1. Ring Theory II

In this chapter we continue to study the structure of rings. We especially consider special types of rings, group rings, polynomial rings, Artinian and Noetherian rings. All these types of rings are important because of their widespread applicability, especially in the context of calculations with algebraic objects. Polynomials are used to generate algebraic extensions of fields, for defining curves and surfaces. They belong to the most important tools in algebra. Noetherian rings will frequently be used in calculations since their ideals have a finite number of generators, hence arithmetic can be done explicitly with those.

2. Group Rings and Polynomial Rings

The study of group rings is a relatively new topic of classical algebra. It was initiated by the idea that rings possess more structural properties than groups, hence, if one associates a suitable ring to a group then the structure of that so-called group ring should reveal structural aspects of the underlying group. We cannot cover this topic in full generality. Hence, we recommend that the reader concentrates on the applications to polynomial rings later in this section.

Definition 2.1. *Let S be a semigroup and R be a ring. Then we define a* **semigroup ring** *$R[S]$ via*

$$R[S] := \{f \,:\, S \to R \mid \; f(s) = 0 \text{ for almost all } s \in S\}$$

with operations

$$\begin{aligned} \textit{addition} \;\; &: \;\; f+g \,:\, S \to R \,:\, s \mapsto f(s)+g(s) \;\;, \\ \textit{multiplication} \;\; &: \;\; f\,g \,:\, S \to R \,:\, s \mapsto \sum_{\substack{t_1 t_2 = s \\ t_1, t_2 \in S}} f(t_1)\, g(t_2) \end{aligned}$$

for all f, $g \in R[S]$.

Whereas the definition of addition is straighforward the notion of multiplication seems to be kind of artificial at first glance. However, if

we look at polynomials in one variable t with coefficients in R (for simplicity's sake let us assume that $R = \mathbb{R}$ as in highschool) then S is just the semigroup $(\mathbb{Z}^{\geq 0}, +)$ and a map f designs the coefficient $f(m)$ to the power t^m, i.e. the map f stands for the polynomial $\sum_{i\geq 0} f(i)t^i$, where the formally infinite sum is actually finite because of the condition imposed on f. On the other hand, when we multiply two polynomials given in their usual representation, say $\sum_{i=0}^{n} a_i t^i$ and $\sum_{j=0}^{m} b_j t^j$, it is quite cumbersome to write down their product:

$$\sum_{k=0}^{n+m}(\sum_{l=0}^{k} a_l b_{k-l})t^k \quad ,$$

where we must additionally require $a_l = 0 \;\; (l > n)$, $b_{k-l} = 0 \;\; (k - l > m)$. This shows why the notion of semigroup rings is advantageous. The advantages will become even more clear when we consider polynomials in several variables. Using the notion of semigroup rings we just choose $S = ((\mathbb{Z}^{\geq 0})^n, +)$ to obtain a polynomial ring over R in n variables. The usual problems, like showing that the order of variables does not matter, are no longer present, this becomes an easy consequence of the analogous property for the direct product of (semi) groups (shown in chapter 2.6).

We leave the verification of the ring axioms for $R[S]$ as an exercise to the reader. As a precedent we establish the law of associativity for multiplication:
For arbitrary $s \in S$ we have

$$\begin{aligned}
(f\,(g\,h))\,(s) &= \sum_{t_1 t_4 = s} f(t_1)\,(g\,h)\,(t_4) \\
&= \sum_{t_1 t_4 = s} f(t_1) \sum_{t_2 t_3 = t_4} g(t_2)\,h(t_3) \\
&= \sum_{t_1 t_2 t_3 = s} f(t_1)\,g(t_2)\,h(t_3) \\
&= \sum_{t_5 t_3 = s} \left(\sum_{t_1 t_2 = t_5} f(t_1) g(t_2) \right) h(t_3) \\
&= \sum_{t_5 t_3 = s} (fg)\,(t_5)\,h(t_3) \\
&= ((f\,g)\,h)\,(s).
\end{aligned}$$

Next we consider the necessary premises for embedding R, S into $R[S]$.

(1) Let S be a monoid with unit element e. We put

$$\iota_R \,:\, R \to R[S] \,:\, r \mapsto f_r \quad \text{with} \quad f_r(s) = \left\{ \begin{array}{ll} r & \text{for } s = e \\ 0 & \text{otherwise} \end{array} \right\} \ .$$

Then ι_R is a ringmonomorphism because of

$$\begin{aligned} f_{r+\tilde{r}}(s) &= \left\{ \begin{array}{ll} r+\tilde{r} & \text{for } s = e \\ 0 & \text{otherwise} \end{array} \right\} \\ &= \left\{ \begin{array}{ll} r & \text{for } s = e \\ 0 & \text{otherwise} \end{array} \right\} + \left\{ \begin{array}{ll} \tilde{r} & \text{for } s = e \\ 0 & \text{otherwise} \end{array} \right\} \\ &= f_r(s) + f_{\tilde{r}}(s) \ , \\ f_{r\tilde{r}}(s) &= \left\{ \begin{array}{ll} r\tilde{r} & \text{for } s = e \\ 0 & \text{otherwise} \end{array} \right\} \\ &= \sum_{t_1 t_2 = s} \left\{ \begin{array}{ll} r & \text{for } t_1 = e \\ 0 & \text{otherwise} \end{array} \right\} \left\{ \begin{array}{ll} \tilde{r} & \text{for } t_2 = e \\ 0 & \text{otherwise} \end{array} \right\} \\ &= (f_r \, f_{\tilde{r}})(s) \ , \end{aligned}$$

and $\ker(\iota_R) = \{0\}$.

(2) Let R be a unital ring. We put

$$\iota_S \,:\, S \to R[S] \,:\, s \mapsto F_s \quad \text{with} \quad F_s(t) = \left\{ \begin{array}{ll} 1 & \text{for } t = s \\ 0 & \text{otherwise} \end{array} \right\} =: \delta_{ts} \ ,$$

where δ_{ts} denotes the **Kronecker symbol** whose value is 1 if both indices coincide and otherwise 0.
ι_S is a homomorphism because of

$$\begin{aligned} F_{s\tilde{s}}(t) &= \delta_{t,s\tilde{s}} \\ &= \left\{ \begin{array}{ll} 1 & \text{for } s\tilde{s} = t \\ 0 & \text{otherwise} \end{array} \right\} \\ &= \sum_{t_1 t_2 = t} \delta_{t_1 s} \, \delta_{t_2 \tilde{s}} \\ &= \sum_{t_1 t_2 = t} \left\{ \begin{array}{ll} 1 & \text{for } t_1 = s \\ 0 & \text{otherwise} \end{array} \right\} \left\{ \begin{array}{ll} 1 & \text{for } t_2 = \tilde{s} \\ 0 & \text{otherwise} \end{array} \right\} \\ &= (F_s \, F_{\tilde{s}})(t) \ . \end{aligned}$$

Obviously, ι_S is injective and therefore a monomorphism.

If additionally S is a monoid then $R[S]$ has a unit element with respect to multiplication, namely F_e:

$$\begin{aligned}(F_e\, f)(t) &= \sum_{t_1 t_2 = t} F_e(t_1)\, f(t_2) \\ &= \sum_{t_1 t_2 = t} \delta_{e t_1}\, f(t_2) \\ &= f(t) \quad \text{for all} \quad f \in R[S] \ .\end{aligned}$$

(3) In case $R \ni 1$ we obtain

$$R[S] \ = \ \left\{ \sum_{s \in S} a_s\, F_s \ \middle|\ a_s \in R,\ a_s = 0 \text{ for almost all } s \in S \right\} \ .$$

If we identify $s \in S$ with its image $F_s = \iota_S(s)$ this becomes

$$R[S] \ = \ \left\{ \sum_{s \in S} a_s\, s \ \middle|\ a_s \in R,\ a_s = 0 \text{ for almost all } s \in S \right\} \ .$$

Then all calculations in $R[S]$ are easy:

$$\alpha \left(\sum_{s \in S} a_s\, s \right) \ = \ \sum_{s \in S} (\alpha\, a_s)\, s \quad \forall \alpha \in R,$$

$$\sum_{s \in S} a_s\, s + \sum_{s \in S} b_s\, s \ = \ \sum_{s \in S} (a_s + b_s)\, s,$$

$$\left(\sum_{s \in S} a_s\, s \right) \left(\sum_{t \in S} b_t\, t \right) \ = \ \sum_{s,t \in S} a_s\, b_t\, s\, t \ = \ \sum_{u \in S} \left(\sum_{st = u} a_s\, b_t \right) u.$$

Examples

(1) $S = \{ t^\nu \mid \nu \in \mathbb{Z}^{\geq 0} \} \cong \mathbb{Z}^{\geq 0}$, R a unital commutative ring.

$$R[S] \ = \ \left\{ \sum_{\nu = 0}^{\infty} a_\nu\, t^\nu \ \middle|\ a_\nu \in R,\ a_\nu \neq 0 \text{ for only finitely many } \nu \right\} =: R[t]$$

is the polynomial ring in the variable t over R. The elements of $R[t]$ are written as

$$f(t) = \sum_{i = 0}^{\infty} a_\nu\, t^\nu$$

with $a_\nu \in R$, almost all $a_\nu = 0$. Polynomials in one varable are usually called **univariate** polynomials.

(2)

$$S = \prod_{i=1}^{n} \{t_i^{\nu_i} \in \mathbb{Z}^{\geq 0}\} \cong (\mathbb{Z}^{\geq 0})^n \,,\ R \text{ a unital commutative ring } .$$

The elements of S can be written in the form $\mathbf{t}^{\underline{\nu}} := t_1^{\nu_1} \cdot \ldots \cdot t_n^{\nu_n}$ with $\underline{\nu} \in (\mathbb{Z}^{\geq 0})^n$. Then

$$R[S] \ = \ \left\{ \sum_{\underline{\nu} \in (\mathbb{Z}^{\geq 0})^n} a_{\underline{\nu}} \mathbf{t}^{\underline{\nu}} \ \middle| \ a_{\underline{\nu}} \in R, a_{\underline{\nu}} \neq 0 \text{ for only finitely many } \underline{\nu} \right\} =: R[\mathbf{t}]$$

is the polynomial ring in n variables $t_1, \ldots, t_n$ over R with elements

$$f(\mathbf{t}) = \sum_{\underline{\nu} \in (\mathbb{Z}^{\geq 0})^n} a_{\underline{\nu}} \mathbf{t}^{\underline{\nu}} \quad (a_{\underline{\nu}} \in R, \text{ almost all } a_{\underline{\nu}} = 0) \ .$$

Polynomials in several variables ($n \geq 2$) are usually called **multivariate** polynomials.

(3) S a group, R a unital ring. $R[S]$ is called **group ring**. Knowledge about the group ring yields information about the group itself. We cite without proof a result of Higman: If G, H are finite abelian groups with $\mathbb{Z}[G] \cong \mathbb{Z}[H]$ then G and H are isomorphic.

As we already mentioned important results on polynomial rings immediately follow from the properties of the semigroup used for their construction.

For example, we get

$$\begin{aligned} R[t_1, ..., t_n, t_{n+1}] &\cong R[t_1, ..., t_n]\,[t_{n+1}], \\ R[t_1, ..., t_n] &\cong R[t_{\pi(1)}, ..., t_{\pi(n)}] \quad \forall \pi \in \mathfrak{S}_n \end{aligned}$$

as an immediate consequence of the corresponding statements for direct products of (semi) groups.

For the elements of the monoid $(\mathbb{Z}^{\geq 0})^n$ we can introduce an ordering via

$$\mathbf{t}^{\underline{\nu}} \geq \mathbf{t}^{\underline{\mu}} \quad \Leftrightarrow \quad \underline{\nu} \geq \underline{\mu} \ .$$

There are various possibilities. We just mention the two most popular ones:

(i) lexicographic ordering

We put $\underline{\nu} \geq \underline{\mu}$ if and only if there exists an index $i \in \{1, ..., n\}$ with $\nu_j = \mu_j \ \ (j < i)$ and $\nu_i > \mu_i$. This means that for the smallest index i for which the coordinates of $\underline{\nu}$ and $\underline{\mu}$ differ the i-th coordinate of $\underline{\nu}$ is

larger than that of $\underline{\mu}$.
(ii) graded lexicographic ordering
We put $\underline{\nu} \geq \underline{\mu}$ if either $\sum_{i=1}^{n} \nu_i^2 > \sum_{i=1}^{n} \mu_i^2$ or, in case both sums are equal, $\underline{\nu}$ is lexicographically greater than $\underline{\mu}$ (including the case $\underline{\nu} = \underline{\mu}$). Here we first compare the Euclidean lengths of $\underline{\nu}$ and $\underline{\mu}$ and only if they are equal we make use of lexicographic ordering.

For a thorough study of (multivariate) polynomials we need to introduce a few definitions which will be mostly familiar from high school arithmetic.

Definition 2.2. *Let*

$$f(\mathbf{t}) = \sum_{\underline{\nu} \in (\mathbb{Z}^{\geq 0})^n} a_{\underline{\nu}} \mathbf{t}^{\underline{\nu}}$$

be an element of the polynomial ring $R[\mathbf{t}]$. The single summands $a_{\underline{\nu}}\mathbf{t}^{\underline{\nu}}$ are called **monomials**. *The* **degree** *of a non-zero monomial is defined as the sum of its exponents: $\nu_1 + \ldots + \nu_n$. The* **degree** *$deg(f)$ of a non-zero polynomial f is the maximum of the degrees of its monomials. The degree of 0 (as monomial or as polynomial) is formally defined to be $-\infty$. If we have a total ordering on the exponents - and therefore on the monomials - the coefficient of the largest monomial is called* **leading coefficient** *$l(f)$, sometimes also* **headterm**. *In case $l(f) = 1$ the polynomial f is called* **monic**.

We shortly consider the behavior of the degree function with respect to the addition and multiplication of polynomials. Comparing the degrees of the occuring monomials we immediately see that

$$\begin{aligned} \deg(f+g) &\leq \max\{\deg(f), \deg(g)\} \\ \deg(f\,g) &\leq \deg(f) + \deg(g) \ . \end{aligned}$$

The last inequality becomes an equation, if $l(f)$, $l(g)$ are no zero divisors.

Hence, the degree of the product of two polynomials equals the sum of their degrees over entire rings R. The property to be entire is therefore transfered from R to $R[\mathbf{t}]$:

$$R[\mathbf{t}] \text{ entire ring } \Leftrightarrow \ R \text{ entire ring } \ .$$

Because of the behavior of the degree function mentioned above we also obtain the result that the units of R and of $R[t]$ coincide:

$$f \in U(R[\mathbf{t}]) \ \Leftrightarrow \ f \in U(R) \ .$$

We will consider further properties of rings with respect to whether they transfer from R to $R[\mathbf{t}]$.

