

Chapter 3

Algebraic numbers and algebraic number fields

Literature:

S. Lang, Algebra, 2nd ed. Addison-Wesley, 1984. Chaps. III,V,VII,VIII,IX.

P. Stevenhagen, Dictaat Algebra 2, Algebra 3 (Dutch).

We have collected some facts about algebraic numbers and algebraic number fields that are needed in this course. Many of the results are stated without proof. For proofs and further background, we refer to Lang's book mentioned above, Peter Stevenhagen's Dutch lecture notes on algebra, and any basic text book on algebraic number theory. We do not require any pre-knowledge beyond basic ring theory. In the Appendix (Section 3.5) we have included some general theory on ring extensions for the interested reader.

3.1 Algebraic numbers and algebraic integers

A number $\alpha \in \mathbb{C}$ is called *algebraic* if there is a non-zero polynomial $f \in \mathbb{Q}[X]$ with $f(\alpha) = 0$. Otherwise, α is called *transcendental*. We define the *algebraic closure* of \mathbb{Q} by

$$\overline{\mathbb{Q}} := \{\alpha \in \mathbb{C} : \alpha \text{ algebraic}\}.$$

Lemma 3.1. (i) $\overline{\mathbb{Q}}$ is a subfield of \mathbb{C} , i.e., sums, differences, products and quotients of algebraic numbers are again algebraic;

(ii) $\overline{\mathbb{Q}}$ is algebraically closed, i.e., if $g = X^n + \beta_1 X^{n-1} \cdots + \beta_n \in \overline{\mathbb{Q}}[X]$ and $\alpha \in \mathbb{C}$ is a zero of g , then $\alpha \in \overline{\mathbb{Q}}$.

(iii) If $g \in \overline{\mathbb{Q}}[X]$ is a monic polynomial, then $g = (X - \alpha_1) \cdots (X - \alpha_n)$ with $\alpha_1, \dots, \alpha_n \in \overline{\mathbb{Q}}$.

Proof. This follows from some results in the Appendix (Section 3.5). Proposition 3.26 in Section 3.5 with $A = \mathbb{Q}$, $B = \mathbb{C}$ implies that $\overline{\mathbb{Q}}$ is a ring. To prove that $\overline{\mathbb{Q}}$ is a field we have to show that if α is a non-zero algebraic number then also α^{-1} is algebraic. Indeed, for such α there are $a_0, \dots, a_d \in \mathbb{Q}$ such that $a_0 \alpha^d + a_1 \alpha^{d-1} + \cdots + a_d = 0$ and $a_0 a_d \neq 0$. But then, $a_d \alpha^{-d} + \cdots + a_0 = 0$. This proves part (i). Part (ii) follows at once from Proposition 3.26, while (iii) is a consequence of (ii). \square

Let $\alpha \in \overline{\mathbb{Q}}$. Among all polynomials with rational coefficients having α as a root there is a unique one of minimal degree that is monic (indeed, take two monic polynomials in $\mathbb{Q}[X]$ of minimal degree, say d , having α as a root; then their difference has degree $< d$, hence must be 0). This polynomial is called the *minimal polynomial* of α , notation f_α .

Lemma 3.2. *Let $\alpha \in \overline{\mathbb{Q}}$. Then f_α is irreducible in $\mathbb{Q}[X]$ and f_α divides all polynomials in $g \in \mathbb{Q}[X]$ with $g(\alpha) = 0$.*

Proof. The irreducibility of f_α is clear since otherwise α were a zero of a polynomial in $\mathbb{Q}[X]$ of degree strictly smaller than that of f_α . Let $g \in \mathbb{Q}[X]$ have $g(\alpha) = 0$. By division with remainder, we have $g = qf_\alpha + r$ for certain $q, r \in \mathbb{Q}[X]$ with $r = 0$ or $\deg r < \deg f_\alpha$. But $r(\alpha) = 0$, so $r \neq 0$ is impossible. \square

Let $\alpha \in \overline{\mathbb{Q}}$. The *degree* of α , notation $\deg \alpha$, is by definition the degree of f_α . We can factorize f_α over $\overline{\mathbb{Q}}$ as

$$(X - \alpha^{(1)}) \cdots (X - \alpha^{(d)}).$$

These $\alpha^{(1)}, \dots, \alpha^{(d)}$ are called the *conjugates* of α . They are necessarily distinct, for otherwise, f_α and its derivative f'_α would have a root in common, which would then be a root of the greatest common divisor of f_α and f'_α . But this is impossible because since f_α is irreducible, this gcd is 1.

An algebraic number α is called *totally real* if all $\alpha^{(i)}$ are in \mathbb{R} , and *totally complex*, if all $\alpha^{(i)}$ are in $\mathbb{C} \setminus \mathbb{R}$. In general, some of the $\alpha^{(i)}$ may be in \mathbb{R} and some not in

\mathbb{R} . For instance, $\alpha = \sqrt[3]{2}$ has minimal polynomial $X^3 - 2$, and conjugates $\sqrt[3]{2}$, $\zeta\sqrt[3]{2}$, and $\zeta^2\sqrt[3]{2}$, where ζ is a primitive cube root of unity.

If $\alpha^{(i)}$ is a non-real conjugate of an algebraic number α , then so is its complex conjugate $\overline{\alpha^{(i)}}$. Hence the non-real conjugates of an algebraic number occur in pairs of complex conjugates, and so their number is even.

Suppose $f_\alpha = X^d + b_1X^{d-1} + \cdots + b_{d-1}X + b_d$; then $b_1, \dots, b_d \in \mathbb{Q}$. Let $a_0 \in \mathbb{Z}_{>0}$ be the least common multiple of the denominators of b_1, \dots, b_d . Put $F_\alpha := a_0f_\alpha$. Then

$$F_\alpha = a_0X^d + a_1X^{d-1} + \cdots + a_d, \quad \text{with } a_0, \dots, a_d \in \mathbb{Z}, \quad \gcd(a_0, \dots, a_d) = 1.$$

We call F_α the *primitive minimal polynomial* of α (terminology invented by the author; not used in general!). We define the *height* of α by

$$H(\alpha) := \max(|a_0|, \dots, |a_d|).$$

Examples. 1. Let $\alpha = a/b$ with $a, b \in \mathbb{Z}$, $b > 0$, $\gcd(a, b) = 1$. Then α has primitive minimal polynomial $bX - a$, hence $H(\alpha) = \max(|a|, b)$.

2. Let $\alpha = \frac{1}{5}(1 + 2\sqrt{3})$. Then

$$\begin{aligned} f_\alpha &= (X - \frac{1}{5}(1 + 2\sqrt{3}))(X - \frac{1}{5}(1 - 2\sqrt{3})) = X^2 - \frac{2}{5}X - \frac{11}{25}, \\ F_\alpha &= 25X^2 - 10X - 11, \quad H(\alpha) = 25. \end{aligned}$$

Exercise 3.1. Let $\alpha \in \overline{\mathbb{Q}}$ be non-zero.

(i) Prove that $H(\alpha^{-1}) = H(\alpha)$.

(ii) Prove that $(H(\alpha) + 1)^{-1} \leq |\alpha| \leq H(\alpha) + 1$.

Definition. A number $\alpha \in \mathbb{C}$ is called an *algebraic integer* if there is a *monic* polynomial $f \in \mathbb{Z}[X]$ such that $f(\alpha) = 0$. Elements of \mathbb{Z} are often called *rational integers*. A number $\alpha \in \mathbb{C}$ is called an *algebraic unit* if both α, α^{-1} are algebraic integers.

We define

$$\overline{\mathbb{Z}} := \{\alpha \in \mathbb{C} : \alpha \text{ algebraic integer}\}, \quad \overline{\mathbb{Z}}^* := \{\alpha \in \mathbb{C} : \alpha \text{ algebraic unit}\}.$$

Lemma 3.3. (i) $\overline{\mathbb{Z}}$ is a ring, i.e., sums, differences and products of algebraic integers are again algebraic integers. Further, $\overline{\mathbb{Z}}^*$ is the multiplicative group of units

(invertible elements) of $\overline{\mathbb{Z}}$.

