I.M. Norms and traces

Let L/K be a Galois field extension, with $G := \mathrm{Aut}(L/K) = \{\sigma_1, \ldots, \sigma_d\}$, and $\alpha \in L$. (We take the convention here that $\sigma_1 = \mathrm{id}_L$.) Many times we have argued that symmetric functions in the $\sigma_i(\alpha)$ are G-invariant hence belong to K, as $K = \mathrm{Inv}(G)$. (And we should remind the reader again that if L/K isn't Galois, this is false.) Amongst these symmetric functions are the coefficients of the minimal polynomial m_α of α over K, as well as its discriminant.

So the idea of harnessing this argument to produce *maps from L to K* seems obvious, even though we haven't needed it thus far in our study of Galois theory.

I.M.1. DEFINITION. (i) The **trace** for L/K is the K-vector space homomorphism

$$\mathrm{Tr}_{L/K} \colon L \to K$$

defined by $\mathrm{Tr}_{L/K}(\alpha) := \sum_{i=1}^{d} \sigma_i(\alpha)$.

(ii) The **norm** for L/K is the (multiplicative) group homomorphism

$$\mathrm{N}_{L/K} \colon L^* \to K^*$$

defined by $\mathrm{N}_{L/K}(\alpha) := \prod_{i=1}^{d} \sigma_i(\alpha)$.

I.M.2. EXAMPLE. For a quadratic number field $\mathbb{Q}(\sqrt{m})/\mathbb{Q}$, since $\sigma_2(a + b\sqrt{m}) = a - b\sqrt{m}$, we get $\mathrm{Tr}_{\mathbb{Q}(\sqrt{m})/\mathbb{Q}}(a + b\sqrt{m}) = 2a$ and $\mathrm{N}_{\mathbb{Q}(\sqrt{m})/\mathbb{Q}}(a + b\sqrt{m}) = a^2 - mb^2$. The latter, of course, has to do with Pell's equation.

To better understand norm and trace, let's back up and consider the Galois conjugates of α. Certainly all the $\sigma_i(\alpha)$ are roots of m_α, because G sends roots to roots, and α is a root. Since m_α is irreducible, G acts transitively on its roots, and so its roots are precisely the $\sigma_i(\alpha)$. Since these roots are distinct, there must be exactly $n_\alpha := \deg(m_\alpha) = [K(\alpha):K]$ distinct Galois conjugates. Indeed, writing $H_\alpha := \mathrm{Aut}(L/K(\alpha))$ and $d_\alpha := |H_\alpha| = d/n_\alpha$, we remark that those σ_i's in the same coset σH_α yield the same $\sigma_i(\alpha)$, and those in

different cosets don't. So essentially by Lagrange, each distinct Galois conjugate is repeated exactly d_α times.

Denote the multiplication-by-α map by $\mu_\alpha\colon L \to L$. As a K-linear transformation, its minimal polynomial is just m_α, and this divides its characteristic polynomial p_α by Cayley-Hamilton. By structure theory we know[45] that $p_\alpha \mid m_\alpha^d$; and since m_α is irreducible (and $K[x]$ a UFD), this makes p_α a power of m_α. By considering degrees this yields

$$p_\alpha(x) = m_\alpha(x)^{d_\alpha} = \prod_{i=1}^{d}(x - \sigma_i(\alpha)), \tag{I.M.3}$$

so that

$$p_\alpha(x) = x^d - \mathrm{Tr}_{L/K}(\alpha)x^{d-1} + \cdots + (-1)^d \mathrm{N}_{L/K}(\alpha). \tag{I.M.4}$$

Of course, if α is a primitive element, then $m_\alpha = p_\alpha$.

In light of (I.M.4), I should remark that one can extend $\mathrm{N}_{L/K}$ and $\mathrm{Tr}_{L/K}$ to arbitrary finite extensions simply as the determinant and trace of μ_α. So then (I.M.4) is *the definition*, and in the separable case I.M.1 and (I.M.3) also hold with the modification that by $\sigma_1, \ldots, \sigma_d$ we mean the embeddings of L in a normal (or algebraic) closure which fix K. In particular, in the situation of the last paragraph (where $K(\alpha)/K$ may not be Galois), this gives

$$m_\alpha(x) = x^{n_\alpha} - \mathrm{Tr}_{K(\alpha)/K}(\alpha)x^{n_\alpha - 1} + \cdots + (-1)^{n_\alpha}\mathrm{N}_{K(\alpha)/K}(\alpha)$$

since m_α is the characteristic polynomial of μ_α on $K(\alpha)$. We now summarize a few more straightforward facts about norm and trace.

I.M.5. PROPERTIES. **(A)** We have $\mathrm{Tr}_{L/K}(\alpha) = d_\alpha \mathrm{Tr}_{K(\alpha)/K}(\alpha)$ and $\mathrm{N}_{L/K}(\alpha) = (\mathrm{N}_{K(\alpha)/K}(\alpha))^{d_\alpha}$ in the above notation. In particular, if $\kappa \in K$, then $\mathrm{N}_{L/K}(\kappa) = \kappa^d$ and $\mathrm{Tr}_{L/K}(\kappa) = d\kappa$. (One consequence is that $\mathrm{N}_{L/K}(\kappa\alpha) = \kappa^d \mathrm{N}_{L/K}(\alpha)$, whilst K-linearity of $\mathrm{Tr}_{L/K}$ gives $\mathrm{Tr}_{L/K}(\kappa\alpha) = \kappa \mathrm{Tr}_{L/K}(\alpha)$.)