Theorem 2.3. *(***Hilbert's Basis Theorem***) Let R be a unital commutative ring. If R is noetherian then also $R[t]$ is noetherian.*

Proof Let $\mathbf{A}$ be an ideal of $R[t]$. Then we consider the polynomials of $\mathbf{A}$ of degree $i \in \mathbb{Z}^{\geq 0}$ and put

$$\mathbf{a}_i := \{x \in R \mid x = \mathrm{lc}(f) \text{ for an } f \in \mathbf{A} \text{ with } \deg(f) = i\} \cup \{0\} \ .$$

The $\mathbf{a}_i$ are ideals in R because of

(1) for $f,\, g \in \mathbf{A}$ with $\deg(f) = \deg(g) = i$ we either have $\mathrm{lc}(f+g) = \mathrm{lc}(f) + \mathrm{lc}(g) \neq 0$ or $\deg(f+g) < i$ and the coefficient of t^i of $f+g$ is zero,
(2) for $a = \mathrm{lc}(f) \in \mathbf{a}_i,\ r \in R$ we have $ra = 0$ or $rf \in \mathbf{A}$ with $\deg(rf) = i$ and $\mathrm{lc}(rf) = ra \in \mathbf{a}_i$.

Since we can multiply elements of $\mathbf{A}$ by t we obtain

$$\mathbf{a}_0 \subseteq \mathbf{a}_1 \subseteq \ldots \subseteq \mathbf{a}_r \subseteq \ldots \ .$$

Since R is noetherian this chain becomes stationary. Let $r \in \mathbb{Z}^{\geq 0}$ be minimal with $\mathbf{a}_r = \mathbf{a}_{r+k}\ \forall k \in \mathbb{N}$. Since R is noetherian each ideal $\mathbf{a}_i$ has finitely many generators $a_{i1}, \ldots, a_{in_i}$ $(n_i \in \mathbb{N})$ for $0 \leq i \leq r$. We fix elements $f_{ij} \in \mathbf{A}$ with $\deg(f_{ij}) = i$ and $\mathrm{lc}(f_{ij}) = a_{ij}$ for $0 \leq i \leq r,\ 1 \leq j \leq n_i$. We will show that $\mathbf{A} = \mathbf{B}$ for

$$\mathbf{B} \ := \ \langle f_{ij} \mid 0 \leq i \leq r\,,\ 1 \leq j \leq n_i \rangle \ .$$

Clearly, $\mathbf{B}$ is contained in $\mathbf{A}$. On the other hand, let $f \in \mathbf{A}$ with $\deg(f) = d$. The proof of $f \in \mathbf{B}$ is carried out by induction on d. For $d = 0$ there is nothing to show since f is contained in $\mathbf{a}_0 \subseteq \mathbf{B}$. We let therefore be $d > 0$ and assume that all elements of $\mathbf{A}$ of degree less than d belong to $\mathbf{B}$. We need to consider two cases.

(1) For $d > r$ we have

$$\mathbf{a}_d = \langle \mathrm{lc}(t^{d-r} f_{r1}), \ldots, \mathrm{lc}(t^{d-r} f_{rn_r}) \rangle \ ,$$

there exist $\gamma_1, \ldots, \gamma_{n_r} \in R$ such that

$$g \ := \ f - \sum_{i=1}^{n_r} \gamma_i t^{d-r} f_{ri}$$

is a polynomial of $\mathbf{A}$ with $\deg(g) < d$.

(2) For $d \leq r$ we analogously obtain a polynomial

$$g \ := \ f - \sum_{i=1}^{n_d} \tilde{\gamma}_i f_{di}$$

of degree less than d in $\mathbf{A}$.

According to our induction assumption in both cases the difference polynomial g belongs to the ideal $\mathbf{B}$, hence the polynomial f itself. This finishes the proof of $\mathbf{A} = \mathbf{B}$, the ideal $\mathbf{A}$ is finitely generated and therefore $R[t]$ noetherian.

□

Applying the preceding theorem n times we obtain that for noetherian rings R the polynomial ring in n variables $R[\mathbf{t}]$ is noetherian, too.

A similar discussion whether the properties of a ring R to be a principal ideal ring or a factorial ring transfer to $R[t]$ (and therefore to $R[\mathbf{t}]$) is postponed to the next section.

3. Univariate Polynomials

Univariate polynomials play a predominant role among all polynomials. This is mainly due to the fact that polynomial rings in one variable over a field have nicer properties than those with several variables. Also, polynomial rings in $n > 1$ variables could be considered as polynomial rings in one variable over a polynomial ring in $n - 1$ variables as base ring. This is usually not the appropriate approach, however, and therefore we shall consider univariate and multivariate polynomials in separate sections.

We begin with basic properties which will be of importance later on.

Definition 3.1. *Let Λ be a unital overring of R, i.e. $1_\Lambda = 1_R$, then for every $x \in \Lambda$ the mapping*

$$\Phi_x \,:\, R[t] \to \Lambda \,:\, f(t) \mapsto f(x)$$

*is a ring homomorphism with $\Phi_x|_R = Id_R$. Hence, it leaves every element of R invariant and is therefore called an R-***homomorphism***. Since Φ_x maps a polynomial to a ring element it is also called a* **specialization** *of the polynomial $f(t)$ to its value $f(x)$.*

That Φ_x is indeed a ring homomorphism can be easily verified and is left as an exercise to the reader.

Definition 3.2. *Let Λ, R be as in the previous definition. An element $x \in \Lambda$ is called* **zero** *of $f(t) \in R[t]$, if f is in the kernel of Φ_x. This is clearly tantamount to the more familiar version that $f(t)$ specializes to 0 at x.*

Proposition 3.3. *Let R be a unital entire ring. An R-homomorphism $\varphi \,:\, R[t] \to R[t]$ is an isomorphism exactly for $\varphi(t) = at + b$ with $a \in U(R)$, $b \in R$.*

Before we actually proof this we emphasize that every R-homomorphism $\varphi : R[t] \to \Lambda$ is uniquely determined by the image $\varphi(t)$. This is because of

$$\varphi\left(\sum_{i=0}^{n} a_i\, t^i\right) \;=\; \sum_{i=0}^{n} \varphi(a_i\, t^i) \;=\; \sum_{i=0}^{n} \varphi(a_i)\, \varphi(t)^i \;=\; \sum_{i=0}^{n} a_i\, \varphi(t)^i \quad .$$

Proof For $\varphi(t) = at + b$ with $a \in U(R),\ b \in R$ the inverse mapping is given by $\varphi^{-1}(t) = a^{-1}(t-b)$ satisfying $\varphi \circ \varphi^{-1} = \mathrm{Id}_{R[t]}$. On the other hand, if φ is an $R[t]$-isomorphism then φ maps t onto some polynomial of $R[t]$, say $\varphi(t) = g(t) := \sum_{i=0}^{n} a_i t^i \in R[t]$ and φ being surjective there exists $f(t) = \sum_{j=0}^{m} b_j t^j \in R[t]$ with $t = \varphi(f(t))$. This yields

$$t = \varphi(f(t)) = \varphi(\sum_{j=0}^{m} b_j t^j) = \sum_{j=0}^{m} b_j \varphi(t)^j = \sum_{j=0}^{m} b_j g(t)^j = f(g(t))$$

and comparing degrees we obtain

$$1 = \deg(t) = \deg(f(g(t)) = \deg(g)\deg(f) \quad .$$

The latter is possible only for $\deg(f) = \deg(g) = 1$, hence $g(t) = at + b,\ f(t) = ct + d$ $(a,\, b,\, c,\, d \in R)$. From

$$\begin{aligned} t &= f(g(t)) \\ &= c\,(at + b) + d \\ &= cat + bc + d \end{aligned}$$

we deduce $1 = ac$, $0 = bc + d$ and therefore $a \in U(R)$.

□

Definition 3.4. *Let Λ be a unital overring of the ring R. An element $x \in \Lambda$ is called* **algebraic** *over R, if the mapping $\varphi_x : R[t] \to \Lambda$ is not injective, i.e. there exists a polynomial $f(t) \in R[t]$ with $f(x) = 0$, x is a zero of a suitable non-constant polynomial of $R[t]$. If x is not algebraic over R it is called* **transcendental** *over R.*

Examples $\sqrt{2} \in \mathbb{R}$ is algebraic over $\mathbb{Z}$ since it is a zero of $f(t) = t^2 - 2 \in \mathbb{Z}[t]$. $e,\ \pi \in \mathbb{R}$ are transcendental over $\mathbb{Z}$ (respectively $\mathbb{Q}$). For a proof of the last statement the reader is refered to [**?**].

Because of the definition of algebraic elements it is important to characterize zeros of polynomials by purely polynomial ring properties. This is achieved upon showing that we can have division with remainder in polynomial rings and will surely have it in case R is a field.

Proposition 3.5. *Let us assume that $f(t)$, $g(t)$ are polynomials in $R[t]$ with the leading coefficient of g being a unit in R. Then there exist polynomials $q(t) := q(f,g)(t)$, $r(t) := r(f,g)(t) \in R[t]$ satisfying $f(t) = q(t)g(t) + r(t)$ and $\deg(r) < \deg(g)$. (This includes $r = 0$ with $\deg(r) = -\infty$ since our assumption on $l(g)$ yields $g \neq 0$.)*

Proof Since the polynomials q, r are given explicitly via calculations in R we give an algorithmic proof.

Algorithm 1. *(division with remainder for polynomials)*
Input *Polynomials $f(t)$, $g(t) \in R[t]$ with $l(g) \in U(R)$.*
Initilization *Set $m := \deg(g)$, $r_0 := f$, $q_0 := 0$, $i := 0$, $\deg(r_i) =: m_i$.*
Step *While $k_i := m_i - m \geq 0$ set*

$$\begin{aligned} \lambda_i &:= l(r_i)/l(g) \\ r_{i+1} &:= r_i - \lambda_i g(t) t^{k_i} \\ q_{i+1} &:= q_i - \lambda_i t^{k_i} \end{aligned}$$

and increase i by 1.
Output *Polynomials $q(t) := q_i(t)$, $r(t) := r_i(t) \in R[t]$ with $f(t) = q(t)g(t) + r(t)$ and $\deg(r) < \deg(g)$.*

In each step from the remaining polynomial r_i (initially $f(t)$) the polynomial $(l(r_i)/l(g)g(t)t^{\deg(r_i)-\deg(g)}$ is subtracted. This decreases the degree of the remainder. It can be carried out until the degree of the remainder becomes smaller than $\deg(g)$. The divison of the leading coefficients is always possible in R since we assumed $l(g) \in U(R)$.
□

Remark We note that division with remainder is always possible in case $g(t)$ is monic.

Proposition 3.6. *Let R be a unital commutative ring, $f(t) \in R[t]$ with $\deg(f) \geq 1$ and Λ be a unital overring of R. $x \in \Lambda$ is a zero of $f(t)$, if and only if $(t - x)$ divides $f(t)$ in $\Lambda[t]$.*

Proof Division with remainder can be carried out in $\Lambda[t]$ since $l(t - x) = 1$ is a unit in Λ. It follows

$$f(t) = Q(f,\, t - x)\,(t - x) + R(f,\, t - x)$$

with $\deg(R(f,\, t-x)) < \deg(t-x) = 1$, hence $R(f,\, t-x)$ is constant. Now we specialize $t \mapsto x$:

$$\begin{aligned} x \text{ zero} &\Leftrightarrow 0 = f(x) \\ &\Leftrightarrow R(f,\, t-x)\,(x) = 0 \\ &\Leftrightarrow R(f,\, t-x) = 0. \end{aligned}$$

□

For division with remainder in polynomial rings (**pseudodivision**) see the detailed exercise 1. Here we just present an illustrative example.

Example For $R = \mathbb{Z}$ the polynomial $f(t) = t^3 - 2$ is not divisible by $g(t) = 2t - 1$ in $R[t]$. However, if we multiply the first polynomial with $l(g)^{deg(f)-deg(g)+1}$ we obtain

$$2^3(t^3 - 2) = (4t^2 + 2t + 1)\,(2t - 1) - 15 \text{ in } \mathbb{Z}[t] \quad .$$

(The reader is advised to carry out this example with the algorithm given above.)

If the underlying ring is a field F then division with remainder is always possible in case $g(t)$ is non–zero. Hence, the polynomial ring $F[t]$ becomes a Euclidean ring with the degree function as Euclidean function.

Theorem 3.7. *A polynomial ring $F[t]$ over a field F is a Euclidean ring.*

We note that repeated divison with remainder yields the greatest common divisor of two polynomials exactly as it did for two rational integers. Since $F[t]$ is Euclidean and therefore a principal ideal ring we even obtain a representation of the greatest common divisor in terms of f, g because the principal ideal $\gcd(f, g)$ equals the ideal $f(t)F[t] + g(t)F[t]$ (see exercise ...). This is of importance for finite extensions of fields, for example.

We model repeated division with remainder for two polynomials f, g as follows. A sequence of polynomials $(f_i)_{i\in\mathbb{Z}^{\geq 0}}$ is calculated via $f_0 := f$, $f_1 := g$ and – for $f_{i+1} \neq 0$ – $f_i = q_{i+1}f_{i+1} + f_{i+2}$ by division with remainder.

Let $f(t)$, $g(t) \in F[t]$ be given. If both polynomials are 0 then their greatest common divisor is also 0 by definition. It is represented as $0 = 0 \cdot f + 0 \cdot g$. If $0 = f \neq g$ then the greatest common divisor is $\frac{1}{l)g)}g(t)$. It is represented via $\gcd(f, g) = 0 \cdot f + \frac{1}{l(g)}g$. Analogously, for $0 = g \neq f$ we obtain $\gcd(f, g) = \frac{1}{l(f)}f + 0 \cdot g$. For the more interesting

case $fg \neq 0$ we present the following algorithm. We note that we will have $f_i = \lambda_i f_0 + \mu_i f_1$ at each step.

Algorithm 2. *(polynomial gcd with presentation)*
Input *Non–zero polynomials* $f(t), g(t) \in F[t]$.
Initialization *Set* $f_0 := f$, $f_1 := g$, $\lambda_0 := \mu_1 := 1$, $\lambda_1 := \mu_0 := 0$ *and* $i := 0$.
Step *While* $f_{i+1} \neq 0$ *set*

$$\begin{aligned} f_{i+2} &:= f_i - q_i f_{i+1} \quad (\textit{division with remainder}) \\ \lambda_{i+2} &:= \lambda_1 - q_i \lambda_{i+1} \\ \mu_{i+2} &:= \mu_i - q_i \mu_{i+1} \end{aligned}$$

then increase i *by 1.*
Output $\gcd(f,g) := \frac{1}{l(f_i)} f_i$, $\lambda := \frac{1}{l(f_i)} \lambda_i$, $\mu := \frac{1}{l(f_i)} \mu_i \in F[t]$ *with* $\gcd(f,g) = \lambda f + \mu g$.

The algorithm is valid as the output polynomial $f_i(t)$ divides $f_{i-1}(t)$ (because of $f_{i+1}(t) = 0$); hence it also divides $f_{i-2}(t), ..., f_1(t), f_0(t)$. On the other hand, any common divisor of $f_0(t)$ and $f_1(t)$ divides $f_2(t)$; hence it also divides $f_3(t), ..., f_i(t)$. Both properties yield $f_i(t) = \gcd(f_0, f_1)$.