(ii) $\overline{\mathbb{Z}}$ is integrally closed, i.e., if $g = X^n + \beta_1 X^{n-1} \dots + \beta_n \in \overline{\mathbb{Z}}[X]$ and $\alpha \in \mathbb{C}$ is a zero of g , then $\alpha \in \overline{\mathbb{Z}}$.

Proof. Apply Proposition 3.26 in Section 3.5 with $A = \mathbb{Z}$, $B = \mathbb{C}$. □

Lemma 3.4. Let $\alpha \in \overline{\mathbb{Q}}$.

(i) $\alpha \in \overline{\mathbb{Z}} \iff f_\alpha \in \mathbb{Z}[X]$.

(ii) $\alpha \in \overline{\mathbb{Z}}^* \iff f_\alpha \in \mathbb{Z}[X]$ and $f_\alpha(0) = \pm 1$.

Proof. (i) \Leftarrow is clear. To prove \Rightarrow , take a monic polynomial $g \in \mathbb{Z}[X]$ with $g(\alpha) = 0$. The polynomial f_α divides g in $\mathbb{Q}[X]$. By factorizing g in $\mathbb{Z}[X]$, we get $g = h_1 h_2$ where $h_1, h_2 \in \mathbb{Z}[X]$ and where h_1 is a constant multiple of f_α . But the leading coefficient of h_1 divides that of g , hence is ± 1 , and so $f_\alpha = \pm h_1 \in \mathbb{Z}[X]$.

(ii) \Leftarrow is again clear. To prove \Rightarrow , we already know that both $f_\alpha \in \mathbb{Z}[X]$ and $f_{\alpha^{-1}} \in \mathbb{Z}[X]$. Further, $f_{\alpha^{-1}}(0) = f_\alpha(0)^{-1} \in \mathbb{Z}$, hence $f_\alpha(0) = \pm 1$. □

Lemma 3.5. $\overline{\mathbb{Z}} \cap \mathbb{Q} = \mathbb{Z}$.

Proof. Let $\alpha \in \overline{\mathbb{Z}} \cap \mathbb{Q}$. Then $f_\alpha = X - \alpha \in \mathbb{Z}[X]$, hence $\alpha \in \mathbb{Z}$. □

An important observation in many Diophantine approximation proofs is that if a is a non-zero rational integer (i.e., in \mathbb{Z}) then $|a| \geq 1$. This cannot be extended to algebraic integers. For instance, $\alpha := \frac{1}{2}(1 - \sqrt{5})$ is a non-zero algebraic integer (being a root of $X^2 - X - 1$) but $|\alpha| < 1$. But in many cases we can use the next lemma instead.

Lemma 3.6. Let α be a non-zero algebraic integer. Then α has a conjugate $\alpha^{(i)}$ with $|\alpha^{(i)}| \geq 1$.

Proof. Let $\alpha^{(1)}, \dots, \alpha^{(d)}$ be the conjugates of α . Then by Lemma 3.4, the minimal polynomial of α , $f_\alpha = \prod_{i=1}^d (X - \alpha^{(i)})$ has its coefficients in \mathbb{Z} . In particular, the product $\alpha^{(1)} \dots \alpha^{(d)} = \pm f_\alpha(0)$ is a non-zero rational integer, whence has absolute value at least 1. This implies the lemma. □

Lemma 3.7. Let α be an algebraic number, and let $F = a_0 X^d + a_1 X^{d-1} + \dots + a_d \in \mathbb{Z}[X]$ be a polynomial with $F(\alpha) = 0$. Then $a_0 \alpha$ is an algebraic integer.

Proof. We clearly have

$$0 = a_0^{d-1}F(\alpha) = (a_0\alpha)^d + a_1(a_0\alpha)^{d-1} + \cdots + a_0^{d-1}a_d,$$

hence $a_0\alpha$ is a zero of a monic polynomial from $\mathbb{Z}[X]$. □

Definition. Given a non-zero algebraic number α , we define the *denominator* of α to be the smallest positive $a \in \mathbb{Z}$ such that $a\alpha$ is an algebraic integer, notation $\text{den}(\alpha)$.

Exercise 3.2. (i) Let $G = b_0X^m + b_1X^{m-1} + \cdots + b_m$ where b_0, \dots, b_m are algebraic integers with $b_0 \neq 0$. Let $\alpha \in \mathbb{C}$ with $G(\alpha) = 0$. Prove that $G/(X - \alpha)$ is a polynomial whose coefficients are algebraic integers.

Hint. Induction on m . In the induction step, use that $b_0\alpha$ is an algebraic integer, after having showed this. Use Lemma 3.3.

(ii) We can express the primitive minimal polynomial of an algebraic number α as $F_\alpha = a_0 \prod_{i=1}^d (X - \alpha^{(i)})$, where $\alpha^{(1)}, \dots, \alpha^{(d)}$ are the conjugates of α . Prove that for each subset $\{i_1, \dots, i_k\}$ of $\{1, \dots, d\}$, the number $a_0\alpha^{(i_1)} \cdots \alpha^{(i_k)}$ is an algebraic integer.

3.2 Algebraic number fields

We first recall a few generalities from field theory. We call $K \supset k$ a field extension, or K an extension of k , if k is a subfield of K , that is, k is a field with the addition and multiplication coming from K . Note that in this case, K is a k -vector space, since it is closed under addition and under scalar multiplication with elements from k (but of course K has much more structure).

Definition. A field extension $K \supset k$ is called *finite* (or K is a finite extension of k) if K is finite dimensional as a k -vector space. In this case, the *degree* of $K \supset k$, notation $[K : k]$, is defined to be the dimension of K as a k -vector space.

Examples. 1. $\mathbb{C} = \{a + bi : a, b \in \mathbb{R}\}$. So $\mathbb{C} \supset \mathbb{R}$ is finite, and $[\mathbb{C} : \mathbb{R}] = 2$.

2. $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$. Verify that this is a field, in particular that it is closed under division. Clearly, $\mathbb{Q}(\sqrt{2}) \supset \mathbb{Q}$ is finite, and $[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2$.

Definition. A(n algebraic) number field is a finite extension of \mathbb{Q} .

Lemma 3.8. *Let $L \supset K \supset k$ be a tower of field extensions (i.e., K is a subfield of L , and k of K). Then $L \supset k$ is finite if and only if $L \supset K$ and $K \supset k$ are finite, and in this case, $[L : k] = [L : K] \cdot [K : k]$.*

Proof. First assume that $L \supset k$ is finite. Then certainly $K \supset k$ is finite since K is a k -linear subspace of L . Further, a k -basis of L also generates L as a k -vector space. Hence $L \supset K$ is finite as well. Conversely, suppose that $K \supset k$ is finite and let $\{\alpha_1, \dots, \alpha_r\}$ a k -basis of K , and suppose that $L \supset K$ is finite and let $\{\beta_1, \dots, \beta_s\}$ be a K -basis of L . Then $\{\alpha_i \beta_j : i = 1, \dots, r, j = 1, \dots, s\}$ is a k -basis of L . This proves our lemma. \square

Let $K \supset k$ be a field extension, and $\alpha_1, \dots, \alpha_r \in K$. Then $k(\alpha_1, \dots, \alpha_r)$ denotes the smallest subfield of K containing both k and $\alpha_1, \dots, \alpha_r$. Thus, $k(\alpha_1, \dots, \alpha_r)$ consists of all entities $f(\alpha_1, \dots, \alpha_r)/g(\alpha_1, \dots, \alpha_r)$, where $f, g \in k[X_1, \dots, X_r]$, and $g(\alpha_1, \dots, \alpha_r) \neq 0$. An extension of the type $k(\alpha) \supset k$ is called *primitive*.

Let $K \supset k$ be an extension and $\alpha \in K$. We say that α is algebraic over k if there is a non-zero polynomial $g \in k[X]$ with $g(\alpha) = 0$. The necessarily unique, monic polynomial of minimal degree with this property is called the *minimal polynomial of α over k* , notation $f_{\alpha, k}$. The degree of α over k is the degree of $f_{\alpha, k}$. The polynomial $f_{\alpha, k}$ is necessarily irreducible in $k[X]$.