[45] This is just the statement that the invariant factors of L as a $K[\lambda]$-module satisfy $\delta_1 \mid \delta_2 \mid \cdots \mid \delta_s = m_\alpha$ and $p_\alpha = \prod_i \delta_i$. Here we deduce that all the $\delta_i = m_\alpha$, $s = d_\alpha$, and $L \cong (K[x]/(m_\alpha(x))^{\oplus d_\alpha}$ as a $K[x]$-module (with x acting by μ_α).

(B) If M is an intermediate field, then $\mathrm{N}_{L/K} = \mathrm{N}_{M/K} \circ \mathrm{N}_{L/M}$ and $\mathrm{Tr}_{L/K} = \mathrm{Tr}_{M/K} \circ \mathrm{Tr}_{L/M}$. (For the general case finite extension case, think in terms of block matrices.)

(C) If L and K are number fields, with rings of integers $\mathcal{O}_L$ and $\mathcal{O}_K$, $\mathrm{N}_{L/K}$ maps $\mathcal{O}_L \setminus \{0\} \to \mathcal{O}_K \setminus \{0\}$ (multiplicative monoid homomorphism), and $\mathrm{Tr}_{L/K}$ sends $\mathcal{O}_L \to \mathcal{O}_K$. This is because Galois conjugates[46] of algebraic integers have the same (monic, integral) minimal polynomial hence remain algebraic integers.

(D) In particular, if $K = \mathbb{Q}$, then $N_{L/\mathbb{Q}}$ sends $\mathcal{O}_L \setminus \{0\} \to \mathbb{Z} \setminus \{0\}$. The units of $\mathcal{O}_L$ are precisely the elements with norm ± 1: if $\mathrm{N}_{L/\mathbb{Q}}(\ell) = \pm 1$, then the "product of other Galois conjugates" furnishes an inverse in $\mathcal{O}_L$; while if the norm of $\ell\ell'$ is ± 1, this must be true of the norms of ℓ and ℓ'.

We state one more property as a

I.M.6. PROPOSITION. $\mathrm{Tr}_{L/K}$ *is surjective.*

PROOF. Let ℓ_0 be such that $\sigma_1(\ell_0), \ldots, \sigma_d(\ell_0)$ are a basis, which we can do by the Normal Basis Theorem I.I.9. In that case their sum certainly isn't zero; so $k_0 := \mathrm{Tr}_{L/K}(\ell_0) \in K^*$. But then any $k = \frac{k}{k_0}\mathrm{Tr}_{L/K}(\ell_0) = \mathrm{Tr}_{L/K}(\frac{k}{k_0}\ell_0)$, using K-linearity of $\mathrm{Tr}_{L/K}$. □

This seems rather important. What about $\mathrm{N}_{L/K}$? Surjectivity of the norm for $\mathbb{Q}(\sqrt{m})/\mathbb{Q}$ (cf. Example) would mean that every $q \in \mathbb{Q}^*$ could be written as $q = a^2 - mb^2$ for some $a, b \in \mathbb{Q}^*$. But this isn't true: for instance, if $q \in \mathbb{N}$ is prime, this equation says[47] that m is a square mod q (already false for $m = -1$ and $q = 3$).

So $\mathrm{N}_{L/K}$ is not surjective, and describing the image is arithmetically delicate. Except for finite fields, the multiplicative groups L^* need not be finitely generated, and so the same goes for the image. It is easier to say something about the kernel, which is what we do next.

[46] Again, in the non-normal case you have to replace these by the distinct embeddings of L in L^c or $\bar{\mathbb{Q}}$, which is to say, by Galois conjugates in L^c.

[47] One easily checks that we cannot have $q \mid b$.

Hilbert's Theorem 90.

In this subsection we assume once more that L/K is Galois. We begin with a

I.M.7. LEMMA. *Let $\sigma \mapsto \ell_\sigma$ be a map from $G \to L^*$ satisfying*

$$\ell_{\sigma\eta} = \sigma(\ell_\eta)\ell_\sigma \quad (\forall \eta, \sigma \in G). \tag{I.M.8}$$

Then there exists $\ell_0 \in L^*$ such that $\ell_\eta = \ell_0/\eta(\ell_0)$ $(\forall \eta \in G)$.

PROOF. Thinking of the $\sigma_i \in G$ $(i = 1, \dots, d)$ as distinct L-valued characters of L^*, the Dedekind Independence Theorem I.L.8 ensures that they are L-linearly independent. In particular, the linear combination $\sum_{i=1}^d \ell_{\sigma_i}\sigma_i(\cdot)$ is not zero on all of L^*, and there exists $\alpha \in L^*$ making $\ell_0 := \sum_i \ell_{\sigma_i}\sigma_i(\alpha) \neq 0$. Now compute (for any $\eta \in G$)

$$\begin{aligned}
\eta(\ell_0) = \textstyle\sum_i \eta(\ell_{\sigma_i})(\eta\sigma_i)(\alpha) &\underset{\text{(I.M.8)}}{=} \textstyle\sum_i \ell_{\eta\sigma_i}\ell_\eta^{-1}(\eta\sigma_i)(\alpha) \\
&= \ell_\eta^{-1}\left(\textstyle\sum_i \ell_{\eta\sigma_i}(\eta\sigma_i)(\alpha)\right) \\
[\text{reindex}] &= \ell_\eta^{-1}\left(\textstyle\sum_i \ell_{\sigma_i}\sigma_i(\alpha)\right) \\
&= \ell_\eta^{-1}\ell_0,
\end{aligned}$$

which gives the desired result. □

I.M.9. THEOREM (Hilbert). *Given L/K cyclic, with $G \cong \langle\varphi\rangle$. Then*

$$\mathrm{N}_{L/K}(\ell) = 1 \quad \Longleftrightarrow \quad \exists\, \ell_0 \in L \text{ such that } \ell = \ell_0/\varphi(\ell_0).$$