After this excursion into computational aspects we proceed with a few consequences of the last theorem. $F[t]$ is a principal ideal ring and a unique factorization ring. For an irreducible polynomial $f(t) \in F[t]$ the factorring $F[t]/f(t)F[t]$ is again a field. In $F[t]$ the number of zeros of a polynomial – counted with respect of their multiplicities – is bounded by the polynomial degree. (This is also true over entire rings, but not in general, as the example $t^2 - 1 \in (\mathbb{Z}/8\mathbb{Z})[t]$ shows, see exercise ...)

Proposition 3.8. *For unital commutative rings* R *the following equivalence holds:*

$$R[t] \textit{ principal ideal ring} \quad \Leftrightarrow \quad R \textit{ field} \ .$$

Proof We already showed that a polynomial ring over a field is a principal entire ring. To show the opposite direction we consider the ring epimorphism

$$\varphi_0 \ : \ R[t] \to R \ : \ f(t) \mapsto f(0) \ .$$

If $R[t]$ is a principal ideal ring then it is a priori an entire ring and therefore R itself must be an entire ring. The homomorphism theorem for rings tells us that $R \cong R[t]/\ker(\varphi_0)$, hence $\ker(\varphi_0)$ is a prime ideal

and therefore maximal in the principal entire ring $R[t]$. Therefore the factorring $R[t]/\ker(\varphi_0)$ is a field. It is isomorphic to R because of the surjectivity of φ_0.
□

Remark An important consequence of this proposition is that polynomial rings in more than one variable are not any more principal ideal rings.

Contrary to this the property of being a factorial ring is transfered from R to $R[t]$. This will be shown below.

Theorem 3.9. *(Gauß) If R is a factorial ring than so is $R[t]$.*

The proof of Gauß' theorem is a bit complicated and will be based on preparatory lemmata.

Proposition 3.10. *Let R be a unital commutative ring. If $\mathfrak{a}$ is an ideal (a prime ideal) of R, then $\mathfrak{a}[t] := \{f(t) = \sum_{i=0}^{n} a_i t^i \in R[t] \mid a_i \in \mathfrak{a} \ \ (0 \leq i \leq n)\}$ is an ideal (a prime ideal) of $R[t]$.*

Proof It is obvious that for an ideal $\mathfrak{a}$ of R also $\mathfrak{a}[t]$ is an ideal of $R[t]$.
Now let us assume that $\mathfrak{a}$ is a prime ideal of R. For

$$f(t) = \sum_{i=0}^{n} a_i t^i,\ g(t) = \sum_{j=0}^{m} b_j t^j \in R[t] \setminus \mathfrak{a}[t] \ \ ,$$

the polynomials f, g have coefficients a_i, $b_j \notin \mathfrak{a}$ for suitable indices i, j; we choose i, j minimal with this property. Then the coefficient of t^{i+j} of the product of f and g satisfies

$$c_{i+j} := \sum_{k=0}^{i+j} a_k\, b_{i+j-k} \equiv a_i b_j \bmod \mathfrak{a} \ \ ,$$

and therefore also c_{i+j} is not in the ideal $\mathfrak{a}$. This implies $fg \notin \mathfrak{a}[t]$.
□

Definition 3.11. *Let R be a factorial ring and $f(t) = \sum_{i=0}^{n} a_i t^i \in R[t]$ with $\deg(f) \geq 0$. Then $I(f) := \gcd\{a_0, a_1, ..., a_n\}$ is called* **content** *of $f(t)$. In case $I(f) = 1$ the polynomial $f(t)$ is said to be* **primitive**.

Remark If R is factorial then any polynomial $f(t) \in R[t]$ with $\deg(f) \geq 0$ can be written as a product of $I(f)$ and a polynomial $f_p(t) \in R[t]$ which is primitive. The polynomial $f_p(t)$ is also called the primitive part of $f(t)$.

Proposition 3.12. *If R is factorial then the product of two primitive polynomials of $R[t]$ is primitive.*

Proof Let $f(t)$, $g(t) \in R[t]$ be primitive and $h := f\,g$. If $I(h)$ is not contained in $U(R)$ then there exists a prime element $\pi \in R$ which divides all coefficients of h. Because of **??** the principal ideal $R\pi$ is a prime ideal, hence $R\pi[t]$ is a prime ideal of $R[t]$ according to the previous proposition. From $f\,g \in R\pi[t]$ we conclude that either $f(t)$ or $g(t)$ is contained in $R\pi[t]$, i.e. all coefficients of f or of g are divisible by π contrary to our assumption $I(f) = I(g) = 1$.
□

Remark For arbitrary polynomials f, g over a factorial ring R the content of their product $I(fg)$ equals the product of their contents $I(f)$ and $I(g)$. This is a direct consequence of the last proposition and the remark preceding it.

The next lemma is known as Gauß' lemma in the literature.

Lemma 3.13. *Let R be a factorial ring with quotient field $K = \mathfrak{Q}(R)$. If $h(t) \in R[t]$ has a positive degree and a factorisation $h = f_1 f_2$ in $K[t]$, then there is also a factorisation $h = cg_1g_2$ in $R[t]$ with primitive polynomials g_1, g_2 and $c \in R$. There exist $\alpha_i \in K$ with $\alpha_i f_i = g_i$ $(i = 1, 2)$.*

Proof Let λ_i be the least common multiples of the denominators of the coefficients of f_i $(i = 1, 2)$. We put $\mu_i := I(\lambda_i f_i)$. Then the primitive parts of $g_i := (\lambda_i f_i)_p$ satisfy

$$\lambda_1\lambda_2\, h = \mu_1\mu_2\, g_1g_2 \quad .$$

From this we conclude $\lambda_1\lambda_2\, I(h) = \mu_1\mu_2$. It follows that $\mu_1\mu_2 = (\lambda_1\lambda_2)\, c$ $(c \in R)$, hence, $h = c\, g_1\, g_2$. The last statement is true with $\alpha_i = \frac{\lambda_i}{\mu_i}$ for $i = 1, 2$.
□

Remarks

(1) If $f(t) \in R[t] \setminus R$ is irreducible then f remains irreducible in $\mathfrak{Q}(R)[t]$. For example, if $n \in \mathbb{Z}$ is not a square then $t^2 - n$ is irreducible in $\mathbb{Z}[t]$. This implies $\sqrt{n} \notin \mathbb{Q}$. Putting it negative: If $f(t) \in R[t]$ is reducible in $K[t]$, then it is also reducible in $R[t]$.

(2) Let $f, g \in R[t]$ and g primitive with $g \mid f$ in $K[t]$. Then g divides f already in $R[t]$.

(3) Two primitive polynomials $f, g \in R[t]$ are associated in $K[t]$ if and only if they are associated in $R[t]$.

After this the proof of Gauß' theorem is straightforward.

Proof of 3.9

We recall that the irreducible elements of $R[t]$ belong to two separate classes:

(1) irreducible elements of R,
(2) irreducible polynomials $f(t) \in R[t]$ of positive degree.

$R[t]$ is a unital entire ring since R has this property. Let $f(t) \in R[t]$ be fixed. Without loss of generality we assume that $\deg(f) > 0$. In the factorial ring $K[t]$, K denoting the quotient field of R, f has a factorisation into irreducible elements: $f = \tilde{q}_1 \cdot \ldots \cdot \tilde{q}_r$ with $\tilde{q}_i \in K[t]$. By Gauß' lemma 3.13 we obtain from this a factorisation

$$f = c\, q_1 \cdot \ldots \cdot q_r$$

with

$$q_i = \alpha_i \tilde{q}_i \in R[t] \;\; (\alpha_i \in K,\; 1 \leq i \leq r)$$

primitive and irreducible, $c \in R$. Since R was assumed to be factorial also c has a factorisation into irreducible elements in R.
If f admits two such factorisations in $R[t]$, say

$$f = d\, q_1 \cdot \ldots \cdot q_r = c\, p_1 \cdot \ldots p_s \;\; (\deg(q_i) > 0,\; \deg(p_j) > 0),$$

then the q_i, p_j are irreducible in $K[t]$, too, hence we obtain $r = s$ and after a potential reordering $q_i = \alpha_i p_i \;\; (\alpha_i \in K,\; 1 \leq i \leq r)$. The q_i and p_i are therefore associated in $R[t]$. It follows that d is also associated to c. Since R is a factorial ring the theorem follows.

□

Since every polynomial of a factorial ring $R[t]$ is a product of irreducible ones the irreducible polynomials are of special interest. We note that polynomials of degree one are necessarily irreducible. There is an easy test whether a first degree polynomial $at + b$ is a potential divisor of an arbitrary polynomial:

$$(at + b) \mid \left(\sum_{i=0}^{n} a_i\, t^{n-i} \right)$$

obviouslyly implies $a \mid a_0,\; b \mid a_n$.

Example Let us discuss for which $a \in \mathbb{Z}$ the polynomial $f(t) = t^5 + at + 1$ is irreducible in $\mathbb{Q}[t]$. Since f is primitive, irreducibility in $\mathbb{Z}[t]$ and in $\mathbb{Q}[t]$ is tantamount. We therefore need to look for potential divisors of f in $\mathbb{Z}[t]$ only. If f is not irreducible, it must have a factor of degree either one or two. Having a linear factor means having a zero. We find that $f(1) = a + 2$, $f(-1) = a$, hence f is reducible for

$a \in \{-2, 0\}$.
Next we are looking for quadratic factors:

$$t^5 + at + 1 = (t^2 + \alpha\, t + \beta)\,(t^3 + \gamma\, t^2 + \delta\, t + \varepsilon) \ .$$

Calculating the right-hand side and comparing coefficients we get the following system of equations:

$$\alpha + \gamma = 0,\ \delta + \alpha\,\gamma + \beta = 0,\ \varepsilon + \alpha\,\delta + \beta\,\gamma = 0,\ \alpha\,\varepsilon + \beta\,\delta = a,\ \beta\,\varepsilon = 1 \ .$$

We eliminate variables by setting

$$\gamma = -\alpha,\ \delta = \alpha^2 - \beta,\ \varepsilon = \alpha\,(2\beta - \alpha^2)$$

and obtain from the two remaining equations either
$\beta = \varepsilon = 1$ in which case the only solution is $1 = a = \alpha = -\gamma\,,\ \delta = 0$,
or
$\beta = \varepsilon = -1$ in which case there is no solution since the second relation for ε becomes $1 = \alpha(2 + \alpha^2)$.
Hence, the given polynomial is reducible if and only if a is 0,1 or -2.

In general there are very few powerful methods which allow to decide irreducibility of a given polynomial. Here we discuss just two. (For R being a finite field or for $R = \mathbb{Q}$ there are better methods which will be introduced later.)

Theorem 3.14. *(Eisenstein criterion) Let R be a factorial ring and $f(t) = \sum_{i=0}^{n} a_i\, t^i \ \in R[t]$ be a polynomial of positive degree. If R contains a prime element π such that $\pi \mid a_i \ \ (0 \le i < n)$, $\pi^2 \nmid a_0$ and $\pi \nmid a_n$, then $f(t)$ is irreducible in $\mathfrak{Q}(R)[t]$.*

Proof We assume that $f(t)$ has a factorisation in $\mathfrak{Q}(R)[t]$ into two polynomials of positive degree. According to 3.13 we also get such a factorisation in $R[t]$. We therefore assume that we have a factorisation in $R[t]$, say $f(t) = g(t)\,h(t)$ with $\deg(g)\deg(h) > 0$. We put

$$g(t) = \sum_{i=0}^{d} b_i\, t^i, \quad h(t) = \sum_{j=0}^{m} c_j\, t^j \ .$$

For the coefficients of the product of g and h we obtain the necessary conditions

$$a_i := \sum_{\substack{k=0 \\ k \le d \\ i-k \le m}}^{i} b_k\, c_i - k \ \ (0 \le i \le n) \ .$$

From $a_0 = b_0\, c_0$ and $\pi \mid a_0,\ \pi^2 \nmid a_0$ we conclude that π either divides b_0 or c_0 but not both. Without loss of generality (g, h are arbitrary so

far) we assume that $\pi \mid b_0$ and $\pi \nmid c_0$. Then we will show by induction that also $\pi \mid b_j \;\; (1 \leq j \leq d)$. Obviously, we have for $d \geq i > 0$:

$$a_i = \sum_{\substack{k=0 \\ i-k \leq m}}^{i} b_k \, c_{i-k} = \sum_{\substack{k=0 \\ i-k \leq m}}^{i-1} b_k \, c_{i-k} + b_i \, c_0 \equiv 0 \bmod \pi$$

because of our induction assumption. We conclude that $\pi | b_i c_0$ and because of $\pi \nmid c_0$ therefore $\pi \mid b_i$. Eventually we obtain for $i = d$ that $\pi \mid b_d c_m = a_n$. This is certainly in contradiction to our premises.
□

Example

(1) Let a be an integer. If there exists a prime number p with $p \mid a$ and $p^2 \nmid a$ then the polynomial $t^n - a$ is irreducible in $\mathbb{Q}[t]$ and in $\mathbb{Z}[t]$. Especially, it follows that $\sqrt[n]{a}$ is not rational.

(2) In $\mathbb{Q}[t]$ the following polynomials are irreducible according to 3.14:

$$\begin{aligned} f_1(t) &= 3\,t^5 - 15 \quad (p = 5), \\ f_2(t) &= 2\,t^{10} - 21 \quad (p = 3,\, 7), \\ f_3(t) &= 5\,t^5 - 12\,t^4 + 24\,t^3 + 2\,t^2 - 4\,t + 34 \quad (p = 2). \end{aligned}$$

We note that only the last two polynomials f_2, f_3 are also irreducible in $\mathbb{Z}[t]$ since the content of f_1 is 3 and therefore not a unit in $\mathbb{Z}$.

(3) For prime numbers p the p–th roots of unity are zeros of $t^p - 1$, they form a cyclic group of order p. $t^p - 1$ ist reducible since $(t-1) \mid (t^p - 1)$. The formula for the sum of the geometric series tells us that

$$\frac{t^p - 1}{t - 1} = \sum_{i=0}^{p-1} t^i =: \Phi_p[t] \;\; .$$

The polynomial $\Phi_p(t)$ is called p–th cyclotomic polynomial. Its zeros lead to the construction of a regular p–gon, they yield a division of the unit circle into p equal parts. Because $t \mapsto t + 1$ is an isomorphism of $R[t]$ (see 3.3) we conclude that $\Phi_p(t)$ is irreducible if and only if $\Phi_p(t+1)$ is irreducible. This trick of changing the variable allows an easy demonstration of the irreducibility of Φ_p. (We remark that it can be employed also to other polynomials to turn them into polynomials satisfying

the Eisenstein criterion.) We calculate

$$\begin{aligned}\Phi_p(t+1) &= \frac{(t+1)^p - 1}{(t+1) - 1} \\ &= \frac{\sum_{i=0}^{p} \binom{p}{i} t^i - 1}{t} \\ &= t^{p-1} + \sum_{i=1}^{p-1} \binom{p}{i} t^{i-1} \quad ,\end{aligned}$$

and in the resulting monic polynomial the coefficients

$$\binom{p}{i} = \frac{p \cdot (p-1) \cdot \ldots \cdot (p-i+1)}{1 \cdot \ldots \cdot i} \quad (1 \leq i \leq p-1)$$

are all divisible by p and the lowest one, $\binom{p}{1} = p$, is not divisible by p^2. Hence, $\Phi_p(t+1)$ is irreducible according to Eisenstein's criterion 3.14.