Lemma 3.9. *Suppose α has degree d over k . Then $k(\alpha)$ is a finite extension of k with basis $\{1, \alpha, \dots, \alpha^{d-1}\}$ over k . Hence $[k(\alpha) : k] = d$.*

Proof. $\{1, \alpha, \dots, \alpha^{d-1}\}$ is certainly linearly independent over k since any non-trivial k -linear combination of these elements would yield a polynomial expression in α of degree $< d$ which is necessarily non-zero. Let $V := \{g(\alpha) : g \in k[X]\}$. This is clearly a k -vector space. Using division with remainder, we can write $g \in k[X]$ as $qf_{\alpha, k} + r$ with $q, r \in k[X]$ and $r = 0$ or $\deg r < d$. Now $g(\alpha) = r(\alpha)$ is a k -linear combination of $1, \alpha, \dots, \alpha^{d-1}$, so these elements form a k -basis of V . To show that V is a field, it remains to show that it is closed under multiplicative inversion. Let $g(\alpha) \in V$ where $g \in k[X]$ is non-zero and of degree $< d$. Since $f_{\alpha, k}$ is irreducible in $k[X]$ it is coprime with g , implying that there are $a, b \in k[X]$ with $ag + bf_{\alpha, k} = 1$. Hence $a(\alpha)g(\alpha) = 1$. This shows that $V = k(\alpha)$. \square

We now specialize to algebraic number fields. Note that by Lemmas 3.8 and 3.9, if $\alpha_1, \dots, \alpha_r$ are algebraic numbers then $\mathbb{Q}(\alpha_1, \dots, \alpha_r)$ is an algebraic number

field. Conversely, any algebraic number field K can be expressed in this form, for instance by taking a \mathbb{Q} -basis of K . The following result, which we state without proof, asserts that a number field can always be generated by a single element.

Theorem 3.10 (Theorem of the primitive element). *Let K be an algebraic number field of degree d . Then there is $\theta \in K$ such that $K = \mathbb{Q}(\theta)$.*

A consequence is that K is a \mathbb{Q} -vector space with basis $\{1, \theta, \dots, \theta^{d-1}\}$.

Example. $\mathbb{Q}(\sqrt{2}, \sqrt{3}) = \mathbb{Q}(\sqrt{2} + \sqrt{3})$.

To verify this, observe first that $\sqrt{3} \notin \mathbb{Q}(\sqrt{2})$. Hence $K := \mathbb{Q}(\sqrt{2}, \sqrt{3})$ has degree $[K : \mathbb{Q}(\sqrt{2})] \cdot [\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 4$. For a \mathbb{Q} -basis of K one may take $\{1, \sqrt{2}, \sqrt{3}, \sqrt{6}\}$. On the other hand, $L := \mathbb{Q}(\sqrt{2} + \sqrt{3})$ is a subfield of K , and it has the four \mathbb{Q} -linearly independent elements $1, \sqrt{2} + \sqrt{3}, (\sqrt{2} + \sqrt{3})^2 = 5 + 2\sqrt{6}, (\sqrt{2} + \sqrt{3})^3 = 11\sqrt{2} + 9\sqrt{3}$. Hence $L = K$.

An *embedding* of a number field K in \mathbb{C} is an injective field homomorphism of K into \mathbb{C} . An embedding of K in \mathbb{C} necessarily leaves the elements of \mathbb{Q} unchanged. This has the following consequences.

First, let σ be an embedding of K in \mathbb{C} , $\alpha_1, \dots, \alpha_r \in K$, and

$$\beta = f(\alpha_1, \dots, \alpha_r) / g(\alpha_1, \dots, \alpha_r) \text{ with } f, g \in \mathbb{Q}[X_1, \dots, X_r].$$

Then $\sigma(\beta) = f(\sigma(\alpha_1), \dots, \sigma(\alpha_r)) / g(\sigma(\alpha_1), \dots, \sigma(\alpha_r))$. So if $K = \mathbb{Q}(\alpha_1, \dots, \alpha_r)$, then σ is uniquely determined by its images on $\alpha_1, \dots, \alpha_r$.

Second, if $f \in \mathbb{Q}[X]$, and $\alpha \in K$ is a zero of f , then also $\sigma(\alpha)$ is a zero of f . For $f(\sigma(\alpha)) = \sigma(f(\alpha)) = 0$.

Proposition 3.11. *Let K be an algebraic number field of degree d . Then there are precisely d distinct embeddings of K in \mathbb{C} .*

These d embeddings can be described explicitly as follows. Suppose that $K = \mathbb{Q}(\theta)$. Let f_θ be the minimal polynomial of θ (over \mathbb{Q}) and suppose f_θ has degree d . Then

$$f_\theta = (X - \theta^{(1)}) \cdots (X - \theta^{(d)}) \text{ with } \theta^{(1)}, \dots, \theta^{(d)} \in \mathbb{C}.$$

As mentioned above, the embeddings of K in \mathbb{C} map θ to the zeros of f_θ , and each embedding is uniquely determined by its image of θ . Hence the d embeddings $\sigma_1, \dots, \sigma_d$ of K in \mathbb{C} can be given by $\sigma_i(\theta) = \theta^{(i)}$.

An embedding σ of K in \mathbb{C} is called *real* if $\sigma(K) \subset \mathbb{R}$, and *complex* if $\sigma(K) \not\subset \mathbb{R}$. If σ is a complex embedding of K , then so is its composition with complex conjugation, $\bar{\sigma} : x \mapsto \overline{\sigma(x)}$. Thus the complex embeddings of K occur in conjugate pairs $\sigma, \bar{\sigma}$. Suppose that K has precisely r_1 real embeddings, and r_2 pairs of conjugate complex embeddings. Then $r_1 + 2r_2 = d$. It will often be convenient to order the embeddings $\sigma_1, \dots, \sigma_d$ in such a way that

- $\sigma_1, \dots, \sigma_{r_1}$ are the real embeddings;
- $\{\sigma_{r_1+1}, \sigma_{r_1+r_2+1} = \overline{\sigma_{r_1+1}}\}, \dots, \{\sigma_{r_1+r_2}, \sigma_{r_1+2r_2} = \overline{\sigma_{r_1+r_2}}\}$ are the pairs of conjugate complex embeddings.

Example. Let $K = \mathbb{Q}(\sqrt[4]{2})$. The minimal polynomial of $\sqrt[4]{2}$ is $X^4 - 2$, and the zeros of $X^4 - 2$ are $i^k \sqrt[4]{2}$ ($k = 0, 1, 2, 3$), where $i^2 = -1$. Thus,

$$K = \left\{ \sum_{j=0}^3 x_j (\sqrt[4]{2})^j : x_j \in \mathbb{Q} \right\},$$

the four embeddings of K in \mathbb{C} are given by

$$\sum_{j=0}^3 x_j (\sqrt[4]{2})^j \mapsto \sum_{j=0}^3 x_j (i^k \sqrt[4]{2})^j \quad (k = 0, 1, 2, 3),$$

and σ_0, σ_2 are real, σ_1, σ_3 are complex, and $\sigma_3 = \bar{\sigma}_1$. So $r_1 = 2, r_2 = 1$.

Let K be an algebraic number field, and L a finite extension of K . Further, let σ, τ be embeddings of respectively K and L in \mathbb{C} . Then τ of L is called a *continuation* of σ if $\tau|_K = \sigma$, i.e., $\sigma(x) = \tau(x)$ for $x \in K$. Obviously, each embedding of L in \mathbb{C} is a continuation of some embedding of K in \mathbb{C} . We state without proof the following:

Proposition 3.12. *Each embedding of K in \mathbb{C} can be continued in precisely $[L : K]$ ways to an embedding of L in \mathbb{C} .*

Example. Let $K = \mathbb{Q}(\sqrt[4]{2}), L = \mathbb{Q}(\sqrt[12]{2})$. Then $L \supset K$ and $[L : K] = 3$. The four embeddings of K in \mathbb{C} are given by $\sigma_k(\sqrt[4]{2}) = i^k \sqrt[4]{2}$ ($k = 0, 1, 2, 3$). The twelve embeddings of L in \mathbb{C} are given by $\tau_l(\sqrt[12]{2}) = \zeta_{12}^l \sqrt[12]{2}$ ($l = 0, 1, \dots, 11$), where $\zeta_{12}^3 = i$. The continuations of σ_k to L are τ_l with $3l \equiv k \pmod{4}, 0 \leq l \leq 11$.

