|} \cdot \left(\frac{4}{\pi}\right)^s \frac{n!}{n^n}.$$

Thus

$$\sqrt{|d_K|} \geq \left(\frac{\pi}{4}\right)^s \frac{n^n}{n!},$$

and the right hand quantity is strictly bigger than 1 for any $s \leq n/2$ and any $n > 1$ (exercise). \square

3.1 An Open Problem

Conjecture 3.8. *There are infinitely many number fields K such that the class group of K has order 1.*

For example, if we consider real quadratic fields $K = \mathbf{Q}(\sqrt{d})$, with d positive and square free, many class numbers are probably 1, as suggested by the MAGMA output below. It looks like 1's will keep appearing infinitely often, and indeed Cohen and Lenstra conjecture that they do. Nobody has found a way to prove this yet.

```

> for d in [2..1000] do
  if d eq SquareFree(d) then
    h := ClassNumber(NumberField(x^2-d));
    if h eq 1 then
      printf "%o, ", d;
    end if;
  end if;
end for;

```

```

2, 3, 5, 6, 7, 11, 13, 14, 17, 19, 21, 22, 23, 29, 31, 33, 37,
38, 41, 43, 46, 47, 53, 57, 59, 61, 62, 67, 69, 71, 73, 77, 83,
86, 89, 93, 94, 97, 101, 103, 107, 109, 113, 118, 127, 129, 131,
133, 134, 137, 139, 141, 149, 151, 157, 158, 161, 163, 166, 167,
173, 177, 179, 181, 191, 193, 197, 199, 201, 206, 209, 211, 213,
214, 217, 227, 233, 237, 239, 241, 249, 251, 253, 262, 263, 269,
271, 277, 278, 281, 283, 293, 301, 302, 307, 309, 311, 313, 317,
329, 331, 334, 337, 341, 347, 349, 353, 358, 367, 373, 379, 381,
382, 383, 389, 393, 397, 398, 409, 413, 417, 419, 421, 422, 431,
433, 437, 446, 449, 453, 454, 457, 461, 463, 467, 478, 479, 487,
489, 491, 497, 501, 502, 503, 509, 517, 521, 523, 526, 537, 541,
542, 547, 553, 557, 563, 566, 569, 571, 573, 581, 587, 589, 593,
597, 599, 601, 607, 613, 614, 617, 619, 622, 631, 633, 641, 643,
647, 649, 653, 661, 662, 669, 673, 677, 681, 683, 691, 694, 701,
709, 713, 717, 718, 719, 721, 734, 737, 739, 743, 749, 751, 753,
757, 758, 766, 769, 773, 781, 787, 789, 797, 809, 811, 813, 821,
823, 827, 829, 838, 849, 853, 857, 859, 862, 863, 869, 877, 878,
881, 883, 886, 887, 889, 893, 907, 911, 913, 917, 919, 921, 926,
929, 933, 937, 941, 947, 953, 958, 967, 971, 973, 974, 977, 983,
989, 991, 997, 998,

```

In contrast, if we look at class numbers of quadratic imaginary fields, only a few at the beginning have class number 1.

```

> for d in [1..1000] do
  if d eq SquareFree(d) then
    h := ClassNumber(NumberField(x^2+d));
    if h eq 1 then
      printf "%o, ", d;
    end if;
  end if;
end for;
1, 2, 3, 7, 11, 19, 43, 67, 163

```

It is a theorem that the above list of 9 fields is the complete list with class number 1. More generally, it is possible (in theory), using deep work of Gross, Zagier, and

Goldfeld involving zeta functions and elliptic curves, to enumerate all quadratic number fields with a given class number.