PROOF. $\underline{(\Longleftarrow)}$: this is obvious since

$$\mathrm{N}_{L/K}(\ell_0/\varphi(\ell_0)) = \{\textstyle\prod_i \varphi^i(\ell_0)\}/\{\prod_i \varphi^i(\varphi(\ell_0))\} = 1.$$

$\underline{(\Longrightarrow)}$: Set $\ell_{\varphi^i} := \ell\varphi(\ell)\varphi^2(\ell)\cdots\varphi^{i-1}(\ell)$ for each i (e.g. $\ell_\varphi = \ell$). This is well-defined since $\ell_{\varphi^d} = \prod_{i=0}^{d-1}\varphi^i(\ell) = \mathrm{N}_{L/K}(\ell) = 1 = \ell_1$, and

$$\ell_{\varphi^i}\varphi^j(\ell_{\varphi^i}) = \ell\varphi(\ell)\cdots\varphi^{j-1}(\ell)\cdot\varphi^j(\ell)\cdots\varphi^{i+j-1}(\ell) = \ell_{\varphi^{i+j}}.$$

So we get a map as in I.M.7, and thus the ℓ_0. □

There is an additive analogue of this involving the trace. You may find the proof in [**Jacobson**]:

I.M.10. PROPOSITION. *Let $\eta \mapsto \lambda_\eta$ be a map from $G \to L$ satisfying $\lambda_{\zeta\eta} = \lambda_\zeta + \zeta(\lambda_\eta)$. Then there exists $\lambda_0 \in L$ such that $\lambda_\eta = \lambda_0 - \eta(\lambda_0)$ for all $\eta \in G$.*

Here are some simple applications, the first of which features Pythagorean triples:

I.M.11. EXAMPLE. Consider $L/K = \mathbb{Q}(\mathbf{i})/\mathbb{Q}$, with $G = \langle\rho\rangle \cong \mathbb{Z}_2$ (ρ = complex conjugation). Suppose $\mathrm{N}_{\mathbb{Q}(\mathbf{i})/\mathbb{Q}}(a + b\mathbf{i}) = a^2 + b^2 = 1$. Then by Hilbert's theorem, there is some $c + d\mathbf{i} \in \mathbb{Q}(\mathbf{i})$ such that

$$a + b\mathbf{i} = \frac{c + d\mathbf{i}}{\rho(c + d\mathbf{i})} = \frac{c + d\mathbf{i}}{c - d\mathbf{i}} = \frac{(c + d\mathbf{i})(c + d\mathbf{i})}{c^2 + d^2} = \frac{c^2 - d^2}{c^2 + d^2} + \frac{2cd}{c^2 + d^2}\mathbf{i},$$

and conversely every element of this form has norm 1. This gives a rational parametrization of all rational points on the unit circle.

The next application revisits a step in the proof of Galois's theorem (which was difficult if d was not prime):

I.M.12. COROLLARY. *Let K contain d distinct d^{th} roots of unity, and L/K be cyclic of degree d. Then $L = K(\alpha)$ for some $\alpha \in L$ with $\alpha^d \in K$.*

PROOF. Let $\zeta \in K$ be a primitive d^{th} root of 1, and $G = \langle\varphi\rangle$. We have $\mathrm{N}_{L/K}(\zeta) = \zeta^d = 1$. By Hilbert's Theorem, there is some $\alpha \in L$ for which $\zeta = \alpha/\varphi(\alpha)$. Rewriting this as $\varphi(\alpha) = \alpha\zeta^{-1}$, we see that $\varphi(\alpha^d) = \alpha^d$ ($\implies \alpha^d \in K$) and $\varphi^i(\alpha) = \alpha\zeta^{-i}$. So we have $G(\alpha) = \{\alpha, \alpha\zeta, \ldots, \alpha\zeta^{d-1}\}$, whence $\deg(m_\alpha) = d$ and $L = K(\alpha)$. □

I.M.13. EXAMPLE. Consider $K = \mathbb{Q}(\zeta_3)$, $L = K(\sqrt[3]{2})$, $\mathrm{Aut}(L/K) = \langle\varphi\rangle \cong \mathbb{Z}_3$. There exists $\alpha \in L$ such that $\varphi(\alpha) = \alpha\zeta_3^{-1}$. Guess what α is?

There is an analogous application of the additive analogue, related to splitting fields of polynomials like $x^p - x + k$:[48]

I.M.14. COROLLARY. *Let $\mathrm{char}(K) = p > 0$, and L/K be cyclic of degree p. Then $L = K(\beta)$ for some $\beta \in L$ with $\beta^p - \beta \in K$.*

[48] Again, see **[Jacobson]** for proof.

Rings of integers revisited [Part 1: cyclotomic case].

We are now in a position to verify the unproved assertions from [**Algebra I**, §§III.J,L] on number rings: the general ones about ideals in $\mathcal{O}_K$; and the specific ones about cyclotomic rings of integers, which we will do first.

The next theorem summarizes those claims made in [**Algebra I**, §III.L] that were used to prove Fermat's last theorem for exponents not divisible by an irregular prime:

I.M.15. THEOREM. *Let $p > 2$ be a prime number and $K = \mathbb{Q}(\zeta_p)$ the corresponding cyclotomic field.*[49]

(i) *The ring of integers is $\mathcal{O}_K = \mathbb{Z}[\zeta_p]$.*

(ii) *The roots of unity in K are the $\pm\zeta_p^j$.*

(iii) *If $u \in \mathcal{O}_K^*$ is a unit, then $u/\bar{u}$ is a root of* 1 (*of the form in* (ii)).