Corollary 3.13. *Let K be an algebraic number field of degree d , and let $\sigma_1, \dots, \sigma_d$ be the embeddings of K in \mathbb{C} . Further, let $\alpha \in K$. Then each of the conjugates of α occurs precisely $[K : \mathbb{Q}(\alpha)]$ times in the sequence $\sigma_1(\alpha), \dots, \sigma_d(\alpha)$.*

Proof. Suppose $\mathbb{Q}(\alpha)$ has degree m . Let $f_\alpha = \prod_{i=1}^m (X - \alpha^{(i)})$. Then the m embeddings τ_1, \dots, τ_m of $\mathbb{Q}(\alpha)$ in \mathbb{C} are determined by $\tau_i(\alpha) = \alpha^{(i)}$ for $i = 1, \dots, m$. Since each τ_i has precisely d/m continuations to K and each embedding of K in \mathbb{C} is a continuation of some τ_i , each of the numbers $\alpha^{(i)}$ occurs precisely $d/m = [K : \mathbb{Q}(\alpha)]$ times among $\sigma_1(\alpha), \dots, \sigma_m(\alpha)$. \square

Let K be an algebraic number field of degree d and $\sigma_1, \dots, \sigma_d$ the embeddings of K in \mathbb{C} . The *characteristic polynomial* of $\alpha \in K$ is defined by

$$\chi_{\alpha, K} := \prod_{i=1}^d (X - \sigma_i(\alpha)).$$

In case that $K = \mathbb{Q}(\alpha)$, we have $\chi_{\alpha, K} = f_\alpha$ is the minimal polynomial of α . In case that $\mathbb{Q}(\alpha)$ is strictly smaller than K we have the following.

Lemma 3.14. *Let $\alpha \in K$ and let f_α be the minimal polynomial of α . Then $\chi_{\alpha, K} = f_\alpha^{[K:\mathbb{Q}(\alpha)]}$. Hence $\chi_{\alpha, K} \in \mathbb{Q}[X]$.*

Proof. Let $m := [\mathbb{Q}(\alpha) : \mathbb{Q}]$. Then $f_\alpha = (X - \alpha^{(1)}) \cdots (X - \alpha^{(m)})$ where $\alpha^{(1)}, \dots, \alpha^{(m)}$ are the conjugates of α . Apply Corollary 3.13. \square

Definition. Let K be an algebraic number field. The *ring of integers* of K is given by

$$O_K := \{\alpha \in K : \alpha \text{ is integral over } \mathbb{Z}\} = K \cap \overline{\mathbb{Z}}.$$

That is, O_K is the integral closure of \mathbb{Z} in K . The group of units (invertible elements) of O_K is denoted by O_K^* . We observe that if $\alpha \in O_K$ and σ is an embedding of K in \mathbb{C} , then $\sigma(\alpha)$ is an algebraic integer. For α is a zero of a monic $f \in \mathbb{Z}[X]$ hence so is $\sigma(\alpha)$.

Lemma 3.15. *Let $\alpha \in K$. Then $\alpha \in O_K \iff \chi_{\alpha, K} \in \mathbb{Z}[X]$.*

Proof. \Leftarrow $\chi_{\alpha, K}$ is a monic polynomial having α as a root.

\Rightarrow Lemma 3.4 implies $f_\alpha \in \mathbb{Z}[X]$, and then Lemma 3.14 implies $\chi_{\alpha, K} \in \mathbb{Z}[X]$. \square

Definition. Let K be an algebraic number field of degree d and $\sigma_1, \dots, \sigma_d$ the embeddings of K in \mathbb{C} . Then the *trace* and *norm* of $\alpha \in K$ over \mathbb{Q} are given by respectively

$$\text{Tr}_{K/\mathbb{Q}}(\alpha) = \sum_{i=1}^d \sigma_i(\alpha), \quad N_{K/\mathbb{Q}}(\alpha) = \prod_{i=1}^d \sigma_i(\alpha).$$

These numbers are coefficients of $\chi_{\alpha, K}$. So by Lemma 3.14 these numbers belong to \mathbb{Q} ; moreover, if $\alpha \in O_K$ then by Lemma 3.15 they belong to \mathbb{Z} . Notice that for $\alpha, \beta \in K$ we have

$$\text{Tr}_{K/\mathbb{Q}}(\alpha + \beta) = \text{Tr}_{K/\mathbb{Q}}(\alpha) + \text{Tr}_{K/\mathbb{Q}}(\beta), \quad N_{K/\mathbb{Q}}(\alpha\beta) = N_{K/\mathbb{Q}}(\alpha)N_{K/\mathbb{Q}}(\beta).$$

Lemma 3.16. *Let $\alpha \in O_K$. Then $\alpha \in O_K^* \iff N_{K/\mathbb{Q}}(\alpha) = \pm 1$.*

Proof. \implies Both α, α^{-1} are in O_K , hence $N_{K/\mathbb{Q}}(\alpha) \in \mathbb{Z}$, $N_{K/\mathbb{Q}}(\alpha^{-1}) \in \mathbb{Z}$. But $N_{K/\mathbb{Q}}(\alpha)N_{K/\mathbb{Q}}(\alpha^{-1}) = 1$, hence $N_{K/\mathbb{Q}}(\alpha) = \pm 1$.

\impliedby Without loss of generality, $\sigma_1(\alpha) = \alpha$. Then $\alpha^{-1} = \pm \sigma_2(\alpha) \cdots \sigma_d(\alpha)$ is an algebraic integer, hence belongs to O_K . \square

Exercise 3.3. *Let K be a quadratic number field, that is an algebraic number field of degree 2.*

(i) *Prove that $K = \mathbb{Q}(\sqrt{d}) = \{a + b\sqrt{d} : a, b \in \mathbb{Q}\}$, where $d \in \mathbb{Z} \setminus \{0, 1\}$ and d is not divisible by the square of an integer $\neq 1$. Also determine the two embeddings of K in \mathbb{C} .*

(ii) *Let $\alpha = a + b\sqrt{d}$ with $a, b \in \mathbb{Q}$. Prove the following:*

if $d \equiv 2 \pmod{4}$ or $d \equiv 3 \pmod{4}$ then $\alpha \in O_K$ if and only if $a, b \in \mathbb{Z}$;

if $d \equiv 1 \pmod{4}$ then $\alpha \in O_K$ if and only if $2a, 2b \in \mathbb{Z}$ and $2a \equiv 2b \pmod{2}$.

Hint. *Determine the minimal polynomial f_α of α and use that its coefficients are in \mathbb{Z} .*

3.3 Galois theory

Let K be an algebraic number field with $K \subset \mathbb{C}$. The field K is called *normal* if $K = \mathbb{Q}(\alpha_1, \dots, \alpha_r)$ where $\alpha_1, \dots, \alpha_r$ are such that $(X - \alpha_1) \cdots (X - \alpha_r) \in \mathbb{Q}[X]$.

Let $K \subset \mathbb{C}$ be a normal algebraic number field and $\alpha_1, \dots, \alpha_r$ as above. Let σ be an embedding of K in \mathbb{C} . In general, if $f \in \mathbb{Q}[X]$ and α is a zero of f in K , then

$f(\sigma(\alpha)) = 0$. As a consequence, σ permutes $\alpha_1, \dots, \alpha_r$. Since K consists of rational functions in $\alpha_1, \dots, \alpha_r$ with coefficients in \mathbb{Q} , this implies that σ is an isomorphism mapping K to itself, i.e., an *automorphism* of K .

The automorphisms of K form a group under composition, the *Galois group* of K , notation $\text{Gal}(K/\mathbb{Q})$. We state without proof some properties of the Galois group.

Proposition 3.17. *Let K be a normal algebraic number field, and $\text{Gal}(K/\mathbb{Q})$ its Galois group.*

(i) $\text{Gal}(K/\mathbb{Q})$ is a group of order $[K : \mathbb{Q}]$. We have

$$\{x \in K : \sigma(x) = x \ \forall \sigma \in \text{Gal}(K/\mathbb{Q})\} = \mathbb{Q}.$$

(ii) There is a bijection between the subgroups of $\text{Gal}(K/\mathbb{Q})$ and the subfields of K , given by

$$\begin{aligned} H \leq \text{Gal}(K/\mathbb{Q}) &\longrightarrow K^H := \{x \in K : \sigma(x) = x \ \forall \sigma \in H\} \\ \text{Gal}(K/L) := \{\sigma \in \text{Gal}(K/\mathbb{Q}) : \sigma|_L = \text{id}\} &\longleftarrow L \end{aligned}$$

and the order of H is equal to $[K : K^H]$.