It should be noted that very few polynomials satisfy the premises of Eisenstein's criterion. In practice the following method is more powerful for proving irreducibility.

Theorem 3.15. *(Reduction) Let R, S be two unital entire rings and $\varphi : R \to S$ be a ring homomorphism with $\varphi(1_R) = 1_S$. Then φ can be canonically extended to a ring homomorphism $\Phi : R[t] \to S[t]$ with $\Phi|_R = \varphi$ via*

$$\sum_{i=0}^{n} a_i t^i \mapsto \sum_{i=0}^{n} \varphi(a_i) t^i \quad .$$

Let $f(t) \in R[t]$ with $\deg(\Phi(f)) = \deg(f) > 0$. If $\Phi(f)$ is irreducible in $S[t]$ then f cannot be written as a product $f = g\,h$ with $\deg(g)\ \deg(h) > 0$ in $R[t]$.

Proof

(1) It is easily verified (see exercises) that

$$\Phi \ : \ R[t] \to S[t] \ : \ \sum_{i=0}^{n} a_i\, t^i \mapsto \sum_{i=0}^{n} \varphi\left(a_i\right) t^i$$

is indeed a ring homomorphism. Its kernel is $\ker(\Phi) = \ker(\varphi)[t]$ because of $\Phi|_R = \varphi$.

(2) If $f = g\,h$ is a proper factorisation, i.e. $(\deg(g)\ \deg(h) > 0)$ in $R[t]$, then there is a proper factorisation $\Phi(f) = \Phi(g)\,\Phi(h)$ and $\deg\left(\Phi\left(g\right)\right) \leq \deg\left(g\right)$, $\deg\left(\Phi\left(h\right)\right) \leq \deg\left(h\right)$. Because of $\deg\left(\Phi\left(f\right)\right) = \deg\left(f\right)$ and S being an entire ring we obtain upon

comparing degrees that $\Phi(g)\,\Phi(h)$ is a proper factorisation of $\Phi(f)$. This is a contradiction to our premises.

□

The last theorem is frequently applied in irreducibility tests for polynomials in $\mathbb{Z}[t]$. In that case we choose $R = \mathbb{Z},\ S = \mathbb{Z}/p\mathbb{Z}$ for a prime number p with $p \nmid l(f)$.

Examples

(1) For $f(t) = t^3 + 39t^2 - 4t + 8 \in \mathbb{Z}[t]$ we choose $p = 3\ :\ \Phi(f) = t^3 - t - 1$ is irreducible in $\mathbb{Z}/3\mathbb{Z}[t]$, since it does not have a zero in $\mathbb{Z}/3\mathbb{Z}$.

(2) For $f(t) = t^2 + (10^{170} + 1)t + (10^{54821} + 343) \in \mathbb{Z}[t]$ we choose $p = 2\ :\ \Phi(f) = t^2 + t + 1$ is irreducible in $\mathbb{Z}/2\mathbb{Z}[t]$. This is of interest since the detection of a zero via the divisors of the constant coefficient is practically impossible in this case.

Our next subject will be the study of calculating solutions of equations. Let R be a unital commutative ring and $f(t) \in R[t]$, say $f(t) = \sum_{i=0}^{n} a_i\, t^{n-i}$. We want to find an element x either in R or in a unital overring Λ which satisfies $f(x) = 0$. If a_0 is not a zero divisor the multiplication by a_o^{n-1} yields:

$$(a_0x)^n + a_1(a_0x)^{n-1} + \ldots + a_n a_0^{n-1} = 0\ .$$

Hence, every solution $y \in R$ of $y^n + a_1y^{n-1} + \ldots + a_n a_0^{n-1} = 0$ corresponds to a solution $x = \frac{y}{a_0}$ in $\mathfrak{Q}(R)$ and vice versa. Therefore we will assume that f is monic from now on. A unital overring Λ of R in which f has a zero is called **solution ring** of the equation $f(x) = 0$.

Lemma 3.16. *Let R be a unital commutative ring and $f(t) \in R[t]$ monic of positive degree. Then $\Lambda := R[t]/f(t)R[t]$ is a solution ring of f. We note that R can be embedded into Λ.*

Proof The ring Λ has an R–basis $t^\nu + f(t)R[t]\ \ (0 \le \nu < \deg(f))$, since every polynomial $g(t) \in R[t]$ can be decomposed in $R[t]$ into

$$g(t) = Q(g,f)(t)\, f(t)\ +\ R(g,f)(t) \quad \text{with} \quad \deg(R(g,f)) < \deg(f)$$

by division with remainder. Hence, every residue class modulo $f(t)R[t]$ contains a unique representative of degree less than $\deg(f)$ and that representative can be written as a linear combination of the $t^\nu + f(t)R[t]$ $(0 \le \nu < \deg(f))$ with coefficients in R. Such a presentation is also unique since the difference of two different polynomials of degree $<$ $\deg(f)$ each is again a polynomial of degree $<$ $\deg(f)$. Since that difference is non zero it does not represent the class $f(t)R[t]$.

By construction of Λ we have $f(t + f(t)R[t]) = f(t) + f(t)R[t] = f(t)R[t]$, i.e. $x := t + f(t)R[t]$ is a zero of f in Λ.
An embedding of R into Λ is given by

$$\tau \,:\, R \to \Lambda \,:\, a \mapsto a + f(t)R[t] \ .$$

□

Remark The ring $\Lambda = R[t]/f(t)R[t]$ has the following 3 properties:

(1) Λ is a unital overring of R.
(2) Λ is generated over R by a zero $x = t + f(t)R[t] \in \Lambda$ of the polynomial f and has an R–basis of $\deg(f)$ elements.
(3) For every solution ring S in which $f(t)$ has a zero y there exists a ring homomorphism

$$\varphi \,:\, \Lambda \to S \,:\, \sum_{i=0}^{\deg(f)-1} a_i t^i + f(t)R[t] \mapsto \sum_{i=0}^{\deg(f)-1} a_i y^i \ .$$

A unital overring of R with these three properties is called **equation ring** for $f(x) = 0$. We emphasize that in equation rings (and similarly in solution rings) all calculations can be carried out easily. This will be demonstrated now. Let us assume that

$$f(t) \ = \ t^n \ + \sum_{i=1}^{n} \xi_i t^{n-i} \in R[t] \ .$$

Then any $\alpha \in \Lambda$ has a unique presentation

$$\alpha = \sum_{i=0}^{n-1} a_i x^i \ \ (a_i \in R) \ .$$

Two elements α and $\beta = \sum_{j=0}^{n-1} b_j x^j$ of Λ can be added by just adding coefficients:

$$\alpha + \beta \ = \ \sum_{i=0}^{n-1} (a_i + b_i) x^i \ .$$

Multiplication is only slightly more difficult. The immediate result

$$\alpha\beta \ = \ \sum_{k=0}^{2n-2} \left(\sum_{l=0}^{k} a_l b_{k-l} \right) x^k \ \ (a_l = 0 \ \ (l \geq n) \,,\ b_{k-l} = 0 \ \ (k - l \geq n))$$

must however be reduced to powers x^k with $k < n$. But this is exactly what the purpose of a solution ring is. Namely, we recursively obtain

basis presentations for all powers of x via

$$\begin{aligned} x^n &= -\sum_{i=1}^{n} \xi_i x^{n-i} \ , \\ x^{n+1} &= -\sum_{i=2}^{n} \xi_i x^{n+1-i} - \xi_1 \left(-\sum_{i=1}^{n} \xi_i x^{n-i} \right) \ . \end{aligned}$$

If we assume that we know the coefficients of the presentation

$$x^k \ = \ \sum_{i=0}^{n-1} a_{ki} x^i$$

we get the coefficients $a_{k+1,i}$ of the presentation for x^{k+1} immediately from

$$x^{k+1} \ = \ \sum_{i=1}^{n-1} a_{k,i-1} x^i \ - \ \sum_{i=0}^{n-1} a_{k,n-1} \xi_{n-i} x^i \ ,$$

hence

$$a_{k+1,i} \ = \ a_{k,i-1} - a_{k,n-1} \xi_{n-i} \ \ (0 \leq i \leq n-1 \, , \ a_{k,-1} := 0) \ .$$

Examples

(1) Let $R = \mathbb{Z}/8\mathbb{Z}$ and $f(t) = t^2 - 1 \in R[t]$. $\Lambda = R[t]/f(t)R[t]$ has an R–basis $1 + f(t)R[t]$, $t + f(t)R[t] =: x$. On the other hand, R itself is a solution ring and we therefore have ring homomorphisms

$\varphi \ : \ \Lambda \to R$ via $t + f(t)R[t] \mapsto \alpha$ with $\alpha \in \{1, 3, 5, 7\}$.

This situation is enlightened by the following diagram:

$$\begin{array}{ccc} a(1 + f(t)R[t]) + b(t + f(t)R[t]) & \mapsto & a + \alpha\, b \\ (a(1 + f(t)R[t]) + b(t + f(t)R[t]))(c(1 + f(t)R[t]) + d(t + f(t)R[t])) & \mapsto & (a + \alpha\, b)\,(c + \alpha\, d) \\ \| & & \| \\ ac(1 + f(t)R[t]) + (ad + bc)(t + f(t)R[t]) + bd(1 + f(t)R[t]) & & (ac + \alpha^2 bd) + \alpha\,(ad + bc) \\ \| & & \| \\ (ac + bd)(1 + f(t)R[t]) + (ad + bc)(t + f(t)R[t]) & & (ac + bd) + \alpha\,(ad + bc) \ . \end{array}$$

(2) If $f(t) = t^3 + pt^2 + qt + r \in R[t]$ has a zero x in Λ then the polynomial $f(t)$ decomposes in $\Lambda[t]$, one factor being $t - x$. Dividing f in $\Lambda[t]$ by $t - x$ we obtain

$$f(t) \ = \ (t - x)\,(t^2 + (p + x)\,t + q + x\,(x + p)) \ .$$

Comparing coefficients we get the following relation for r in Λ: $r = -x\,(x\,(x + p) + q)$.

(3) Let us assume that $t^2 + m$ is irreducible over R (for example, $m = 1$, $R = \mathbb{R}$ or $m = -2$, $R = \mathbb{Z}/5\mathbb{Z}$). Then $\Lambda := R[t]/f(t)R[t]$ has a basis $1, x := t + f(t)R[t]$. Therefore we have $R[t]/f(t)R[t] \cong R \times R$. But what does the ring structure on $R \times R$ look like? Addition is clearly done coordinatewise. Multiplication needs to be transfered from Λ, however. Because of $x^2 = -m$ we obtain

$$\begin{aligned} (a + bx) \cdot (c + dx) &= ac - mbd + (bc + da)x \quad \text{, hence} \\ (a,\, b) \cdot (c,\, d) &= (ac - mbd,\, bc + da) \end{aligned}$$

in $R \times R$.

Corollary 3.17. *By a* $(\deg(f)-1)$*–fold application of the construction in 3.16 we obtain a ring* $S(f, R)$ *(***splitting ring** *of* f *over* R*) with* $\deg(f)!$ *basis elements over* R*.*

Proposition 3.18. *Let* F *be a field and* $f(t) \in F[t]$ *of positive degree. Then there exists an extension field* E *of* F *in which* f *has a zero.*

Proof In $F[t]$ the polynomial f splits into a product of irreducible polynomials. We assume that g is such an irreducible factor. In case $\deg(g) = 1$ the polynomial g (and therefore f) has a zero in F. Now let us assume that $\deg(g) > 1$. Since $F[t]$ is a principal ideal ring the ideal $g(t)F[t]$ is maximal. Hence, $E := F[t]/g(t)F[t]$ is a field. According to 3.16 the polynomial g (and therefore f) has a zero in E.
□

Theorem 3.19. *Let* F *be a field and* $f(t) \in F[t]$ *of positive degree. Then there exists an extension field* E *of* F *such that* f *splits in* $E[t]$ *into a product of linear factors:*

$$f(t) = l(f) \prod_{i=1}^{\deg(f)} (t - x_i) \quad (x_i \in E) \quad .$$

Hence, all zeros of f *are contained in* E*.*

Proof We do this by induction over the degree $n := \deg(f)$ of f. For $n = 0$ the polynomial f is constant and equals its leading coefficient. In this case, the product over the monic linear factors is empty. By induction hypothesis we assume that the theorem is true for all polynomials of degree less than or equal to n. Let $f(t) \in F[t]$ be a polynomial of degree $n+1$. By 3.18 there exists an extension field E_1 of F in which f has a zero, say x_1. In $E_1[t]$ the polynomial $f(t)$ therefore

splits into two factors: $f(t) = (t - x_1)\, g(t)$, with a polynomial g of degree n and with $l(f) = l(g)$. According to our induction assumption there exists an extension field E of E_1 so that g splits in $E[t]$ into a product of linear factors

$$g(t) = l(g) \prod_{i=2}^{n+1} (t - x_i) \quad (x_2, ..., x_{n+1} \in E) \quad ,$$

hence,

$$f(t) = l(f) \prod_{i=1}^{n+1} (t - x_i) \quad .$$

□

Example Let $f(t) \in F[t]$ be monic. Then f splits in $E[t]$ with zeros $x_i \in E$ in the following way:

$$\begin{aligned} f(t) &= \prod_{i=1}^{n} (t - x_i) \\ &= (t - x_1)\,(t - x_2)\, \ldots\, (t - x_n) \\ &= t^n - t^{n-1} \sum_{i=1}^{n} x_i + t^{n-2} \sum_{i<j} x_i\, x_j \\ &\quad + \ldots + (-1)^{n-k} t^k \sum_{1 \le i_1 < i_2 < \ldots < i_k \le n} x_{i_1} \cdot \ldots \cdot x_{i_k} \\ &\quad + \ldots + (-1)^n\, x_1 \cdot \ldots \cdot x_n \quad . \end{aligned}$$

This will be used in the next section when we discuss elementary symmetric functions.