We begin with the easiest part:

PROOF OF (ii). If there are any other roots of unity in K besides the powers of ζ_{2p}, then we have $\mathbb{Q}(\zeta_m) \subseteq \mathbb{Q}(\zeta_p)$ for some m having $2p$ as proper divisor. By the Tower Law, $\phi(m) = [\mathbb{Q}(\zeta_m):\mathbb{Q}]$ divides $[K:\mathbb{Q}] = p-1$, which contradicts the formula for $\phi(m)$. □

Let's deal with (i) next. We have $\mathbb{Z}_p^* \cong \mathrm{Aut}(K/\mathbb{Q}) = \{\sigma_j\}_{j=1}^{p-1}$, with $\sigma_j(\zeta_p) := \zeta_p^j$. So for example $\mathrm{Tr}_K(1) = p-1$ while $\mathrm{Tr}_K(\zeta_p^j) = \sum_{i=1}^{p-1} \zeta_p^{ij} = \sum_{i'=1}^{p-1} \zeta_p^{i'} = -1$ for $j = 1, \ldots, p-1$.

Though we won't discuss this, computing rings of integers in general is difficult, though there is an algorithm. The idea is to take a basis of $K/\mathbb{Q}$ inside $\mathcal{O}_K$ (like powers of a primitive element in $\mathcal{O}_K$) and consider possible enlargements so long as the square of any prime divides the discriminant of said basis, cf. I.M.18 below. (Though you have to be able to decide whether specific elements are in $\mathcal{O}_K$, and this requires computing characteristic polynomials, which can be labor intensive.) Fortunately, nothing like this is required to deal with $\mathcal{O}_{\mathbb{Q}(\zeta_p)}$, only the following

[49] We keep this notation through the end of this subsection.

I.M.16. LEMMA. (a) $(1-\zeta_p)\mathcal{O}_K \cap \mathbb{Z} = p\mathbb{Z}$.
(b) *If* $\alpha \in \mathcal{O}_K$, *then* $\tau_\alpha := \mathrm{Tr}_K(\alpha(1-\zeta_p)) \in p\mathbb{Z}$.

PROOF. (a) From $p = \Phi_p(1) = \prod_{j=1}^{p-1}(1-\zeta_p^j) = \mathrm{N}_K(1-\zeta_p)$, we have that $p\mathbb{Z} \subset (1-\zeta_p)\mathcal{O}_K \cap \mathbb{Z} \subset \mathbb{Z}$. Clearly the first or second inclusion is an equality. If the first, we win. If the second, then $\exists \alpha \in \mathcal{O}_K$ such that $1 = (1-\zeta_p)\alpha$, and taking norms gives $1 = N_K(1-\zeta_p)N_K(\alpha) = pN_K(\alpha)$; but $N_K(\alpha) \in \mathbb{Z}$ (as $\alpha \in \bar{\mathbb{Z}}$), and this is a contradiction.

(b) First, $\alpha(1-\zeta_p) \in \mathcal{O}_K \implies \tau_\alpha \in \mathbb{Z}$. Second, since each $\sigma_j(1-\zeta_p) = 1-\zeta_p^j = (1-\zeta_p)(1+\zeta_p+\cdots+\zeta_p^{j-1}) \in (1-\zeta_p)\mathcal{O}_K$ and each $\sigma_j(\alpha) \in \mathcal{O}_K$, we have $\tau_\alpha \in (1-\zeta_p)\mathcal{O}_K$. Conclude by (a) that $\tau_\alpha \in p\mathbb{Z}$. □

PROOF OF I.M.15(i). We need to prove $\mathcal{O}_K \subset \mathbb{Z}[\zeta_p]$. Given $\alpha \in \mathcal{O}_K$, we can write $\alpha = \sum_{j=0}^{p-2} q_j\zeta_p^j$ for some $q_j \in \mathbb{Q}$; the goal is to show each $q_j \in \mathbb{Z}$. We compute

$$\begin{aligned}\tau_\alpha &= \mathrm{Tr}_K\left(\textstyle\sum_{j=0}^{p-2} q_j(\zeta_p^j - \zeta_p^{j+1})\right) = \textstyle\sum_{j=0}^{p-2} q_j(\mathrm{Tr}_K\zeta_p^j - \mathrm{Tr}_K\zeta_p^{j+1}) \\ &= q_0((p-1)-(-1)) = q_0 p,\end{aligned}$$

whence I.M.16(b) $\implies q_0 \in \mathbb{Z}$. Now set $\alpha_1 := \sum_{j=1}^{p-2} q_j\zeta_p^{j-1} = (\alpha - q_0)\zeta_p^{p-1} \in \mathcal{O}_K$, repeat the argument to get $q_1 \in \mathbb{Z}$, and continue in this fashion with $\alpha_2 := (\alpha_1 - q_1)\zeta_p^{p-1} \in \mathcal{O}_K$ etc. □

Turning finally to the units, what we need is the famous and *extremely useful*

I.M.17. LEMMA (Kronecker). *Let* $\alpha \in \bar{\mathbb{Z}}$ *be an algebraic integer, so that its minimal polynomial* $m_\alpha(x) = \prod_{i=1}^{n}(x-\alpha_i) \in \mathbb{Z}[x]$. *If all of* α*'s Galois conjugates have absolute values* $|\alpha_i| = 1$, *then* α *is a root of unity.*