(iii) Let $f \in \mathbb{Q}[X]$ be an irreducible polynomial having at least one root in K . Then all roots of f lie in K , for any two roots α, β of f there is $\sigma \in \text{Gal}(K/\mathbb{Q})$ such that $\sigma(\alpha) = \beta$, and each $\sigma \in \text{Gal}(K/\mathbb{Q})$ permutes the roots of f .

Remarks. 1. The bijection in (ii) reverses inclusions: if H_1 is a subgroup of H_2 , then K^{H_2} is a subfield of K^{H_1} .

2. Every algebraic number field K can be extended to a normal number field. Let $K = \mathbb{Q}(\theta)$. Then for the minimal polynomial f_θ of θ we have $f_\theta = (X - \theta^{(1)}) \cdots (X - \theta^{(d)})$. Clearly, $N := \mathbb{Q}(\theta^{(1)}, \dots, \theta^{(d)})$ is normal, and N contains K .

Example. Let $K = \mathbb{Q}(\sqrt{2}, \sqrt{3})$. Then $K = \mathbb{Q}(\sqrt{2}, -\sqrt{2}, \sqrt{3}, -\sqrt{3})$ is generated by the four roots of $(X^2 - 2)(X^2 - 3)$ hence it is normal. We have seen before that $[K : \mathbb{Q}] = 4$, and that K has basis $\{1, \sqrt{2}, \sqrt{3}, \sqrt{6}\}$ over \mathbb{Q} . Hence $G := \text{Gal}(K/\mathbb{Q})$ has order 4. Any $\sigma \in \text{Gal}(K/\mathbb{Q})$ maps $\sqrt{2}$ to a root of $X^2 - 2$, hence to $\pm\sqrt{2}$. Likewise, σ maps $\sqrt{3}$ to $\pm\sqrt{3}$. Thus, $G = \{\sigma_{ab} : a, b \in \{1, -1\}\}$, where $\sigma_{ab}(\sqrt{2}) = a\sqrt{2}$, $\sigma_{ab}(\sqrt{3}) = b\sqrt{3}$. It is easy to check that G is the Klein four group, with σ_{11} the identity. The table below gives the subgroups of G and corresponding subfields of K .

H	$K^H = \{x \in K : \sigma(x) = x \ \forall \sigma \in H\}$
$\{\text{id}\}$	$\mathbb{Q}(\sqrt{2}, \sqrt{3})$
$\{\text{id}, \sigma_{1,-1}\}$	$\mathbb{Q}(\sqrt{2})$
$\{\text{id}, \sigma_{-1,1}\}$	$\mathbb{Q}(\sqrt{3})$
$\{\text{id}, \sigma_{-1,-1}\}$	$\mathbb{Q}(\sqrt{6})$
G	\mathbb{Q}

As an example, we compute the subfield K^H corresponding to the subgroup $H = \{\text{id}, \sigma_{-1,-1}\}$ of G . Recall that $\{1, \sqrt{2}, \sqrt{3}, \sqrt{6}\}$ is a basis of K . Thus, every element of K can be expressed uniquely as $x_0 + x_1\sqrt{2} + x_2\sqrt{3} + x_3\sqrt{6}$ with $x_i \in \mathbb{Q}$. Now $\sigma_{-1,-1}$ maps $\beta = x_0 + x_1\sqrt{2} + x_2\sqrt{3} + x_3\sqrt{6}$ to $x_0 - x_1\sqrt{2} - x_2\sqrt{3} + x_3\sqrt{6}$, and thus $\sigma_{11}(\beta) = \beta$ if and only if $x_1 = x_2 = 0$. This shows that $K^H = \mathbb{Q}(\sqrt{6})$.

Exercise 3.4. Let $\sqrt[3]{2}$ be the real cube root of 2, and ζ a primitive cube root of unity.
(i) Prove that the field $K := \mathbb{Q}(\sqrt[3]{2}, \zeta)$ is normal and that $[K : \mathbb{Q}] = 6$.
(ii) Determine the subfields of K .

3.4 Siegel's Lemma

We develop a tool which is very important in Diophantine approximation, the so-called *Siegel's Lemma*, which was formally stated for the first time by Siegel in 1929, but was known before. Essentially, it states that under certain hypotheses, a system of M homogeneous linear equations in N unknowns

$$(3.1) \quad \begin{cases} a_{11}x_1 + \cdots + a_{1N}x_N = 0 \\ \vdots \\ a_{M1}x_1 + \cdots + a_{MN}x_N = 0 \end{cases}$$

has a non-trivial solution in integer coordinates, the absolute values of which are not too large.

Theorem 3.18 (Siegel's Lemma). *Assume that $N > M > 0$, $A \geq 1$, and*

$$a_{ij} \in \mathbb{Z}, \quad |a_{ij}| \leq A \quad \text{for } i = 1, \dots, M, \ j = 1, \dots, N.$$

Then (3.1) has a solution $\mathbf{x} = (x_1, \dots, x_N) \in \mathbb{Z}^N \setminus \{\mathbf{0}\}$ with

$$\max_{1 \leq i \leq N} |x_i| \leq (NA)^{M/(N-M)}.$$

Proof. For $i = 1, \dots, M$, $\mathbf{x} \in \mathbb{Z}^N$, put $l_i(\mathbf{x}) := \sum_{j=1}^N a_{ij}x_j$ and let

$$-C_i := \sum_{j=1}^N \min(a_{ij}, 0), \quad D_i := \sum_{j=1}^N \max(a_{ij}, 0).$$

Notice that $C_i + D_i \leq NA$. Let B be a positive integer, and let $S_B := \{0, \dots, B\}^N$. For each $\mathbf{y} \in S_B$ we have

$$-C_i B \leq l_i(\mathbf{y}) \leq D_i B \quad \text{for } i = 1, \dots, M.$$

Notice that S_B has cardinality $(B + 1)^N$. Further, if \mathbf{y} runs through S_B , then $(l_1(\mathbf{y}), \dots, l_M(\mathbf{y}))$ runs through a set of cardinality at most

$$\prod_{i=1}^M (C_i B + D_i B + 1) \leq (NAB + 1)^M.$$

We choose B such that $(B + 1)^N > (NAB + 1)^M$. Then by the box principle, there are distinct $\mathbf{y}_1, \mathbf{y}_2 \in S_B$ with $l_i(\mathbf{y}_1) = l_i(\mathbf{y}_2)$ for $i = 1, \dots, M$. Take $\mathbf{x} = \mathbf{y}_1 - \mathbf{y}_2$. Then \mathbf{x} satisfies (3.1) and $|x_i| \leq B$ for $i = 1, \dots, N$.

We finish our proof by showing that the choice $B = \lceil (NA)^{M/(N-M)} \rceil$ is valid. Indeed, with this choice of B we have $(B + 1)^{N-M} > (NA)^M$, hence

$$(B + 1)^N > (NA(B + 1))^M > (NAB + 1)^M.$$

□

We often need a generalization where the coefficients a_{ij} are algebraic integers instead of just rational integers. To deduce this, we need some preparations.

Let K be an algebraic number field of degree d . Denote as usual by O_K its ring of integers. Assume K has r_1 real embeddings, and r_2 pairs of conjugate complex embeddings, so that $r_1 + 2r_2 = d$. We order the embeddings of K in \mathbb{C} in such a way that $\sigma_1, \dots, \sigma_{r_1}$ are the real embeddings, and $\{\sigma_{r_1+1}, \sigma_{r_1+r_2+1} = \overline{\sigma_{r_1+1}}\}, \dots, \{\sigma_{r_1+r_2}, \sigma_{r_1+2r_2} = \overline{\sigma_{r_1+r_2}}\}$ are the pairs of conjugate complex embeddings. Define the map

$$\begin{aligned} \varphi : K &\rightarrow \mathbb{R}^d : \\ x &\mapsto (\sigma_1(x), \dots, \sigma_{r_1}(x), \operatorname{Re} \sigma_{r_1+1}(x), \operatorname{Im} \sigma_{r_1+1}(x), \dots, \operatorname{Re} \sigma_{r_1+r_2}(x), \operatorname{Im} \sigma_{r_1+r_2}(x)). \end{aligned}$$

The *house* of an algebraic number α is defined by

$$\overline{|\alpha|} := \max(|\alpha^{(1)}|, \dots, |\alpha^{(m)}|),$$

where $\alpha^{(1)}, \dots, \alpha^{(m)}$ denote the conjugates of α . We have

$$(3.2) \quad \overline{|\alpha|} = \max(|\sigma_1(\alpha)|, \dots, |\sigma_d(\alpha)|) \quad \text{for } \alpha \in K,$$

for by Corollary 3.13, the sequence $\sigma_1(\alpha), \dots, \sigma_d(\alpha)$ consists of the conjugates of α , each repeated exactly $[K : \mathbb{Q}(\alpha)] = d/m$ times.