4. Symmetric Polynomials and the Fundamental Theorem of Algebra

Definition 4.1. *Let R be a unital commutative ring. A polynomial $f(\mathbf{t}) \in R[\mathbf{t}]$ $(\mathbf{t} = (t_1, ..., t_n))$ is called* **symmetric** *if it satisfies $f(\mathbf{t}) = f(t_{\pi(1)}, ..., t_{\pi(n)})$ for all $\pi \in \mathfrak{S}_\mathbf{n}$. Special symmetric polynomials are*

$$\begin{aligned} \sigma_0(\underline{t}) &:= 1 \\ \sigma_j(\underline{t}) &:= \sum_{1 \le i_1 < i_2 < \ldots < i_j \le n} t_{i_1} \cdot \ldots \cdot t_{i_j} \quad (1 \le j \le n) \end{aligned}$$

which are called **elementary symmetric functions** *in $t_1, ..., t_n$. Here we shortly wrote σ_j instead of the more precise $\sigma_j^{(n)}$.*

Examples

(1) Among the elementary symmetric functions

$$\begin{aligned}\sigma_1(\mathbf{t}) &= t_1 + \ldots + t_n \ , \\ \sigma_2(\mathbf{t}) &= t_1t_2 + \ldots + t_1t_n + t_2t_3 + \ldots + t_2t_n + ... + t_{n-1}t_n \ , \\ \sigma_n(\mathbf{t}) &= t_1 \ldots t_n\end{aligned}$$

are used most frequently.

(2) Let $f(\mathbf{t}, t) \in R[t_1, ..., t_n, t]$ be monic of degree n with $f(\mathbf{t}, t_i) = 0 \ \ (1 \leq i \leq n)$ and R be a factorial ring. Then we obtain

$$\begin{aligned}f(\mathbf{t}, t) &= \prod_{i=1}^{n}(t - t_i) \\ &= \sum_{j=0}^{n}(-1)^{n-j}\,\sigma_{n-j}(\mathbf{t})\,t^j \\ &= \sum_{i=0}^{n}(-1)^{i}\,\sigma_i(\mathbf{t})\,t^{n-i} \ .\end{aligned}$$

(3) Specializing the variables $t_{l+1}, ..., t_n$ to zero we obtain symmetric polynomials

$$\sigma_k^{(l)}(t_1, \ldots, t_n) \ := \ \sigma_k(t_1, \ldots, t_l, 0, \ldots, 0) \ \ (1 \leq k \leq l) \ .$$

of the same degree. For $k > l$ such a specialization yields zero.

Remark The symmetric polynomials form a subring of $R[t_1, ..., t_n]$. The specialization

$$\Phi_{\underline{\sigma}} \ : \ R[t_1, ..., t_n] \to R[t_1, ..., t_n] \ : \ f \mapsto f(\sigma_1, ..., \sigma_n)$$

maps f onto a symmetric polynomial.

Theorem 4.2. *(Principal theorem for elementary symmetric functions) Let R be a factorial ring. Then every symmetric polynomial $f(\mathbf{t}) \in R[\mathbf{t}]$ can be written as $g(\sigma_1, ..., \sigma_n)$ for a uniquely determined polynomial $g(\mathbf{t}) \in R[\mathbf{t}]$.*

Proof We introduce a **weight** w for monomials. For

$$g(\mathbf{t}) = a\,t_1^{m_1} \cdot \ldots \cdot t_n^{m_n} = a\mathbf{t}^{\underline{m}}$$

we set $w(t_k) = k, \ w(g) := \sum_{i=1}^{n} i\,m_i$. Accordingly the **weight of a polynomial** is defined as the maximum of the weights of the occuring monomials. For

$$f(\mathbf{t}) \ = \ \sum_{\underline{\nu} \in (\mathbb{Z}^{\geq 0})^n} a_{\underline{\nu}}\mathbf{t}^{\underline{\nu}}$$

we set

$$w(f) = \max\{w(\mathbf{t}^{\underline{\nu}}) \mid a_{\underline{\nu}} \neq 0\} \ .$$

Hence, the weight of a polynomial $w(f(\mathbf{t}))$ is exactly the degree of the polynomial $f(\sigma_1, \ldots, \sigma_n)$.

Proof of the existence of a polynomial g.
We show that for every symmetric polynomial $f(\mathbf{t}) \in R[\mathbf{t}]$ of degree d there exists a polynomial $g(\mathbf{t}) \in R[\mathbf{t}]$ of weight $w(g) \leq d$ such that $f(\mathbf{t}) = g(\sigma_1, ..., \sigma_n)$. The proof is by induction on n.
For $n = 1$ we can put $f = g$ because we have $\sigma_1 = t_1$ in this case.
Hence, we assume that we have shown the theorem for polynomials in $n-1$ variables. To obtain the result also for polynomials in n variables we apply induction on the degree d of f.
For $d \leq 0$ the polynomial f is constant and $g = f$ does the job. We now assume that $d > 0$ and that the statement is true for polynomials of degree less than d. Let f be an arbitrary polynomial of degree d. Specializing $t_n \mapsto 0$ in f we obtain $f^{(n-1)}(t_1, ..., t_{n-1})$ of degree at most d. According to our induction hypothesis about the number of variables there exists a polynomial $g_1(t_1, ..., t_{n-1}) \in R[t_1, ..., t_{n-1}] \subseteq R[t_1, ..., t_n]$ of weight $\leq d$ with

$$\begin{aligned} f(t_1, ..., t_{n-1}, 0) &= g_1(\sigma_1^{(n-1)}, ..., \sigma_{n-1}^{(n-1)}) \\ &\quad (\sigma_i^{(n-1)} := \sigma_i(t_1, ..., t_{n-1}, 0)) \ . \end{aligned}$$

We put $h(\mathbf{t}) := f(\mathbf{t}) - g_1(\sigma_1, ..., \sigma_{n-1})$. The polynomial h then is again symmetric of degree $\leq d$. We have $h(t_1, ..., t_{n-1}, 0) = 0$, the polynomial h is therefore divisible by t_n. h being symmetric it is then divisible by all t_i $(1 \leq i \leq n)$. Hence, $R[\mathbf{t}]$ being a factorial ring, h is divisible by σ_n. We therefore get $h(\mathbf{t}) = \sigma_n h_1(\mathbf{t})$ with h_1 again being symmetric. We either have $h_1 = 0$ or $\deg(h_1) = \deg(h) - n < d$. According to our induction assumption on d there exists $g_2 \in R[t_1, ..., t_n]$ of weight $\leq d - n$ with $h_1(t_1, ..., t_n) = g_2(\sigma_1, ..., \sigma_n)$. Putting things together we obtain $f(\mathbf{t}) = g(\sigma_1, ..., \sigma_n)$ for

$$g(\sigma_1, ..., \sigma_n) \ = \ g_1(\sigma_1, ..., \sigma_n) \ + \ \sigma_n \, g_2(\sigma_1, ..., \sigma_n) \ .$$

Proof of uniqueness of g.
To prove uniqueness we show that for $f(\mathbf{t}) \in R[\mathbf{t}]$ with $f(\sigma_1, ..., \sigma_n) = 0$ we must have $f = 0$. The proof is by induction on the number n of variables. For $n = 1$ the statement is trivial because of $\sigma_1 = t_1$. Now we assume that uniqueness is guaranteed for polynomials of at most $n - 1$ variables. We let $0 \neq f(\mathbf{t}) \in R[\mathbf{t}]$ be a symmetric polynomial in n variables and of minimal degree with $f(\sigma_1, ..., \sigma_n) = 0$. We write

f as a polynomial in t_n: $f(t_1,...,t_n) = \sum_{i=0}^{k} f_i(t_1,...,t_{n-1})\, t_n^i$. The coefficient $f_0(t_1,...,t_{n-1})$ cannot be zero. Otherwise the polynomial f would be divisible by t_n and therefore – as we saw above – by σ_n in contradiction to our degree assumption. Specializing $t_n \mapsto 0$ we obtain that f_0 is symmetric in $n-1$ variables with

$$0 = f(\sigma_1^{(n-1)},...,\sigma_n^{(n-1)},0) = f_0(\sigma_1^{(n-1)},...,\sigma_{n-1}^{(n-1)}) \ ,$$

again a contradiction to our induction assumption.

□

Example For $f(t_1,t_2,t_3) = (t_1-t_2)^2\,(t_1-t_3)^2\,(t_2-t_3)^2$ we have

$$\begin{aligned} f(t_1,t_2,0) &= (t_1-t_2)^2\,(t_1t_2)^2 \\ &= ((\sigma_1^{(2)})^2 - 4\,\sigma_2^{(2)})\,(\sigma_2^{(2)})^2 \ ; \end{aligned}$$

and consequently $h(\mathbf{t}) = \sigma_3\, h_1(\mathbf{t})$ with

$$\begin{aligned} h_1(t_1,t_2,0) &= 18\,\sigma_1^{(2)}\,\sigma_2^{(2)} - 4\,(\sigma_1^{(2)})^3 \\ h_2(\mathbf{t}) &= h_1(\mathbf{t}) - 18\,\sigma_1^{(3)}\,\sigma_2^{(3)} + 4\,\sigma_1^{(3)} \\ &= -27\,t_1t_2t_3 \\ &= -27\,\sigma_3 \ . \end{aligned}$$

Putting things together we obtain

$$f(t_1,t_2,t_3) = \sigma_1^2\,\sigma_2^2 - 4\,\sigma_2^3 + \sigma_3\,(18\,\sigma_1\,\sigma_2 - 4\,\sigma_1^3 - 27\,\sigma_3) \ .$$

Besides the elementary symmetric functions there is another set of symmetric polynomials, the so–called **power sums**:

$$S_k := S_k(\mathbf{t}) := \sum_{i=0}^{n} t_i^k \ \ (k \in \mathbb{Z}^{\geq 0}) \ .$$

Calculations with them are usually easier than with the σ_i. However, a transfer from power sums to elementary symmetric functions is generally possible only in characteristic zero, as we will see below.

Theorem 4.3. *The power sums S_k and the elementary symmetric functions σ_j are connected via* **Newton's relations***:*

(1)

$$\sum_{i=0}^{k-1} (-1)^i\,\sigma_i(\mathbf{t})\,S_{k-i}(\mathbf{t}) + k\,(-1)^k\,\sigma_k(\mathbf{t}) = 0 \quad (0 \leq k \leq n) \ ,$$

(2)

$$\sum_{i=0}^{n}(-1)^i\,\sigma_i(\mathbf{t})\,S_{k-i}(\mathbf{t})=0\quad(k\geq n)\ .$$

Proof The polynomial

$$f(t_1,...,t_n,t):=\sum_{j=0}^{n}(-1)^j\,\sigma_j(\mathbf{t})\,t^{n-j}=\prod_{j=1}^{n}(t-t_j)$$

in $n+1$ variables $t_1,...,t_n,t$ satisfies

$$0=\sum_{j=0}^{n}(-1)^j\,\sigma_j(\mathbf{t})\,t_i^{n-j}\quad(1\leq i\leq n)\ ,$$

respectively,

$$0=\sum_{j=0}^{n}(-1)^j\,\sigma_j(\mathbf{t})\,t_i^{k-j}\quad(1\leq i\leq n,\ k\geq n)\ .$$

Summing up these n equations yields

$$\sum_{j=0}^{n}(-1)^j\,\sigma_j(\mathbf{t})\,S_{k-j}(\mathbf{t})=0\ ,$$

hence 2., respectively 1. in case $k=n$. The remaining part of 1. will now be proved for fixed k via induction on the number of variables n. For the initial value $n=k$ we have already proved it. Therefore we assume that $n>k$ and that the theorem is true for $n-1$ variables. We put

$$F(t_1,...,t_n):=\sum_{i=0}^{k-1}(-1)^i\,\sigma_i(\mathbf{t})\,S_{k-i}(\mathbf{t})+(-1)^k\,k\,\sigma_k(\mathbf{t})\ .$$

F is certainly a symmetric function of degree $\leq k$ and because of $k<n$ also less than n. By induction assumption we have $F(t_1,...,t_{n-1},0)=0$. Hence, $F(\mathbf{t})$ is divisible by t_n – and since it is symmetric – also by $\sigma_n(\mathbf{t})$. Because of $\deg(F)<n$ the polynomial F must therefore be 0.

□

Example We list the first few of Newton's relations:

$$\begin{aligned}
S_1(\mathbf{t}) &= \sigma_1(\mathbf{t}),\\
S_2(\mathbf{t}) &= \sigma_1(\mathbf{t})\,S_1(\mathbf{t})-2\,\sigma_2(\mathbf{t})\\
&= \sigma_1^2(\mathbf{t})-2\,\sigma_2(\mathbf{t}),\\
S_3(\mathbf{t}) &= \sigma_1(\mathbf{t})\,S_2(\mathbf{t})-\sigma_2(\mathbf{t})\,S_1(\mathbf{t})+2\,\sigma_3(\mathbf{t})\\
&= \sigma_1^3(\mathbf{t})-3\,\sigma_1(\mathbf{t})\,\sigma_2(\mathbf{t})+3\,\sigma_3(\mathbf{t})\ .
\end{aligned}$$

If the natural numbers are no zerodivisors in R then we can also express the σ_k by the S_k over $\mathfrak{Q}(R)$:

$$\begin{aligned}
\sigma_1(\mathbf{t}) &= S_1(\mathbf{t}), \\
\sigma_2(\mathbf{t}) &= \frac{1}{2}\left(S_1(\mathbf{t})^2 - S_2(\mathbf{t})\right), \\
\sigma_3(\mathbf{t}) &= \frac{1}{3}\left(S_2(\mathbf{t}) - S_1(\mathbf{t})^3 + 3\,S_1(\mathbf{t})\,\frac{1}{2}\left(S_1(\mathbf{t})^2 - S_2(\mathbf{t})\right)\right) \\
&= \frac{1}{6}\left(2\,S_3(\mathbf{t}) + S_1(\mathbf{t})^3 - 3\,S_1(\mathbf{t})\,S_2(\mathbf{t})\right) \ .
\end{aligned}$$

Definition 4.4. *The polynomial*

$$d(f) := a_0^{2n-2} \prod_{1\le i<j\le n} (t_i - t_j)^2$$

of $R[\mathbf{t}]$ is called **discriminant** *of the polynomial*

$$f(t) \;=\; a_0 \prod_{i=1}^{n}(t - t_i) \;\in R[\mathbf{t}][t] \ .$$

The exponent of a_0 is chosen minimal such that $d(f)$ belongs to $R[\mathbf{t}]$. This will be shown in the next proposition.

Example A monic quadratic polynomial $t^2 + at + b \in \mathbb{R}[t]$ has the zeros $x_{1,2} = -\frac{a}{2} \pm \sqrt{\frac{a^2-4b}{4}}$. Its discriminant is therefore $(x_1 - x_2)^2 = a^2 - 4b$. We note that the sign of the discriminant decides whether both zeros are real or complex. The discriminant vanishes if and only if both zeros coincide.

Proposition 4.5. *The discriminant of the polynomial*

$$f(t) \;=\; \sum_{i=0}^{n} a_i t^{n-i} \;=\; a_0 \prod_{i=1}^{n}(t - x_i)$$

satisfies

$$a_0\, d(f) \;=\; (-1)^{\binom{n}{2}} \operatorname{res}(f, f') \ .$$

The discriminant $d(f)$ is an element of R.