PROOF. Write $m_\alpha = x^n + a_1x^{n-1} + \cdots + a_n$, with all $a_k \in \mathbb{Z}$. Since each $|\alpha_i| \le 1$, we have also $|a_k| \le \binom{n}{k}$ $(\forall k)$. There are only *finitely many* such polynomials (monic integral, of the same degree or smaller, with these bounds on the coefficients). Moreover, if α satisfies the assumptions of the lemma, so do its powers; and they lie

in the same $\mathbb{Q}(\alpha)$ hence have minimal polynomials of no higher degree than n $(= [\mathbb{Q}(\alpha):\mathbb{Q}])$. So α, α^2, α^3, etc. are all among the *finitely many* roots of the finite set of polynomials just exhibited. Two distinct powers of α must therefore be equal. □

PROOF OF I.M.15(iii). Given our unit $u \in \mathcal{O}_K^*$, also $\bar{u} = \sigma_{p-1}(u) \in \mathcal{O}_K^*$ hence $u/\bar{u} \in \mathcal{O}_K$. Indeed, since σ_{p-1} is just complex conjugation on K, and $\mathrm{Aut}(K/\mathbb{Q})$ is abelian, complex conjugation commutes with all the σ_j's. So $\sigma_j(u/\bar{u}) = \sigma_j(u)/\overline{\sigma_j(u)}$ evidently has absolute value 1 for all j, and we are done by Kronecker's lemma. □

Rings of integers revisited [Part 2: ideals and norms].

It will be useful to have the following bit of language: a **full lattice** Λ in a $\mathbb{Q}$-vector space V of dimension n is a subgroup which is free abelian of rank n. Equivalently, Λ is of the form $\mathbb{Z}\langle\underline{\beta}\rangle (\cong \mathbb{Z}^n)$ with $\underline{\beta} = \{\beta_1, \ldots, \beta_n\}$ a $\mathbb{Q}$-basis of V, and we say that $\underline{\beta}$ is a **basis** of Λ.

We begin with a generalization of our earlier notion of discriminant. Throughout $K/\mathbb{Q}$ is a number field of degree $[K:\mathbb{Q}] = n$, and we write $\mathrm{Tr}_K := \mathrm{Tr}_{K/\mathbb{Q}}$ and $\mathrm{N}_K := \mathrm{N}_{K/\mathbb{Q}}$.

I.M.18. DEFINITION. The **discriminant** of $\underline{\gamma} = (\gamma_1, \ldots, \gamma_n) \in K^n$ is $\Delta_K(\underline{\gamma}) := \det(Q(\underline{\gamma}))$, where $Q(\underline{\gamma}) := [\mathrm{Tr}_K(\gamma_i\gamma_j)]_{1\le i,j\le n}$.

One should think of this as the square of a measure of the covolume of the lattice spanned by $\underline{\gamma}$, a point of view which is justified by I.M.19 and I.M.21 below:

I.M.19. PROPOSITION. (i) $\Delta_K(\underline{\gamma}) \in \mathbb{Q}$. *Moreover, if* $\gamma_i \in \mathcal{O}_K$, *then* $\Delta_K(\underline{\gamma}) \in \mathbb{Z}$.

(ii) *For* $M \in M_n(\mathbb{Q})$ *and* $\underline{\delta} := M\underline{\gamma}$ (*thinking of* $\underline{\delta}, \underline{\gamma}$ *as column vectors*), $\Delta_K(\underline{\delta}) = (\det(M))^2\Delta_K(\underline{\gamma})$.

(iii) *If* $\underline{\gamma}$ *and* $\underline{\gamma}' = M\underline{\gamma}$ *are bases of full lattices* Λ *and* Λ', *with* $\Lambda' \subseteq \Lambda$, *then* $|\Lambda/\Lambda'| = |\det M|$, *and* $\Lambda = \Lambda' \iff |\det M| = 1$. *Hence* $\Delta_K(\Lambda) := |\Delta_K(\text{any basis of } \Lambda)|$ *is well-defined, and*

$$\Delta_K(\Lambda') = |\Lambda/\Lambda'|^2\Delta_K(\Lambda). \tag{I.M.20}$$

PROOF. (i) is evident since the entries of $Q(\underline{\gamma})$ are in $\mathbb{Q}$ resp. $\mathbb{Z}$. For (ii), the $(i,j)^{\text{th}}$ entry of $M \cdot Q(\underline{\gamma}) \cdot {}^tM$ is

$$\begin{aligned}\textstyle\sum_{k=1}^n \sum_{\ell=1}^n M_{ik}\mathrm{Tr}_K(\gamma_k\gamma_\ell)M_{j\ell} &= \mathrm{Tr}_K\left(\textstyle\sum_k \sum_\ell M_{ik}\gamma_k\gamma_\ell M_{j\ell}\right)\\ &= \mathrm{Tr}_K\left((\textstyle\sum_k M_{ik}\gamma_k)(\sum_\ell M_{j\ell}\gamma_\ell)\right)\\ &= \mathrm{Tr}_K(\delta_i\delta_j).\end{aligned}$$

So $M\,Q(\underline{\gamma})\,{}^tM = Q(\underline{\delta})$, and taking det of both sides gives the desired relation.

Finally, for (iii), use row and column operations on M to bring it into the form $\mathrm{diag}(d_1,\dots,d_n)$, whence $\Lambda/\Lambda' \cong \mathbb{Z}_{d_1} \times \cdots \times \mathbb{Z}_{d_n}$. Since the matrices producing those operations are in $\mathrm{GL}_n(\mathbb{Z})$, they have ± 1 determinants and so $|\det M| = d_1 \cdots d_n = |\Lambda/\Lambda'|$. The remaining statements are self-explanatory. □

Let θ be a primitive element, so that $K = \mathbb{Q}(\theta)$, and write $p_\theta(x) =: \prod_{i=1}^n (x - \theta_i)$. Obviously $p_\theta = m_\theta$ is irreducible and the θ_i are distinct. Denote by Θ the basis $\{1, \theta, \theta^2, \dots, \theta^{n-1}\}$ of $K/\mathbb{Q}$, and note that $Q(\Theta) = [\mathrm{Tr}_K(\theta^{i+j-2})]$.