Lastly, we define $\|\mathbf{x}\|_\infty := \max_i |x_i|$ for $\mathbf{x} = (x_1, \dots, x_d) \in \mathbb{R}^d$.

Lemma 3.19. *Let $\alpha \in O_K$ with $\|\varphi(\alpha)\|_\infty \leq \frac{2}{3}$. Then $\alpha = 0$.*

Proof. We prove that $|\sigma_i(\alpha)| < 1$ for $i = 1, \dots, n$. By Lemma 3.6, this implies $\alpha = 0$.

Our assumption implies that $|\sigma_i(\alpha)| \leq \frac{2}{3} < 1$ for $i = 1, \dots, r_1$. Further, for $i = r_1 + 1, \dots, r_1 + r_2$ the real and imaginary parts of $\sigma_i(\alpha)$ have absolute values at most $\frac{2}{3}$. But this is then also true for their complex conjugates, which are $\sigma_i(\alpha)$ for $i = r_1 + r_2 + 1, \dots, r_1 + 2r_2 = d$. Hence $|\sigma_i(\alpha)| \leq \sqrt{(2/3)^2 + (2/3)^2} < 1$ for $i = r_1 + 1, \dots, d$. \square

We consider again systems (3.1), but now the coefficients a_{ij} are from O_K .

Theorem 3.20. *Let $[K : \mathbb{Q}] = d$, let M, N be integers with $N > dM > 0$, let A be a real ≥ 1 , and suppose that*

$$a_{ij} \in O_K, \quad \overline{|a_{ij}|} \leq A \quad \text{for } i = 1, \dots, M, \quad j = 1, \dots, N.$$

Then the system

$$(3.1) \quad \begin{cases} a_{11}x_1 + \dots + a_{1N}x_N = 0 \\ \vdots \\ a_{M1}x_1 + \dots + a_{MN}x_N = 0 \end{cases}$$

has a solution $\mathbf{x} = (x_1, \dots, x_N) \in \mathbb{Z}^N \setminus \{\mathbf{0}\}$ such that

$$(3.3) \quad \max_{1 \leq i \leq N} |x_i| \leq (3NA)^{dM/(N-dM)}.$$

Remark. Notice that the condition $N > M$ from Theorem 3.18 has been replaced by $N > dM$, while there is also an occurrence of d in the exponent on $3NA$. The reason for this is that if we apply $\sigma_1, \dots, \sigma_D$ to the equations in (3.1) we get in fact a system of dM equations in N unknowns.

Proof. Write $l_i(\mathbf{x}) := a_{i1}x_1 + \dots + a_{iN}x_N$ for $i = 1, \dots, M$ and define the linear map $\psi : \mathbb{Z}^N \rightarrow \mathbb{R}^{Md}$ by

$$\psi(\mathbf{x}) := (\varphi(l_1(\mathbf{x})), \dots, \varphi(l_M(\mathbf{x}))).$$

This is well defined since $l_i(\mathbf{x}) \in O_K$ for $i = 1, \dots, M$. Note that for $\mathbf{x} = (x_1, \dots, x_N) \in \mathbb{Z}^N$, and for an embedding σ of K in \mathbb{C} we have $\sigma(l_i(\mathbf{x})) = \sum_{j=1}^N \sigma(a_{ij})x_j$. Thus, we see that the components of $\psi(\mathbf{x})$ are linear expressions $b_1x_1 + \dots + b_Nx_N$, where each coefficient b_j is either $\sigma_k(a_{ij})$ or the real or imaginary part of $\sigma_k(a_{ij})$, for some $k \in \{1, \dots, d\}$, $i \in \{1, \dots, M\}$. By (3.2), all b_j have absolute value at most A .

Let B be a positive integer to be specified later and consider the vectors $\mathbf{y} \in S_B := \{0, \dots, B\}^N$. If $\mathbf{y} \in S$ then all coordinates of $\psi(\mathbf{y})$ have absolute value at most NAB , that is, $\psi(\mathbf{y})$ lies in the cube $[-NAB, NAB]^{Md}$.

We can partition this cube into $([3NAB] + 1)^{Md}$ small cubes of side length $\leq 2/3$. Now suppose that the cardinality of S_B is larger than the number of small cubes, that is,

$$(3.4) \quad (B + 1)^N > ([3NAB] + 1)^{Md}.$$

Then there must be distinct $\mathbf{y}_1, \mathbf{y}_2 \in S_B$ such that $\psi(\mathbf{y}_1), \psi(\mathbf{y}_2)$ lie in the same small cube. Let $\mathbf{x} := \mathbf{y}_1 - \mathbf{y}_2$. Then

$$\|\psi(\mathbf{x})\|_\infty = \|\psi(\mathbf{y}_1) - \psi(\mathbf{y}_2)\|_\infty \leq \frac{2}{3}.$$

This implies $\|\varphi(l_i(\mathbf{x}))\|_\infty \leq \frac{2}{3}$ for $i = 1, \dots, M$. Since $l_i(\mathbf{x}) \in O_K$, by Lemma 3.19 we have $l_i(\mathbf{x}) = 0$ for $i = 1, \dots, M$. Further, $\|\mathbf{x}\|_\infty \leq B$.

We verify that (3.4) is satisfied with $B = [(3NA)^{Md/(N-Md)}]$. Indeed, for this value of B we have

$$(B + 1)^N = (B + 1)^{N-Md}(B + 1)^{Md} > (3NA)^{Md}(B + 1)^{Md} > (3NAB + 1)^{Md}.$$

□

3.5 Appendix: Ring extensions

A ring A is always supposed to have a (necessarily unique) unit element 1 such that $x \cdot 1 = 1 \cdot x = x$ for all $x \in A$. An element $a \in A$ is called a *unit* of A if there is $b \in A$ with $ba = ab = 1$; this necessarily unique element b is denoted by a^{-1} . The units of A form a multiplicative group, the unit group of A , which we denote by A^* . In particular, the unit group of a field K is $K^* = K \setminus \{0\}$.

An *integral domain* is a commutative ring without divisors of 0 , i.e., it does not contain non-zero elements a, b such that $ab = 0$. The quotient field of an integral domain A is denoted by Q_A . The field Q_A consists of all fractions a/b with $a, b \in A$, $b \neq 0$, where two fractions $a/b, c/d$ are identified if $ad = bc$.

Given two rings A, B , when writing $A \subset B$ or $B \supset A$ we always mean that A is a subring of B , i.e., A is a ring with the addition, multiplication and unit element of B . We call $A \subset B$ or $B \supset A$ a ring extension. It is possible to set up a theory for ring extensions similar to that for field extensions. We restrict ourselves to extensions of commutative rings.

A role similar to that of vector spaces in the theory of field extensions is played by *modules* in the theory of ring extensions. In general, if A is a commutative ring, then an A -module is a set M , endowed with an addition $+$: $M \times M \rightarrow M$ and scalar multiplication \cdot : $A \times M \rightarrow M$, which satisfy precisely the same axioms as the addition and scalar multiplication of a vector space, except that the scalars are taken from a ring instead of a field. A \mathbb{Z} -module is simply an abelian group.

Let A be a commutative ring, and M an A -module. We call M' an A -submodule of M if it is closed under the addition and scalar multiplication of M . That is, M' is an A -submodule of M if and only if for all $\alpha, \beta \in M', r, s \in A$ we have $r\alpha + s\beta \in M'$.