Proof The derivative of the given polynomial is

$$f'(t) \;=\; a_0 \sum_{i=1}^{n} \prod_{\substack{j=1 \\ j\neq i}}^{n} (t - x_j) \ .$$

Hence, we obtain

$$f'(x_i) \;=\; a_0 \prod_{\substack{j=1\\ j\neq i}}^{n}(x_i - x_j)$$

and can easily calculate the resultant of f and f':

$$\begin{aligned}
\operatorname{res}(f,f') &= a_0^{n-1}\prod_{i=1}^{n} f'(x_i)\\
&= a_0^{n-1}\prod_{i=1}^{n}\left(a_0\prod_{\substack{j=1\\ j\neq i}}^{n}(x_i-x_j)\right)\\
&= a_0^{2n-1}\prod_{i=1}^{n}\left(\prod_{j>i}(x_i-x_j)\prod_{j<i}(-(x_j-x_i))\right)\\
&= a_0^{2n-1}(-1)^{\sum_{i=1}^{n}(i-1)}\prod_{1\le i<j\le n}(x_i-x_j)^2\\
&= (-1)^{\binom{n}{2}}a_0\,d(f)\;\;.
\end{aligned}$$

To prove the second statement we consider $\operatorname{res}(\frac{f}{a_0},\frac{f'}{a_0})$ as determinant of a $(2n-1)\times(2n-1)$ matrix. The first column of that matrix has entries of $\{0,1,n\}$. The remaining $2n-2$ columns contain – besides zeros – entries of the form

$$\frac{a_i}{a_0}\;\;(1\le i\le n)\;\text{ or }\;(n-i)\,\frac{a_i}{a_0}\;\;(1\le i<n)\;\;.$$

Hence, upon multiplication with a_0^{2n-2} that resultant belongs to the ring R.

□

For readers familiar with matrices and determinants we give the following proposition as an exercise.

Proposition 4.6. *In $R[\mathbf{t}]$ the discriminant satisfies*

$$d(f) = a_0^{2n-2}\,\det\left((S_{i+j-2}(\mathbf{t})_{1\le i,j\le n}\right)\;\;.$$

(Hint: The proof makes use of Vandermonde's determinant.)

Theorem 4.7. (Fundamental Theorem of Algebra)
Every polynomial $f(t)\in\mathbb{C}[t]$ of positive degree n has a zero in $\mathbb{C}$. This

is tantamount to the statement that f can be factored in $\mathbb{C}[t]$ into linear factors:

$$f(t) = l(f)\prod_{j=1}^{n}(t - x_j) \quad (x_j \in \mathbb{C}) \ .$$

We also say that $\mathbb{C}$ is **algebraically closed** *since all elements which are algebraic over $\mathbb{C}$ already belong to $\mathbb{C}$.*

Proof

(1) In a first step the statement is reduced to polynomials with real coefficients. For $f(t) \in \mathbb{C}[t]$ we form the product of $f(t)$ and its complex conjugate $\overline{f(t)}$ which is obtained from f by applying complex conjugation to every coefficient. Since that product is invariant under complex conjugation it must be contained in $\mathbb{R}[t]$:

$$g(t) := f(t)\overline{f(t)} \ \in \mathbb{R}[t] \ .$$

Assuming that the statement is true for polynomials in $\mathbb{R}[t]$ we get

$$g(t) := |l(f)|^2 \prod_{j=1}^{2n}(t - c_j) \ .$$

Because of the preceding remark with c_j also $\overline{c_j}$ is a zero of $g(t)$. We order the c_j such that $c_{n+j} = \overline{c_j} \ \ (1 \leq j \leq n)$. This yields

$$f(t) = l(f)\prod_{j=1}^{n}(t - c_j) \quad (1 \leq j \leq n) \ .$$

(2) We need to show that every polynomial of positive degree in $\mathbb{R}[t]$, say $f(t) = t^n + a_1 t^{n-1} + \ldots + a_n \ \in \mathbb{R}[t]$, has a zero in $\mathbb{C}$. It is remarkable that all known proofs for this are not purely algebraic inasmuch as they require some elements from analysis, usually a form of the intermediate value theorem. The intermediate value theorem tells us that a polynomial of $\mathbb{R}[t]$ of odd degree – being interpreted as a continuous function from $\mathbb{R}$ to $\mathbb{R}$ – has a zero in $\mathbb{R}$. Hence, we can assume that $\deg(f) = 2^k q$ with $k \in \mathbb{N}$, $q \in \mathbb{N}$ odd. The proof for the latter is by induction on k. For the initial value $k = 0$ we already saw this. Hence, we assume that $k > 0$ and that the statement is true for degrees of f divisible at most by 2^{k-1}. According to 3.19 there

is an extension field K of $\mathbb{R}$ over which $f(t)$ decomposes into a product of linear factors:

$$f(t) = l(f) \prod_{j=1}^{n} (t - x_j) \ .$$

Following an idea of Laplace we consider the polynomials

$$L_r(t) := \prod_{1 \le \mu < \nu \le n} (t - x_\mu - x_\nu - r\, x_\mu x_\nu) \ \in K[t] \ \ (r \in \mathbb{R}) \ .$$

We note that the coefficients of $L_r(t)$ are symmetric functions in the zeros x_j. Hence, the coefficients of L_r can be written as polynomials in the elementary symmetric functions $\sigma_j(\mathbf{x}) = (-1)^j a_j$ implying $L_r(t) \in \mathbb{R}[t]$. Then

$$\deg(L_r) = \frac{n}{2}\,(n-1) = 2^{k-1} q\,(2^k q - 1) = 2^{k-1}\tilde{q} \ ,$$

$\tilde{q}$ is odd, and L_r has a zero z_r in $\mathbb{C}$ for every $r \in \mathbb{R}$ according to our induction assumption. For every $r \in \mathbb{R}$ there exist indices μ, ν with

$$z_r := x_\mu + x_\nu + r\, x_\mu x_\nu \ \in \mathbb{C} \ .$$

Since the number of pairs of indices μ, ν is finite, the number of parameters $r \in \mathbb{R}$ is infinite there must exist $r \neq \tilde{r}$ in R, $1 \le \mu < \nu \le n$ with

$$x_\mu + x_\nu + r\, x_\mu x_\nu,\ x_\mu + x_\nu + \tilde{r}\, x_\mu x_\nu \ \in \mathbb{C} \ .$$

This has the consequences $x_\mu x_\nu \in \mathbb{C}$, $x_\mu + x_\nu \in \mathbb{C}$, i.e. x_μ, x_ν are zeros of

$$t^2 - (x_\mu + x_\nu)\, t + x_\mu x_\nu \ \in \mathbb{C}[t] \ ,$$

and therefore also x_μ, x_ν belong to $\mathbb{C}$. Here we use that square roots of complex numbers again belong to $\mathbb{C}$. Hence, $f(t)$ has at least 2 zeros in $\mathbb{C}$.

□

Corollary 4.8. *Every polynomial $f(t) \in \mathbb{R}[t]$ of positive degree can be decomposed as*

$$f(t) = l(f) \prod_{i=1}^{k} (t - c_i) \prod_{j=1}^{l} q_j(t)$$

with $c_i \in \mathbb{R}$, $q_j(t) = t^2 + u_j t + v_j \in \mathbb{R}[t]$ irreducible $(1 \le j \le l))$. This presentation is unique up to the order of the factors.

Proof We let $c_1, ..., c_k$ denote all real zeros of $f(t)$. Then we obtain

$$f(t) = l(f) \prod_{i=1}^{k} (t - c_i)\, g(t) \quad ,$$

where the polynomial $g(t) \in \mathbb{R}[t]$ is uniquely determined. Then for every zero $x \in \mathbb{C}$ of $g(t)$ also $\overline{x}$ is a zero. We therefore order the zeros of g appropriately to obtain $g(t) = \prod_{j=1}^{l} (t - z_j)(t - \overline{z_j})$. Then we put

$$q_j(t) = t^2 - (z_j + \overline{z_j})\, t + z_j \overline{z_j} \ \in \mathbb{R}[t] \ \ (1 \leq j \leq l) \quad .$$

The uniqueness of that presentation – up to the order of the factors – is a consequence of $\mathbb{R}[t]$ being a factorial ring..

□

Corollary 4.9. *The irreducible elements of $\mathbb{C}[t]$ are the polynomials of degree one. The irreducible elements of $\mathbb{R}[t]$ are the polynomials of degree one and those polynomials $t^2 + ut + v$ of degree two with $u^2 - 4v < 0$.*

The next theorem is an appendix for readers already familiar with vector spaces.

Theorem 4.10. *Let Λ be a unital entire commutative overring of $\mathbb{R}$ in which every element is algebraic over $\mathbb{R}$. Then Λ is isomorphic either to $\mathbb{R}$ or to $\mathbb{C}$.*

Proof Let us assume that $\Lambda \neq \mathbb{R}$. For $x \in \Lambda \setminus \mathbb{R}$ we obtain a 2–dimensional $\mathbb{R}$–vectorspace $V := \mathbb{R}1 + \mathbb{R}x$. Also there exists $f(t) \in \mathbb{R}[t]$ of positive degree with $f(x) = 0$. Because of the fundamental theorem the minimal polynomial of $x \in \Lambda \setminus \mathbb{R}$ is necessarily of degree 2, say

$$g(t) = t^2 + ut + v \ \in \mathbb{R}[t] \quad (u^2 - 4v < 0) \quad .$$

This implies $x^2 = -ux - v$ in Λ. Therefore we can introduce a multiplication in V via

$$\begin{aligned} (a + bx)\,(c + dx) \ &:= \ ac + (ad + bc)\, x + (-ux - v)\, bd \\ &= \ (ac - vbd) + (ad + bc - ubd)\, x \quad . \end{aligned}$$

Thus V becomes a commutative unital entire overring of $\mathbb{R}$ with a 2–element basis. We show that V is isomorphic to $\mathbb{C}$ via

$$a + bx \mapsto a + \frac{b}{2}\,(-u + iD) \quad \text{for} \quad D = \sqrt{4v - u^2} \quad .$$

That mapping is a priori surjective and injective and also additive. Its multiplicativity follows from the diagram

$$\begin{array}{ccc}
(a+bx)\,(c+dx) & \mapsto & \left(a+\frac{b}{2}\,(-u+iD)\right)\,\left(c+\frac{d}{2}\,(-u+iD)\right) \\
\| & & \| \\
(ac-bdv)+x\,(bc+ad-ubd) & & ac+(\frac{ad}{2}+\frac{bc}{2})(-u+iD)+\frac{bd}{4}\,(u^2-2uDi-D^2) \\
\downarrow & & \| \\
ac-bdv+(bc+ad-ubd)\,\frac{1}{2}\,(-u+iD) & & ac-\frac{u}{2}\,(ad+bc-bdu)-bdv+\frac{i}{2}\,D\,(ad+bc-bdu)
\end{array}$$

We still need to prove that $\Lambda = V$. For this we let $y \in \Lambda \setminus \mathbb{R}$ arbitrary, $f(t) \in \mathbb{R}[t]$ with $f(y) = 0$. Making use of $V \cong \mathbb{C}$ we conclude that f decomposes into linear factors $t - \lambda \;\; (\lambda \in V)$, so we get $y = \lambda$ for a suitable choice of λ.

□

Remark The last theorem shows that any $\mathbb{R}$–vectorspace V of dimension r cannot be a field for $r > 2$.

5. Multivariate Polynomials and Gröbner Bases

To make the presentation easier we assume in this paragraph that all polynomials have coefficients in a base field F. We emphasize, however, that all concepts which we develop can be generalized to polynomials with coefficients in a Noetherian ring R. In any case, every ideal in the ring of polynomials is finitely generated. The goal of this section is to develop an algorithm for the computation of special sets of generators for arbitrary ideals, so-called Gröbner bases. They have turned out to be one of the strongest tools in computer algebra. For example, using Gröbner bases it is easy to decide whether a polynomial belongs to a given ideal. Another application is to the solution of non linear systems of algebraic equations.

We recall several notations about multivariate polynomials. Usually, we will consider polynomials in n variables, i.e. from $F[t_1, ..., t_n]$ which we abbreviate by $F[\mathbf{t}]$. Any polynomial $f(\mathbf{t})$ is a finite sum of monomials $m(\mathbf{t}) = a \prod_{i=1}^{n} t_i^{m_i} =: a_{\mathbf{m}} \mathbf{t}^{\mathbf{m}}$. The sum $m_1 + ... + m_n$ is called the **degree** of the monomial $m(\mathbf{t})$. If the coefficient $a \in F$ of $m(\mathbf{t})$ is one the monomial is called monic. We note that the least common multiple lcm and the greatest common divisor gcd of two monic monomials $m(\mathbf{t})$, $k(\mathbf{t})$ are given by

$$\mathrm{lcm}(m,k) \;=\; \prod_{i=1}^{n} t_i^{\max(m_i,k_i)} \;,\; \gcd(m,k) \;=\; \prod_{i=1}^{n} t_i^{\min(m_i,k_i)} \quad .$$

Analogously to the case of univariate polynomials we would like to put the monomials of a polynomial into a specific order. Of course, we can do this with respect to their degrees, but in case $n > 1$ there

exist monic monomials of the same degree which do not coincide. For example, we must decide whether $t_1^2 t_2$ or $t_1 t_2^2$ should come first.

This means to introduce a total ordering on the set $\mathcal{S}$ of all monic monomials. Clearly, once the variables have been fixed we can identify each monic monomial with its vector $\mathbf{m} = (m_1, ..., m_n)$ of exponents. This establishes a monoid isomorphism between the multiplicative monoid $\mathcal{S}$ and the additive monoid $(\mathbb{Z}^{\geq 0})^n$. The ordering to be chosen should be compatible with the law of composition of the monoid, i.e. we require that for elements α, β, γ of the monoid the ordering $\alpha > \beta$ implies $\alpha\gamma > \beta\gamma$. Also the property that every non-zero subset of the monoid contains a minimal element will be useful. This element is then unique since we requested a total ordering.

In practice, the following orderings on $\mathcal{S}$ have turned out to be of special interest.

(1) **Lexicographical Ordering** $>_{lex}$
For elements of $(\mathbb{Z}^{\geq 0})^n$ we say that $\mathbf{m} = (m_1, ..., m_n)$ is bigger than $\mathbf{k} = (k_1, ...k_n)$ if there is a smallest index, say i, such that $m_j = k_j$ for $1 \leq j < i$ and $m_i > k_i$.
For example, we have $(1, 2, 0) >_{lex} (0, 3, 4)$ and $(3, 2, 4) >_{lex} (3, 2, 1)$.

(2) **Inverse Lexicographical Ordering** $>_{ilex}$
For elements of $(\mathbb{Z}^{\geq 0})^n$ we say that $\mathbf{m} = (m_1, ..., m_n)$ is bigger than $\mathbf{k} = (k_1, ...k_n)$ if there is a largest index, say i, such that $m_j = k_j$ for $i < j \leq n$ and $m_i > k_i$.
For example, we have $(4, 7, 4) >_{ilex} (4, 2, 3)$ and $(5, 1, 3) >_{ilex} (4, 1, 3)$.