We recall from (I.M.4) that $\prod_i \theta_i = \mathrm{N}_K(\theta)$ and $\sum_i \theta_i = \mathrm{Tr}_K(\theta)$, and that one has $\theta_i = \sigma_i(\theta)$ where $\sigma_1, \dots, \sigma_n$ are the distinct embeddings of K in $\mathbb{C}$ (or equivalently, a system of representatives of the n cosets of $\mathrm{Aut}(K^c/\mathbb{Q})/\mathrm{Aut}(K^c/K)$).

I.M.21. PROPOSITION. $\Delta_K(\Theta) = \prod_{i<j}(\theta_i - \theta_j)^2$ $(\neq 0)$.

PROOF. Writing $A = [\theta_i^{j-1}]_{1\le i,j\le n}$, $\det(A) = \prod_{i<j}(\theta_j - \theta_i)$ is the Vandermonde determinant. Now

$$[{}^tA \cdot A]_{(i,j)} = \textstyle\sum_{k=1}^n \theta_k^{i-1}\theta_k^{j-1} = \mathrm{Tr}_K(\theta^{i-1}\theta^{j-1}) = Q(\Theta)_{(i,j)}.$$

So $Q(\Theta) = {}^tAA$, and taking determinants gives $\Delta_K(\Theta) = (\det(A))^2$ as desired. □

So we recover the polynomial discriminant in this case: $\Delta_K(\Theta)$ is just the discriminant Δ_{m_θ} of the *polynomial* m_θ from I.K.1.

I.M.22. COROLLARY. $\Delta_K(\underline{\gamma}) \neq 0 \iff \underline{\gamma}$ *is a basis of* $K/\mathbb{Q}$.

PROOF. Whether or not $\underline{\gamma}$ is a basis, Θ is one; so there exists $M \in M_n(\mathbb{Q})$ such that $\underline{\gamma} = M\Theta$, and then $\Delta_K(\underline{\gamma}) = (\det M)^2 \Delta_K(\Theta)$ by I.M.19(ii), where $\Delta_K(\Theta) \neq 0$ by I.M.21. Now $\underline{\gamma}$ is a basis $\iff$ M is invertible $\iff$ $\det(M) \neq 0 \iff \Delta_K(\underline{\gamma}) \neq 0$. □

Another useful formula for the discriminant of a basis is

$$\Delta_K(\underline{\gamma}) = (\det \Sigma_\gamma)^2 := (\det[\sigma_i(\gamma_j)])^2. \tag{I.M.23}$$

We see this by writing the RHS as $\det({}^t\Sigma_\gamma \Sigma_\gamma)$ then computing

$$({}^t\Sigma_\gamma \Sigma_\gamma)_{(i,j)} = \sum_k \sigma_k(\gamma_i)\sigma_k(\gamma_j) = \sum_k \sigma_k(\gamma_i\gamma_j) = \mathrm{Tr}_K(\gamma_i\gamma_j) = Q(\underline{\gamma})_{(i,j)}.$$

Next recall that $\mathcal{I}(K)$, the *monoid of integral ideals*, comprises the nonzero ideals in $\mathcal{O}_K$. We are ready to prove [**Algebra I**, III.L.5(iii)]:[50]

I.M.24. THEOREM. *Any $I \in \mathcal{I}(K)$ is a full lattice in K: that is, I is finitely generated as an abelian group, and isomorphic to $\mathbb{Z}^{[K:\mathbb{Q}]}$.*

PROOF. By [**Algebra I**, III.J.17(a)(i)], I contains a basis $\mathfrak{B}$ for $K/\mathbb{Q}$. [To recap: multiply a given basis of K by a suitably large integer, so that the minimal polynomials of the basis elements become monic integral, making them elements of $\mathcal{O}_K$. Then multiply them by a single nonzero element of I.] Any such $\mathfrak{B}$ satisfies $\Delta_K(\mathfrak{B}) \in \mathbb{Z}\backslash\{0\}$ by I.M.19(i) and I.M.22.

Choosing $\mathfrak{B} = \langle \beta_1, \ldots, \beta_n \rangle \subset I$ with minimal $|\Delta_K(\mathfrak{B})|$, suppose that $I \setminus \mathbb{Z}\langle\mathfrak{B}\rangle$ is nonempty, and take any α in it. Then $\alpha = \sum_i q_i\beta_i$ with all $q_i \in \mathbb{Q}$ (since $\mathfrak{B}$ is a $\mathbb{Q}$-basis of K), but some q_i (say q_1) not in $\mathbb{Z}$. Defining $\mathfrak{B}' := (\alpha - \lfloor q_1 \rfloor \beta_1, \beta_2, \ldots, \beta_n)$, we note that:

- $\mathfrak{B}' = \left(\begin{array}{c|ccc} q_1 - \lfloor q_1 \rfloor & q_2 & \cdots & q_n \\ \hline 0 & 1 & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & & 1 \end{array}\right) \mathfrak{B} =: M'\mathfrak{B}$ is clearly a basis of $K/\mathbb{Q}$ (as M' is invertible);
- $\alpha - \lfloor q_1 \rfloor \beta_1 \in I \implies \mathfrak{B}' \subset I$ ($\implies \Delta_K(\mathfrak{B}') \in \mathbb{Z}$); and
- $0 < q_1 - \lfloor q_1 \rfloor < 1$.