We say that M is *finitely generated over A* , if there is a finite set of elements $\alpha_1, \dots, \alpha_r \in M$ such that $M = \{\sum_{i=1}^r x_i \alpha_i : x_i \in A\}$. We call $\{\alpha_1, \dots, \alpha_r\}$ an *A -basis* for M if $\alpha_1, \dots, \alpha_r$ generate M as an A -module, and if they are A -linearly independent, i.e., there is no $(x_1, \dots, x_r) \in A^r \setminus \{\mathbf{0}\}$ with $\sum_{i=1}^r x_i \alpha_i = 0$.

One notable difference between vector spaces and modules is that finitely generated vector spaces always have a basis, whereas finitely generated modules over a ring need not have a basis. An A -module which does have a basis is called *free*.

Example. View the residue class ring $\mathbb{Z}/n\mathbb{Z}$ (with n a positive integer) as a \mathbb{Z} -

module. Write the residue class $a \bmod n$ as \bar{a} . As a \mathbb{Z} -module, $\mathbb{Z}/n\mathbb{Z}$ is generated by $\bar{1}$. But for each $\bar{a} \in \mathbb{Z}/n\mathbb{Z}$ we have $n \cdot \bar{a} = \bar{0}$, Hence each element of $\mathbb{Z}/n\mathbb{Z}$ is linearly dependent over \mathbb{Z} , and so $\mathbb{Z}/n\mathbb{Z}$ does not have a \mathbb{Z} -basis.

Now let $A \subset B$ be an extension of commutative rings. We denote the unit element of A by 1 . Clearly, B may be viewed as an A -module. We call $A \subset B$ or $B \supset A$ a *finite ring extension* and say that B is *finite over A* , if B is finitely generated as an A -module.

Given $\alpha_1, \dots, \alpha_r \in B$, we denote by $A[\alpha_1, \dots, \alpha_r]$ the smallest subring of B containing A and $\alpha_1, \dots, \alpha_r$. Thus,

$$A[\alpha_1, \dots, \alpha_r] = \{f(\alpha_1, \dots, \alpha_r) : f \in A[X_1, \dots, X_r]\}.$$

An element $\alpha \in B$ is said to be *integral over A* if there is a *monic* polynomial $f \in A[X]$ with $f(\alpha) = 0$.

Before proceeding, we recall a version of the division with remainder algorithm for rings.

Lemma 3.21. *Let A be a commutative ring, and $f, g \in A[X]$ polynomials such that the leading coefficient of f is in A^* . Then there are polynomials $q, r \in A[X]$ such that $g = qf + r$, $\deg q \leq \deg g - \deg f$ and $\deg r < \deg f$. The polynomials q, r are uniquely determined by f, g .*

Proof. Induction on the degree of g . First assume that $\deg g < \deg f$. If $g = qf + r$ for some $q, r \in A[X]$ with $q \neq 0$ and $\deg r < \deg f$, then since the leading coefficient of f is in A^* , we have $\deg qf \geq \deg f$, while on the other hand $\deg qf = \deg(g - r) < \deg f$, which is impossible. Hence $q = 0$, $r = g$.

Suppose that $\deg g = m$ and $\deg f = n$ with $m \geq n$. let a, b be the leading coefficients of f, g , respectively. So $a \in A^*$. Then apply the induction hypothesis to $g - ba^{-1}X^{m-n}f$, which is in $A[X]$ and has degree smaller than m . \square

Lemma 3.22. *Let $A \subset B$ be an extension of commutative rings, and $\alpha \in B$. Then the following are equivalent:*

- (i) α is integral over A ;
- (ii) $A[\alpha]$ is finite over A ;

(iii) there is a non-zero, finitely generated A -submodule M of B such that $1 \in M$ and $\alpha M \subseteq M$, where $\alpha M = \{\alpha x : x \in M\}$.

Proof. (i) \implies (ii). Let $f \in A[X]$ be a monic polynomial with $f(\alpha) = 0$. Let $\beta \in A[\alpha]$. Then $\beta = g(\alpha)$ with $g \in A[X]$. Since f is monic, using division with remainder we find $q, r \in A[X]$ with $g = qf + r$, and $\deg r < \deg f = n$. We may write $r = c_0 + c_1X + \cdots + c_{n-1}X^{n-1}$ with $c_i \in A$. Thus,

$$\beta = g(\alpha) = q(\alpha)f(\alpha) + r(\alpha) = r(\alpha) = \sum_{i=0}^{n-1} c_i \alpha^i.$$

It follows that $A[\alpha]$ is generated as an A -module by $1, \alpha, \dots, \alpha^{n-1}$.

(ii) \implies (iii). Trivial.

(iii) \implies (i). let $\{\omega_1, \dots, \omega_r\}$ be a set of A -module generators for M . Since $1 \in M$ we may assume that $\omega_1 = 1$. Further, since $\alpha\omega_i \in M$ we have

$$\alpha\omega_i = c_{i1}\omega_1 + \cdots + c_{ir}\omega_r \quad \text{with } c_{ij} \in A \text{ for } i, j = 1, \dots, r.$$

This can be rewritten as

$$(\alpha I - C) \begin{pmatrix} \omega_1 \\ \vdots \\ \omega_r \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

where I is the $r \times r$ unit matrix and $C = (c_{ij})_{i,j=1,\dots,r}$ is the $r \times r$ -matrix with c_{ij} on the i -th row and j -th column. We multiply both sides of this identity on the left with the matrix consisting of the minors of $\alpha I - C$, i.e., with $D = (M_{ij})_{i,j=1,\dots,r}$, where $M_{ij} = (-1)^{i+j}$ times the determinant of the matrix, obtained by removing the j -th row and i -th column from $\alpha I - C$. Then since

$$D \cdot (\alpha I - C) = \det(\alpha I - C) \cdot I,$$

we obtain $\det(\alpha I - C)\omega_i = 0$ for $i = 1, \dots, r$, so $\det(\alpha I - C) = 0$. That is, α is a zero of $\det(XI - C)$, which is a monic polynomial from $A[X]$. \square

Corollary 3.23. *Let $A \subset B$ be a finite extension of commutative rings. Then every $\alpha \in B$ is integral over A .*

Proof. Apply Lemma 3.22, (iii) \implies (i) with $M = B$. □

Lemma 3.24. *Let A, B, C be commutative rings such that B is finite over A and C is finite over B . Then C is finite over A .*

Proof. Suppose that B is generated as an A -module by $\alpha_1, \dots, \alpha_m$, and C is generated as a B -module by β_1, \dots, β_n . A straightforward computation shows that C is generated as an A -module by $\alpha_i \beta_j$ ($i = 1, \dots, m, j = 1, \dots, n$). □

Lemma 3.25. *Let A, B be commutative rings. Then B is finite over A if and only if $B = A[\alpha_1, \dots, \alpha_r]$ for certain $\alpha_1, \dots, \alpha_r$ that are integral over A .*

Proof. Suppose that B is finite over A , say B is generated as an A -module by $\alpha_1, \dots, \alpha_r$. By Corollary 3.23, $\alpha_1, \dots, \alpha_r$ are integral over A , and clearly, $B = A[\alpha_1, \dots, \alpha_r]$.

Conversely, suppose that $B = A[\alpha_1, \dots, \alpha_r]$ where $\alpha_1, \dots, \alpha_r$ are integral over A . Let $B_0 := A$ and $B_i := A[\alpha_1, \dots, \alpha_i]$ for $i = 1, \dots, r$. Thus, $B_i = B_{i-1}[\alpha_i]$ for $i = 1, \dots, r$. By Lemma 3.22, (i) \implies (ii), B_i is finite over B_{i-1} for $i = 1, \dots, r$, and then by Lemma 3.24, $B = B_r$ is finite over A . □

Proposition 3.26. *Let $A \subset B$ be an extension of commutative rings.*

- (i) *Let $\alpha, \beta \in B$ be integral over A . Then $\alpha \pm \beta$ and $\alpha\beta$ are integral over A .*
- (ii) *Let $\beta_1, \dots, \beta_n \in B$ be integral over A , and let $\alpha \in B$ be a zero of $X^n + \beta_1 X^{n-1} + \dots + \beta_n$. Then α is integral over A .*

Proof. (i). By Lemma 3.25, the ring $A[\alpha, \beta]$ is finite over A , and then by Corollary 3.23, all elements of $A[\alpha, \beta]$ are integral over A .