(3) **Graded Lexicographical Ordering** $>_{glex}$
For elements of $(\mathbb{Z}^{\geq 0})^n$ we say that $\mathbf{m} = (m_1, ..., m_n)$ is bigger than $\mathbf{k} = (k_1, ...k_n)$ if either $m_1 + ... + m_n > k_1 + ... + k_n$ or $m_1 + ... + m_n = k_1 + ... + k_n$ and $\mathbf{m} >_{lex} \mathbf{k}$. For example, we have $(1, 2, 3) >_{glex} (3, 2, 0)$ and $(1, 2, 4) >_{glex} (1, 1, 5)$.

(4) **Graded Inverse Lexicographical Ordering** $>_{gilex}$
For elements of $(\mathbb{Z}^{\geq 0})^n$ we say that $\mathbf{m} = (m_1, ..., m_n)$ is bigger than $\mathbf{k} = (k_1, ...k_n)$ if either $m_1 + ... + m_n > k_1 + ... + k_n$ or $m_1 + ... + m_n = k_1 + ... + k_n$ and $\mathbf{m} >_{ilex} \mathbf{k}$.
For example, we have $(4, 7, 1) >_{gilex} (4, 2, 3)$ and $(1, 4, 3) >_{gilex} (4, 1, 3)$.

It is straightforward that all three orderings are total orderings of $(\mathbb{Z}^{\geq 0})^n$. Also the compatibility of these orderings with addition is immediate. We leave it as an exercise to the reader to show that every non-empty subset of $(\mathbb{Z}^{\geq 0})^n$ has a minimal element. We note that

the inverse lexicographical ordering is used to look at the elements of $F[\mathbf{t}]$ as polynomials in the variable t_n with coefficients in $F[t_1, ..., t_{n-1}]$ (recursive representation).

Since every polynomial f is a finite sum of monomials any (total) ordering of the monomials can be used to introduce a (partial) ordering of the polynomials. Especially, we can define the **leading monomial (leading term)** $\mathrm{lt}(f)$ as the largest monomial $a_{\mathbf{m}}\mathbf{t}^{\mathbf{m}}$ occuring in the presentation of f. We denote the corresponding monic part $\mathbf{t}^{\mathbf{m}}$ by $\mathrm{mlt}(f)$ (**monic leading term**) and the coefficient $a_{\mathbf{m}}$ by $\mathrm{lc}(f)$ (**leading coefficient**) of f. The partial ordering on $F[\mathbf{t}]$ is then obtained via

$$f > g \;\Leftrightarrow\; \mathrm{mlt}(f) > \mathrm{mlt}(g) \quad .$$

With these prerequisites at hand we turn our interest to computations in a polynomial ideal $\mathbf{I}$. We assume that it is given by a finite number of generators, say $f_1, ..., f_k$. Then every element g of $\mathbf{I}$ can be written as

$$g \;=\; \sum_{i=1}^{k} r_i f_i \quad (r_i \in F[\mathbf{t}]) \quad .$$

We are interested in elements of small degree of $\mathbf{I}$ since they will play a decisive role for Gröbner bases. In case r_i is not constant the degree of $r_i f_i$ is larger then the degree of f_i. We can therefore expect g to be of small degree only if the sum of the leading monomials of several summands in the presentation of g is zero. If we just consider two polynomials instead of k this phenomenon can be enforced in the following way.

Definition 5.1. *For two polynomials $f, g \in F[\mathbf{t}]$ we define their* **S-polynomial** *$S(f, g)$ as*

$$S(f,g) \;:=\; \frac{\mathrm{lcm}(\mathrm{mlt}(f), \mathrm{mlt}(g))}{\mathrm{lt}(f)} f \;-\; \frac{\mathrm{lcm}(\mathrm{mlt}(f), \mathrm{mlt}(g))}{\mathrm{lt}(g)} g \quad .$$

Hence, the leading term of the S-polynomial $S(f, g)$ is smaller than the least common multiple of the leading terms of f and g. This property will be of importance in a characterization of Gröbner bases in 5.5.

Examples

1. We calculate the S-polynomial of $f = t_1^3 t_2^2 - t_1^2 t_2^3 + t_1$ and $g = 3t_1^4 t_2 + t_2^2$

in $\mathbb{Q}[\mathbf{t}]$ with respect to $>_{glex}$.

$$\begin{aligned} S(t_1^3t_2^2 - t_1^2t_2^3 + t_1, 3t_1^4t_2 + t_2^2) &= \frac{t_1^4t_2^2}{t_1^3t_2^2}\, f \;-\; \frac{t_1^4t_2^2}{3t_1^4t_2}\, g \\ &= t_1 f \;-\; \frac{1}{3}t_2 g \\ &= -t_1^3t_2^3 + t_1^2 - \frac{1}{3}t_2^3 \end{aligned}$$

This computation remains valid for $>_{lex}$.
2. For the polynomials $f = t_1^2 - t_2$ and $g = t_1^3 - t_3$ of $\mathbb{Q}[\mathbf{t}]$ we compute their S-polynomial for two different orderings.
(a) $t_1 > t_2 > t_3$ (lexicographic ordering)

$$S(f,g) = t_1 f - g = -t_1t_2 + t_3 \;\; .$$

(b) $t_2 > t_3 > t_1$

$$S(f,g) = t_3 f - t_2 g = -t_2t_1^3 + t_1^2t_3 \;\; .$$

The following lemma will also be used in characterizing Gröbner bases in 5.5.

Lemma 5.2. *Let $f, f_1, ..., f_s \in F[\mathbf{t}]$ and $f = \sum_{i=1}^{s} c_i f_i \;\; (c_i \in F)$ with $\delta = \mathrm{mlt}(f_1) = ... = \mathrm{mlt}(f_s) \neq \mathrm{mlt}(f)$. Then f is also an F-linear combination of the $S(f_i, f_{i+1}) \;\; (1 \leq i < s)$.*

Proof Since the monic leading terms of all f_i coincide the monic leading term of $\sum_{i=1}^{s} c_i f_i$ either equals δ or it is smaller. According to our assumption $\delta \neq \mathrm{mlt}(f)$ we must therefore have $\sum_{i=1}^{s} c_i \,\mathrm{lt}(f_i) \;=\; 0$. We let $\mathrm{lt}(f_i) = a_i\delta \;\; (a_i \in F^{\times})$ and set $b_i := a_i c_i, \;\; p_i := f_i / \,\mathrm{lc}(f_i)$ for $(1 \leq i \leq s)$ and obtain

$$\sum_{i=1}^{s} b_i \;=\; 0 \;\; ,$$

$$S(f_i, f_j) = \frac{\delta}{a_i\delta}\, f_i \;-\; \frac{\delta}{a_j\delta}\, f_j = \frac{f_i}{a_i} - \frac{f_j}{a_j} = S(p_i, p_j)$$

and eventually

$$\begin{aligned}\sum_{i=1}^{s} c_i f_i &= \sum_{i=1}^{s} b_i(f_i/\operatorname{lc}(f_i)) \\ &= b_1(p_1-p_2)+(b_1+b_2)(p_2-p_3)+...+ \\ &\quad (b_1+...+b_{s-1})(p_{s-1}-p_s)+(b_1+...+b_s)p_s \\ &= \sum_{i=1}^{s-1}(\sum_{j=1}^{i} b_j)S(f_i,f_{i+1}) \ .\end{aligned}$$

□

We note that we have $\operatorname{mlt}(S(f_i,f_j)) < \delta$ in the preceding lemma.

If the leading term of a non-zero polynomial g divides the leading term of a polynomial f ($\operatorname{lt}(g) \mid \operatorname{lt}(f)$), i.e. there exists a monomial h with $\operatorname{lt}(f) = h\operatorname{lt}(g)$) then we can subtract hg from f so that this divisibility property is no longer satisfied for the polynomials $f-hg$ and g. We note that $f-hg = S(f,g)$. This remains valid even without the divisibility condition if we set $h=0$ in case $\operatorname{lt}(g)$ does not divide $\operatorname{lt}(f)$. Repeatedly replacing f by $S(f,g)$ in case $\operatorname{lt}(g)$ divides $\operatorname{lt}(f)$ until this divisibility condition does not hold anymore we say that the polynomial f is **reduced modulo** g. This concept can be easily generalized to the reduction of a polynomial modulo a non-empty finite set of non-zero polynomials.

Definition 5.3. *Let* $\mathbf{I}$ *be a non-zero ideal of* $F[t_1,...,t_n]$ *with ordered basis* $G = \{g_1,...,g_k\}$. *We say that an element* $f \in F[t_1,...,t_n]$ **reduces to zero modulo** G *if the sequence* $f_0 = f$,

$$f_i = \textit{reduction of } f_{i-1} \textit{ modulo } g_i \ \ (1 \le i \le k)$$

satisfies $f_k = 0$.

We remark that a polynomial which reduces to 0 modulo G necessarily belongs to the ideal $\mathbf{I}$. However, not every element of an ideal must have this property.

Example Let $f = t_1t_2^2 - t_1$ and $G = \{g_1, g_2\}$ with $g_1 = t_1t_2 + 1$, $g_2 = t_2^2 - 1$. Then we obtain $f_1 = f - t_2(t_1t_2+1) = -t_1 - t_2 = f_2$, and f does not reduce to 0 modulo $\{g_1,g_2\}$. If we change the order of the basis elements, however, we compute $f_1 = f - t_1(t_2^2-1) = 0$ and f does reduce to 0 modulo that newly ordered basis. The reason for this phenomenon is that the basis $\{g_1,g_2\}$ is not a Gröbner basis (see

definition and theorem below). It does not contain the S-polynomial

$$S(g_1, g_2) = \frac{t_1 t_2^2}{t_1 t_2}(t_1 t_2 + 1) - \frac{t_1 t_2^2}{t_2^2}(t_2^2 - 1) = t_1 + t_2 \ .$$

If we add $g_3 := t_1 + t_2$ to the basis we get $G = \{g_1, g_2, g_3\}$ and, clearly, f reduces to 0 modulo G. We note that the same holds for the S-polynomials $S(g_1, g_3)$, $S(g_2, g_3)$.

Definition 5.4. *Let* $\mathbf{I}$ *be a non-zero ideal of* $F[t_1, ..., t_n]$ *with basis* G. *If every* $f \in \mathbf{I}$ *reduces to zero modulo* G *then* G *is called a* **Gröbner basis** *of* $\mathbf{I}$.

We remark that Gröbner bases are by no means unique since every superset of a Gröbner basis also satisfies the condition of the definition.

Theorem 5.5. *Let* $\mathbf{I}$ *be a non-zero ideal of* $F[t_1, ..., t_n]$ *with basis* $G = \{g_1, ..., g_s\}$. *Then* G *is a Gröbner basis for* $\mathbf{I}$ *if and only if every S-polynomial* $S(g_i, g_j)$ $(1 \leq i < j \leq s)$ *reduces to 0 modulo* G.

Proof If G is a Gröbner basis of $\mathbf{I}$ then every polynomial of $\mathbf{I}$, hence a priori every $S(g_i, g_j)$, reduces to 0 modulo G.

Now let us assume that every S-polynomial $S(g_i, g_j)$ reduces to 0 modulo G but that there exists $f \in \mathbf{I}$ which does not. Obviously, f is not zero. If $0 \neq f$ cannot be reduced modulo G anymore then in the basis representation

$$f = \sum_{i=1}^{s} h_i g_i \ \ (h_i \in F[t_1, ..., t_n]) \tag{1}$$

the leading monomial of f is not divisible by any of the leading monomials of the g_i. We assume that (1) is a presentation of f in which the monic part of the largest occuring monomial on the right-hand side is as small as possible, say δ. We observe that $\mathrm{mlt}(f) \neq \delta$. Then we rewrite (1) by separating those summands with monic leading monomial δ from the other terms. We put $J_1 := \{i \mid 1 \leq i \leq s, \mathrm{mlt}(h_i g_i) = \delta\}$ and $J_2 := \{1, ..., s\} \setminus J_1$ and obtain

$$\begin{aligned} f &= \sum_{i \in J_1} h_i g_i + \sum_{i \in J_2} h_i g_i & (2) \\ &= \sum_{i \in J_1} lt(h_i) g_i + \sum_{i \in J_1} (h_i - \mathrm{lt}(h_i)) g_i + \sum_{i \in J_2} h_i g_i & (3) \end{aligned}$$

so that the leading monomials of the summands in the two last sums are smaller than δ. Because of $\mathrm{mlt}(f) \neq \delta$ the first sum satisfies the

prerequisites of the preceding lemma for the polynomial

$$f \; - \; \left(\sum_{i \in J_1} (h_i - \mathrm{lt}(h_i)) g_i + \sum_{i \in J_2} h_i g_i \right)$$

if we put $f_i = \mathrm{mlt}(h_i) g_i$ and $c_i = \mathrm{lc}(h_i)$. For $i, j \in J_1$ we set $g_{ij} := \mathrm{lcm}(\mathrm{mlt}(g_i), \mathrm{mlt}(g_j))$ and observe that $g_{ij} > \mathrm{mlt}(S(g_i, g_j))$. Hence, the first sum of (3) becomes an F-linear combination of S-polynomials of the form

$$\begin{aligned} S(\mathrm{mlt}(h_i) g_i, \mathrm{mlt}(h_j) g_j) &= \frac{\delta \, \mathrm{mlt}(h_i) g_i}{\mathrm{mlt}(h_i) \, \mathrm{lt}(g_i)} \; - \; \frac{\delta \, \mathrm{mlt}(h_j) g_j}{\mathrm{mlt}(h_j) \, \mathrm{lt}(g_j)} \\ &= \frac{\delta}{\mathrm{lt}(g_i)} g_i \; - \; \frac{\delta}{\mathrm{lt}(g_j)} g_j \\ &= \frac{\delta}{g_{ij}} S(g_i, g_j) \; . \end{aligned}$$

Since the S-polynomials $S(g_i, g_j)$ reduce to zero modulo G they have a basis presentation

$$S(g_i, g_j) \; = \; \sum_{\nu=1}^{s} h_\nu g_\nu$$

with $\mathrm{mlt}(h_\nu g_\nu) \; \leq \; \mathrm{mlt}(S(g_i, g_j))$. Because of $\mathrm{mlt}(\frac{\delta}{g_{ij}} S(g_i, g_j)) < \delta$ inserting those presentations into (3) yields a presentation of f by G in which all occuring monomials are smaller than δ contradicting our assumption.

□

The following algorithm (Buchberger's algorithm) constructs a Gröbner basis from an arbitrary ideal basis.