[50] as well as the assumption of [**Algebra I**, III.J.17(a)(ii)].

So $|\Delta_K(\mathfrak{B}')| = (\det M')^2|\Delta_K(\mathfrak{B})| = (q_1 - \lfloor q_1 \rfloor)^2|\Delta_K(\mathfrak{B})| < |\Delta_K(\mathfrak{B})|$ contradicts minimality of $|\Delta_K(\mathfrak{B})|$.

Conclude that $I \setminus \mathbb{Z}\langle\mathfrak{B}\rangle$ is in fact empty, i.e. $I = \mathbb{Z}\langle\mathfrak{B}\rangle$. □

I.M.25. DEFINITION. A basis of $\mathcal{O}_K$, which exists by I.M.22, is called an *integral basis of K*. The **discriminant of K**, written $\boldsymbol{\Delta_K}$, is the discriminant $\Delta_K(\mathcal{O}_K)$ of any integral basis of K.

I.M.26. EXAMPLE. Given $D \in \mathbb{Z}\setminus\{0\}$ squarefree and $\underset{(4)}{\equiv} 2$ or 3, the ring of integers $\mathcal{O}_K$ of $K = \mathbb{Q}(\sqrt{D})$ has basis $(1, \sqrt{D})$ and discriminant $\Delta_K = \det \begin{pmatrix} \mathrm{Tr}(1) & \mathrm{Tr}(\sqrt{D}) \\ \mathrm{Tr}(\sqrt{D}) & \mathrm{Tr}(D) \end{pmatrix} = \det \left(\begin{smallmatrix} 2 & 0 \\ 0 & 2D \end{smallmatrix}\right) = 4D$. For $D \underset{(4)}{\equiv} 1$, a similar computation gives $\Delta_K = D$.

The next result settles [**Algebra I**, III.L.5(ii)], except for multiplicativity:

I.M.27. PROPOSITION-DEFINITION. (i) *For any $I \in \mathcal{I}(K)$, the **ideal norm** $\mathfrak{N}(I) := |\mathcal{O}_K/I|$ is finite.*

(ii) *If $I \supseteq J$, then $\mathfrak{N}(I) \cdot |I/J| = \mathfrak{N}(J)$; in particular, $\mathfrak{N}(I) \leq \mathfrak{N}(J)$.*

(iii) ["strictness"] *If $I \supsetneq J$, then $\mathfrak{N}(I) < \mathfrak{N}(J)$.*

(iv) *If $\alpha \in \mathcal{O}_K$, then $\mathfrak{N}((\alpha)) = |\mathrm{N}_K(\alpha)|$.*

PROOF. By I.M.24, $\mathcal{O}_K$ and I are full lattices, and so (I.M.20) $\implies |\mathcal{O}_K/I|^2 = \frac{\Delta_K(I)}{\Delta_K}$, in which numerator and denominator are nonzero by I.M.22. This gives (i).

For (ii) and (iii), note that as abelian groups, $J \leq I \leq \mathcal{O}_K \implies \mathcal{O}_K/J \cong \frac{\mathcal{O}_K/I}{I/J} \implies |\mathcal{O}_K/I| = \frac{|\mathcal{O}_K/I|}{|I/J|}$. Hence, if $\mathfrak{N}(I) = \mathfrak{N}(J)$, then $|I/J| = 1$ and $I = J$.

Finally, if we write $\mathfrak{B} = \{\beta_1, \ldots, \beta_n\}$ for a basis of $\mathcal{O}_K$, then $\mathfrak{B}' = \{\alpha\beta_1, \ldots, \alpha\beta_n\}$ is a basis of (α). But then, if $M_\alpha = (m_{ij})$ is the matrix of multiplication by α with respect to $\mathfrak{B}$, we have $\alpha\beta_j = \sum_i m_{ij}\beta_i$, which is to say $\mathfrak{B}' = {}^tM_\alpha\mathfrak{B}$. Thus, by I.M.19(iii) $\mathfrak{N}((\alpha)) \overset{\text{def}}{=} |\mathcal{O}_K/(\alpha)| = \left|\det {}^tM_\alpha\right| = |\det M_\alpha| \overset{\text{def}}{=} |\mathrm{N}_K(\alpha)|$. □

Turning to maximal ideals and their invertibility, we need two lemmas which make essential use of the ideal norm. The first highlights something very special about number rings:

I.M.28. LEMMA. *Prime ideals in a number ring $\mathcal{O}_K$ are maximal.*[51]

PROOF. Let P be a prime ideal in $\mathcal{O}_K$. By I.M.27(i), $\mathcal{O}_K/P$ is finite. Since P is prime, $\mathcal{O}_K/P$ is a domain. So by Wedderburn's little theorem [**Algebra I**, III.F.18], $\mathcal{O}_K/P$ is a field, and P is maximal. □

I.M.29. LEMMA. *Any $I \in \mathcal{I}(K)$ contains a product of maximal ideals.*

PROOF. Induce on $\mathfrak{N}(I)$. By strictness, $\mathfrak{N}(I) = 1 \implies I = \mathcal{O}_K$ (base case). Assuming the statement for ideals with norm $\leq k$, suppose $\mathfrak{N}(I) = k+1$, and also that I isn't prime (if it is, we're done). Then there exist $\alpha_1, \alpha_2 \in \mathcal{O}_K \backslash I$ such that $\alpha_1\alpha_2 \in I$. Writing $J_i := I + (\alpha_i)$, strictness implies $\mathfrak{N}(J_i) < k+1$. So $J_i \supseteq \prod_j P_{ij}$ by the inductive hypothesis, and

$$(\textstyle\prod_j P_{1j})(\prod_j P_{2j}) \subset J_1J_2 = I^2 + \alpha_1 I + \alpha_2 I + \alpha_1\alpha_2 \subset I$$

completes the induction. □

We can now establish [**Algebra I**, III.L.5(i)], i.e. that each maximal/prime ideal has a fractional-ideal inverse:

I.M.30. THEOREM. *For any maximal ideal $P \subset \mathcal{O}_K$, there exists $\gamma \in K\backslash\mathcal{O}_K$ such that $\gamma P \subseteq \mathcal{O}_K$; and then $((\gamma) + \mathcal{O}_K)P = \mathcal{O}_K$, i.e. $(\gamma) + \mathcal{O}_K$ yields an inverse $P^{-1} \in \mathcal{J}(K)$.*

PROOF. Let $\alpha \in P\backslash\{0\}$. By I.M.29, there exists a product of maximal ideals $P_1 \cdots P_t \subseteq (\alpha)$, where we take t to be as small as possible. Renumbering the P_i if necessary, $P \supset (\alpha) \supset \prod_{i=1}^t P_i$ (i.e. $P \mid \prod P_i$) $\implies P \supset P_1$ (i.e. $P \mid P_1$) since P is prime; and since P_1 is maximal, $P = P_1$.

Now minimality of $t \implies (\alpha) \not\supset P_2 \cdots P_t \implies \exists\, \beta \in P_2 \cdots P_t \backslash (\alpha)$ $\implies \frac{\beta}{\alpha} =: \gamma \notin \mathcal{O}_K$ (since otherwise $\beta = \alpha\frac{\beta}{\alpha} \in \alpha\mathcal{O}_K = (\alpha)$)

$$\implies \quad \gamma P = \alpha^{-1}P \cdot (\beta) \subset \alpha^{-1}P_1P_2 \cdots P_t \subset \alpha^{-1}(\alpha) = \mathcal{O}_K$$

proves the first statement.

[51] Remember that maximal ideals are always prime, but not conversely in general.

Writing $P' = (\gamma) + \mathcal{O}_K$, we have $P \subset \gamma P + P\,(= P'P) \subset \mathcal{O}_K$; so by maximality of P, either $\gamma P + P = \mathcal{O}_K$ (and we are done) or $\gamma P + P = P \implies \gamma P \subset P$. In the latter event, writing the matrix of μ_γ with respect to a basis of P, we see that it has a monic integral characteristic polynomial, whence $\gamma \in \mathcal{O}_K$, a contradiction. □

It remains to show that the ideal norm is multiplicative. We already know (by I.M.27(ii)) that for $I \supseteq J$, $\mathfrak{N}(I) \mid \mathfrak{N}(J)$, and thus that an ideal with prime norm is prime (why?). (The converse isn't quite true: the norm of a prime ideal is in general a prime power.) I also claim that for any $I \in \mathcal{I}(K)$, we have $\mathfrak{N}(I) \in I$ hence $I \supseteq (\mathfrak{N}(I))$.[52] This is because the order of $1 + I \in \mathcal{O}_K/I$ must divide $|\mathcal{O}_K/I| = \mathfrak{N}(I)$, and so $\mathfrak{N}(I)(1 + I) = 0 \implies \mathfrak{N}(I) \in I$.

We also know that any $I \in \mathcal{I}(K)$ is a product of maximal ideals, since this didn't use multiplicativity of the ideal norm (only I.M.27(ii-iii) and I.M.30).

I.M.31. LEMMA. *Given $I, P \in \mathcal{I}(K)$, with P prime, we have $|I/PI| = |\mathcal{O}_K/P|\,(= \mathfrak{N}(P))$, hence*

$$\mathfrak{N}(PI) = |\mathcal{O}_K/PI| = |\mathcal{O}_K/I||I/PI| = \mathfrak{N}(I)\mathfrak{N}(P).$$

PROOF. First of all, $I \backslash PI$ is nonempty, since otherwise $I = PI \implies \mathcal{O}_K = II^{-1} = PII^{-1} = P$. Given $\alpha \in I \backslash PI$, we have $(\alpha) \not\subset PI \implies \alpha I^{-1} \not\subset P$ and $(\alpha) \subset I \implies \alpha I^{-1} \subset II^{-1} = \mathcal{O}_K$. Since P is maximal and $P \subsetneq \alpha I^{-1} + P \subseteq \mathcal{O}_K$, we get $\alpha I^{-1} + P = \mathcal{O}_K$, and $\alpha\beta + \pi = 1$ for some $\beta \in I^{-1}$ and $\pi \in P$.

Clearly multiplication by α induces a homomorphism of abelian groups from $\mathcal{O}_K/P$ to I/PI, and multiplication by β gives a map back. To check that these are mutually inverse, just note that by $\alpha\beta + \pi = 1$, $\lambda + P = \alpha\beta\lambda + P$ and $\mu + PI = \alpha\beta\mu + PI$. □

Writing an arbitrary ideal as a product of maximals, $I = \prod_i P_i$, and iteratively applying I.M.31, gives $\mathfrak{N}(I) = \prod_i \mathfrak{N}(P_i)$. By factoring any other ideal J, and thus the product IJ, we arrive at the

[52] This was used in the proof of [**Algebra I**, III.L.20], if you want it to work for arbitrary number fields.

I.M.32. THEOREM. $I, J \in \mathcal{I}(K) \implies \mathfrak{N}(IJ) = \mathfrak{N}(I)\mathfrak{N}(J)$.

One can now have complete confidence in the results of [**Algebra I**, §III.L] for general number fields.