(ii). Let $C = A[\beta_1, \dots, \beta_n]$. By Lemma 3.25, C is finite over A . The number α is integral over C , so by Lemma 3.22, $C[\alpha]$ is finite over C . Again by Lemma 3.25, $C[\alpha]$ is finite over A , and then by Corollary 3.23, α is integral over A . □

Definition. Let $A \subset B$ be an extension of commutative rings. By Proposition 3.26 (i), the set

$$C = \{\alpha \in B : \alpha \text{ is integral over } A\}$$

is a subring of B , called the *integral closure* of A in B . Note that by Proposition 3.26(ii) every element of B that is integral over C already belongs to C .

We say that A is *integrally closed in B* if $C = A$, i.e., every element of B that is integral over A already belongs to A .

Let A be an integral domain. The *integral closure* of A is defined as its integral closure in its quotient field. If this integral closure is equal to A itself, we call A *integrally closed*.

Examples. 1. \mathbb{Z} is integrally closed.

2. Let $A = \mathbb{Z}[2\sqrt{2}]$. The quotient field of A is $\mathbb{Q}(\sqrt{2})$, and its integral closure is $\mathbb{Z}[\sqrt{2}]$.

3.6 Further exercises

Let $\alpha \in \overline{\mathbb{Q}}$ be an algebraic number of degree d .

Recall that the *denominator* of α is the smallest positive $m \in \mathbb{Z}$ such that $m\alpha$ is an algebraic integer, notation $\text{den}(\alpha)$.

Further, we have defined the *house* of an algebraic number α by

$$|\overline{\alpha}| := \max(|\alpha^{(1)}|, \dots, |\alpha^{(d)}|)$$

where $d = \deg \alpha$ and $\alpha^{(1)}, \dots, \alpha^{(d)}$ denote the conjugates of α .

Lemma 3.6 implies that if α is a non-zero algebraic integer, then $|\overline{\alpha}| \geq 1$. In the next exercises you are asked to prove some further properties of the house.

Exercise 3.5. Let α, β be algebraic numbers. Prove that

$$|\overline{\alpha + \beta}| \leq |\overline{\alpha}| + |\overline{\beta}|, \quad |\overline{\alpha \cdot \beta}| \leq |\overline{\alpha}| \cdot |\overline{\beta}|, \quad |\overline{\alpha^n}| = |\overline{\alpha}|^n \text{ for } n \in \mathbb{Z}_{>0}.$$

Hint. Apply (3.2) with $K = \mathbb{Q}(\alpha, \beta)$.

Exercise 3.6. (i) Let α be an algebraic integer of degree d that is not equal to 0 and that is not an algebraic unit. Prove that $|\overline{\alpha}| \geq 2^{1/d}$.

(ii) Let α be a non-zero algebraic integer of degree d . Prove that $H(\alpha) \leq (2 \cdot |\overline{\alpha}|)^d$ (consider the minimal polynomial of α).

(iii) Compute an explicit expression $f(C, d)$ depending only on C and d , such that the number of algebraic integers $\alpha \in \mathbb{C}$ with $|\overline{\alpha}| \leq C$, $\deg \alpha \leq d$ is at most $f(C, d)$.

(iv) Let α be a non-zero algebraic integer. As mentioned above, we have $|\overline{\alpha}| \geq 1$. Prove that $|\overline{\alpha}| = 1 \iff \alpha$ is a root of unity.

(v) Let α be a non-zero algebraic integer of degree d which is not a root of unity. Compute an explicit expression $C(d) > 1$ depending only on d such that $|\overline{\alpha}| \geq C(d)$.

Hint. Consider the set $\{\alpha^n : 0 \leq n \leq n_0\}$ where n_0 is the largest integer n such that $|\overline{\alpha}|^n \leq 2$.

Remark. For a non-zero algebraic integer α , define the quantity

$$f(\alpha) := (\deg \alpha) \log |\overline{\alpha}|.$$

From (i) of the above exercise, one derives that $f(\alpha) \geq \log 2$ if α is a non-zero algebraic integer which is not an algebraic unit. From (v) it follows that there is $c(d) > 0$ depending only on d such that $f(\alpha) \geq c(d)$ for every algebraic unit α of degree d not equal to a root of unity; probably what you get from your answer in (v) is a function in d decreasing rapidly to 0 as $d \rightarrow \infty$.

The *Schinzel-Zassenhaus conjecture*, posed by them in 1965, asserts that there is a constant $c > 0$ independent of the degree of α such that $f(\alpha) \geq c$ for every algebraic unit α that is not a root of unity.

This conjecture was open until recently. It came as a surprise when in 2019, Vesselin Dimitrov proved that for every algebraic unit α not equal to a root of unity one has

$$f(\alpha) \geq \frac{\log 2}{4} = 0.1732\dots,$$

thereby establishing the conjecture of Schinzel and Zassenhaus.

In 1985, Boyd conjectured the following, which is still open: $f(\alpha)$ assumes its minimum at the zeros of $X^3 + X^2 - 1$, i.e., for every algebraic unit α not equal to a root of unity one has

$$f(\alpha) \geq f(\theta) = 0.4217\dots,$$

where θ is any zero of $X^3 + X^2 - 1$. We mention here that $X^3 + X^2 - 1$ has one real zero θ_1 with $0 < \theta_1 < 1$ and two complex conjugate zeros θ_2, θ_3 with $|\theta_3| = |\theta_2| = \theta_1^{-1/2} > 1$. Thus, if $\theta \in \{\theta_1, \theta_2, \theta_3\}$ is any zero of $X^3 + X^2 - 1$ then $f(\theta) = 3 \log |\theta_2| = 3 \log |\theta_3| = -\frac{3}{2} \log \theta_1$ for $i = 1, 2, 3$.

Exercise 3.7. (i) Let α be a non-zero algebraic number of degree d . Prove that

$$|\alpha| \geq \text{den}(\alpha)^{-d} \cdot |\overline{\alpha}|^{1-d}.$$

(ii) Give a proof for the following inequality of Liouville (1844):

let α be an algebraic number in \mathbb{R} of degree $d \geq 2$, not necessarily an algebraic integer. Then there is a constant $c(\alpha) > 0$ such that

$$\left| \alpha - \frac{x}{y} \right| \geq c(\alpha)y^{-d} \quad \text{for all } x, y \in \mathbb{Z} \text{ with } y > 0.$$

Hint. Assuming that $c(\alpha) < 1$, the inequality is certainly true if $|\alpha - x/y| > 1$, so we may assume that $|\alpha - x/y| \leq 1$. This gives a bound for $|x/y|$. Now apply (i) to $\alpha - x/y$.

(iii) Using (ii), prove that $\sum_{n=1}^{\infty} 10^{-n!}$ is transcendental.

Hint. Assume that $\alpha := \sum_{n=1}^{\infty} 10^{-n!}$ is algebraic of degree d . Take $x/y = \sum_{n=1}^N 10^{-n!}$.

Exercise 3.8. Denote by $\mathbb{Z}[[X]]$ the ring of formal power series with coefficients in \mathbb{Z} . Addition and multiplication of two power series $f(X) = \sum_{n=0}^{\infty} a_n X^n$, $g(X) = \sum_{n=0}^{\infty} b_n X^n$ are given by

$$f(X) + g(X) = \sum_{n=0}^{\infty} (a_n + b_n) X^n, \quad f(X)g(X) = \sum_{n=0}^{\infty} \left(\sum_{k=0}^n a_{n-k} b_k \right) X^n.$$

Further, we write $f(X) \equiv 0 \pmod{X^s}$ if $f(X)$ is divisible by X^s , i.e., if $a_0 = \dots = a_{s-1} = 0$.

Let $f_1(X), \dots, f_m(X) \in \mathbb{Z}[[X]]$ be formal power series whose coefficients have absolute values at most A , where $A \geq 1$. Let r, s be integers with $mr > s > 0$. Prove that there are polynomials $p_1(X), \dots, p_m(X) \in \mathbb{Z}[X]$, not all zero, of degree $< r$ and with coefficients with absolute values at most $(mrA)^{s/(mr-s)}$, such that

$$p_1(X)f_1(X) + \dots + p_m(X)f_m(X) \equiv 0 \pmod{X^s}.$$