Buchberger Algorithm

Input A basis $G = \{g_1, ..., g_s\}$ of an ideal $\mathbf{I}$.
Output A Gröbner basis $G = \{g_1, ..., g_t\}$ of $\mathbf{I}$.
Initialization Set $t := s$, $B := \{(i, j) \mid 1 \leq i < j \leq t\}$.
Step If $B \neq \emptyset$ choose (i, j) from B and remove (i, j) from B; reduce $S(g_i, g_j)$ modulo G to f; if $f \neq 0$ add $g_{t+1} := f$ to G and $\{(i, t+1) \mid 1 \leq i \leq t\}$ to B and increase t by 1.

We still need to show that Buchberger's algorithm terminates. For this we consider the sequence of ideals $\mathbf{I}_s := \langle \mathrm{mlt}(g_i) \mid 1 \leq i \leq s \rangle$. We show that every enlargement of G (increase of s) yields a strictly larger ideal $\mathbf{I}_s$. Since any ascending chain of ideals becomes stationary s is bounded.

Let us therefore assume that f is an S-polynomial of two elements of G which is already reduced modulo G but is still non-zero. Hence, f will be inserted into G thus increasing $\sharp G$. We will show that $\mathrm{mlt}(f)$ is not contained in $\mathbf{I}_s$. Namely, if $\mathrm{mlt}(f)$ belongs to $\mathbf{I}_s$ there exists a presentation

$$\mathrm{mlt}(f) \;=\; \sum_{i=1}^{s} h_i \,\mathrm{mlt}(g_i)$$

with polynomials $h_i \in F[\mathbf{t}]$. Comparing monomials on both sides we get a non-empty subset J_1 of $\{1, 2, ..., s\}$ and monomials $a_i\mathbf{t}^{m_i}$ which are summands of h_i such that

$$\mathrm{mlt}(f) \;=\; \sum_{i\in J_1} a_i\mathbf{t}^{\mathbf{m_i}} \,\mathrm{mlt}(g_i) \quad .$$

But then $\mathrm{mlt}(f)$ is a multiple of $\mathrm{mlt}(g_i)$ for $i \in J_1$ and f can be further reduced modulo g_i contradicting our assumption.

We remark that the previous considerations also show that every monomial which is contained in an ideal with a basis of monomials is a linear combination of monomials each summand being divisible by one of the basis elements.

6. Multivariate Polynomials – Resultants

Besides Gröbner bases there is another important tool for eliminating variables in a system of polynomial equations: resultants. We introduce them in a generic way, i.e. we assume that their coefficients are algebraically independent over $\mathbb{Z}$. Let $R = \mathbb{Z}[a_0, ..., a_n, b_0, ..., b_m]$ be a polynomial ring in $n + m + 2$ variables. Then the polynomials

$$\begin{aligned} A(t) &= a_0t^n + a_1t^{n-1} + ... + a_{n-1}t + a_n \;\text{ and} \\ B(t) &= b_0t^m + b_1t^{m-1} + ... + b_{m-1}t + b_m \end{aligned} \tag{4}$$

of $R[t]$ are said to be **generic** inasmuch as any two polynomials f, g over a unital commutative ring Λ with $\deg(f) \le n$, $\deg(g) \le m$ can be obtained as homomorphic images of A, B by mapping

$$1_{\mathbb{Z}} \mapsto 1_\Lambda,\; t \mapsto t,\; a_i \mapsto \alpha_i \in \Lambda,\; b_j \mapsto \beta_j \in \Lambda \;\; (0 \le i \le n,\, 0 \le j \le m)$$

for suitable elements α_i, β_j of Λ. If S denotes a common splitting ring of A, B over R we obtain

$$A(t) = a_0 \prod_{i=1}^{n}(t - x_i),\;\; B(t) = b_0 \prod_{j=1}^{m}(t - y_j) \tag{5}$$

in $S[t]$. From this we conclude

$$\frac{a_i}{a_0} \;=\; (-1)^i \sum_{1\le j_1<\ldots<j_i\le n} x_{j_1}...x_{j_i} \;=:\; (-1)^i\sigma_i \;\; (1\le i\le n) \qquad (6)$$

with the σ_i being symmetric functions in the zeros of A (so-called elementary symmetric functions). Analogously, we get

$$\frac{b_i}{b_0} \;=\; (-1)^i \sum_{1\le j_1<\ldots<j_i\le m} y_{j_1}...y_{j_i} \;\; (1\le i\le m) \;\; . \qquad (7)$$

It follows that $\mathbb{Z}[a_0,...,a_n,b_0,...,b_m]\subseteq\mathbb{Z}[a_0,x_1,...,x_n,b_0,y_1,...,y_m]$ and therefore also $a_0,x_1,...,x_n,b_0,y_1,...,y_m$ are algebraically independent.

From the theorem of Gauß we know that $R[t]$ is a unique factorization domain. Hence, the greatest common divisor of A,B is well defined. Any common zero of A,B is also a zero of $\gcd(A,B)$ and vice versa. Whereas the zeros of A,B usually do not belong to R, the greatest common divisor $\gcd(A,B)$ is calculated in $R[t]$. Hence, the existence of common zeros can be decided without the need of constructing ring extensions of R.

Lemma 6.1. *The greatest common divisor of $A,B\in R[t]$ given in (4) is different from 1, if and only if there exist non-zero polynomials $U,V\in R[t]$ satisfying $\deg(U)<m$, $\deg(V)<n$ and $UA=VB$.*

Proof If $C:=\gcd(A,B)$ is different from 1 we write $A=C\tilde{A}$, $B=C\tilde{B}$ with $\deg(\tilde{A})<n$, $\deg(\tilde{B})<m$ and obtain

$$C\tilde{A}\tilde{B} \;=\; \tilde{B}A \;=\; \tilde{A}B$$

so that we can choose $U=\tilde{B}$, $V=\tilde{A}$.

If U,V with the properties of the lemma exist we consider the factorizations of UA and of VB into prime polynomials. Clearly, not every prime polynomial dividing A can divide V because of $\deg(V)<\deg(A)$. Therefore at least one such prime polynomial must divide B and consequently $\gcd(A,B)$.

□

Setting

$$U(t)=\sum_{i=0}^{m-1} u_i t^{m-1-i}\,,\; V(t)=\sum_{j=0}^{n-1} v_j t^{n-1-j} \;\in R[t] \qquad (8)$$

the equation $UA=VB$ yields a linear system of equations for the coefficients u_i, v_j. For the coefficent of t^μ $(0\le\mu\le m+n-1)$ we

obtain

$$\sum_{\max\{0,\mu-n\}\le\nu\le\min\{\mu,m-1\}} u_\nu a_{\mu-\nu} \;=\; \sum_{\max\{0,\mu-m\}\le\nu\le\min\{\mu,n-1\}} v_\nu b_{\mu-\nu} \;\;.$$

These are $(m+n)$ equations. Moving everything to the left-hand side we get a homogenous linear system of equations

$$(u_0,...,u_{m-1},-v_0,...,-v_{n-1})\begin{pmatrix} a_0 & ... & a_n & & & & & \\ & a_0 & ... & a_n & & & & \\ & & & ... & & & & \\ & & & ... & & & & \\ & & & & a_0 & a_1 & ... & a_n \\ b_0 & b_1 & . & . & . & b_m & & \\ & b_0 & b_1 & . & . & . & b_m & \\ & & & & & & ... & \\ & & & & b_0 & b_1 & ... & b_m \end{pmatrix} = \mathbf{0} \qquad (9)$$

We denote the formidable coefficient matrix by Δ. We recall that A, B have a common zero if and only if that homogenous system has a non-trivial solution. Of course, the latter holds exactly for $\det(\Delta) = 0$.

Definition 6.2. *The determinant of the coefficient matrix Δ of (9) is called* **resultant** *of A and B. It is denoted by res(A,B).*

Next we show that $\mathrm{res}(A,B)$ is in the ideal of $R[t]$ generated by A and B. This is of importance when we use resultants for eliminating variables from multivariate polynomials. Looking at the structure of Δ we immediately find the following identity:

$$\mathbf{v} := \begin{pmatrix} t^{m-1}A \\ t^{m-2}A \\ . \\ . \\ t^0 A \\ t^{n-1}B \\ . \\ . \\ t^0 B \end{pmatrix} \;=\; \Delta \begin{pmatrix} t^{n+m-1} \\ t^{n+m-2} \\ . \\ . \\ t^n \\ t^{n-1} \\ . \\ . \\ t^0 \end{pmatrix} \quad . \qquad (10)$$

Denoting the columns of Δ by $\mathbf{d}_1,...,\mathbf{d}_{n+m}$ this is tantamount to

$$\mathbf{v} \;=\; \sum_{\kappa=1}^{n+m} t^{n+m-\kappa}\mathbf{d}_\kappa \;\;.$$

Then Cramer's rule tells us that

$$\mathrm{res}(A,B) = \det(\Delta)t^0 = \det(\mathbf{d}_1,...,\mathbf{d}_{m+n-1},\mathbf{v}) \;\;.$$

Calculating the last determinant we note that the variable t only occurs in the last column $\mathbf{v}$ so that we indeed obtain polynomials ϕ, $\psi \in R[t]$ with $\deg(\phi) < m$, $\deg(\psi) < n$ and

$$\phi A + \psi B = \operatorname{res}(A, B) \ . \tag{11}$$

We note that in the first m rows of Δ we have entries zero or coefficients of A and in the last n rows the entries are zero or coefficients of B. According to Laplace's theorem we have

$$\operatorname{res}(A, B) = \sum_{\pi \in S_{m+n}} \operatorname{sign}(\pi)\Delta(1, \pi(1))...\Delta(m+n, \pi(m+n)) \tag{12}$$

if $\Delta(i, j)$ denotes the entry of Δ in row i and column j. Therefore any non-zero summand of the sum in (12) must consist of m factors a_μ (from the first m rows) and n factors b_ν (from the last n rows). We conclude that $\operatorname{res}(A, B)$ is a homogenous polynomial of degree m in the a_μ and of degree n in the b_ν. We can write it in the form

$$\operatorname{res}(A, B) = a_0^m b_0^n F(\frac{a_1}{a_0}, ..., \frac{a_n}{a_0}, \frac{b_1}{b_0}, ..., \frac{b_m}{b_0}) \ .$$

According to our remark on the elementary symmetric functions of the zeros of A, respectively B, we know that F can also be written as a polynomial in the variables $x_1, ..., x_n, y_1, ..., y_m$ which we again denote by F. Since the resultant vanishes if zeros of A and B coincide and since $\mathbb{Z}[a_0, x_1, ..., x_n, b_0, y_1, ..., y_m]$ is a factorial ring the polynomial F must be divisible by the polynomials $x_i - y_j$ $(1 \leq i \leq n,\ 1 \leq j \leq m)$ and therefore by the polynomial

$$\tilde{F} := \prod_{i=1}^{n} \prod_{j=1}^{m} (x_i - y_j) \ .$$

Lemma 6.3. *The resultant res(A, B) of the generic polynomials A, B of (4) coincides with each of the three polynomials*

(1) $a_0^m b_0^n \tilde{F}$,
(2) $a_0^m \prod_{i=1}^{n} B(x_i)$,
(3) $(-1)^{mn} b_0^n \prod_{j=1}^{m} A(y_j)$.

Proof The equality of the polynomials in the lemma is immediate from (5) :

$$a_0^m b_0^n \tilde{F} = a_0^m \prod_{i=1}^{n} \left(b_0 \prod_{j=1}^{m} (x_i - y_j) \right) = (-1)^{mn} b_0^n \prod_{j=1}^{m} \left(a_0 \prod_{i=1}^{n} (y_j - x_i) \right) \ .$$

We also know that $a_0^m b_0^n \tilde{F}$ divides $\operatorname{res}(A, B)$. From 6.3(2.) we conclude that $a_0^m b_0^n \tilde{F}$ is homogenous of degree n in the b_ν and from (3.) that

it is homogenous of degree m in the a_μ. Since $\mathrm{res}(A,B)$ has the same properties the quotient of $\mathrm{res}(A,B)$ and $a_0^m b_0^n \tilde{F}$ must be constant. The constant will be determined by comparing the coefficients of $a_0^m b_m^n$. In $\det(\Delta)$ we obtain that monomial as the product of all diagonal elements of Δ, it has coefficient 1. But looking at its coefficient in $a_0^m b_0^n \tilde{F}$ we find from 6.3(2.) that it is 1, too.

□

We list a few direct consequences of the last lemma which will be of help in the actual computation of resultants:

$$\begin{aligned} \mathrm{res}(A,B) &= (-1)^{mn}\mathrm{res}(B,A)\ , \\ \mathrm{res}(rA,B) &= r^m \mathrm{res}(A,B)\ , \\ \mathrm{res}(A,rB) &= r^n \mathrm{res}(A,B)\ \ (r \in R)\ . \end{aligned} \tag{13}$$

If one or even both polynomials involved are constant we get

$$\begin{aligned} \mathrm{res}(a_0,B) &= a_0^m\ , \\ \mathrm{res}(A,b_0) &= b_0^n\ , \\ \mathrm{res}(a_0,b_0) &= 1\ . \end{aligned} \tag{14}$$

Hence, we will try to evaluate $\mathrm{res}(A,B)$ by pseudo-division. In case $\deg(B) > \deg(A)$ we compute $\mathrm{res}(A,B) = (-1)^{mn}\mathrm{res}(B,A)$. Hence, we may assume that $\deg(A) \geq \deg(B)$. Applying pseudo-division we get polynomials $Q = Q(A,B)$, $R = R(A,B) \in R[t]$, $\deg(R) < \deg(B)$ satisfying

$$b_0^{n-m+1} A \ = \ Q\, B \ + \ R\ . \tag{15}$$

Then the last lemma yields

$$\begin{aligned} \mathrm{res}(A,B) &= (-1)^{mn} b_0^n \prod_{j=1}^{m} A(y_j) \\ &= (-1)^{mn} b_0^{n-m(n-m+1)} \prod_{j=1}^{m} (Q\,B + R)(y_j) \\ &= (-1)^{mn} b_0^{n-\deg(R)-m(n-m+1)} \left(b_0^{\deg(R)} \prod_{j=1}^{m} R(y_j) \right) \\ &= (-1)^{m(n-\deg(R))} b_0^{n-\deg(R)-m(n-m+1)} \mathrm{res}(R,B) \\ &= (-1)^{mn} b_0^{n-\deg(R)-m(n-m+1)} \mathrm{res}(B,R)\ . \end{aligned} \tag{16}$$

We note that the exponent of b_0 is likely to become negative so that these calculations can only be carried out in the quotient field of R.

However, since the resultant itself is an element of R we are guaranteed that the final result will not contain denominators.

In the actual calculation of the resultant of two polynomials we successively replace the polynomial of larger degree via pseudodivision by a polynomial whose degree is less than the original lower degree. In this way we eventually obtain a constant remainder polynomial. If that constant is zero, the original resultant is zero, too. Otherwise the last resultant is evaluated by (14). This leads to the following algorithm.

Algorithm for computing resultants

Input $A, B \in R[t]$ with $\deg(A) \geq \deg(B) > 0$.
Output $\mathrm{res}(A,B) \in R$ and polynomials $\phi, \psi \in R[t]$ satisfying $\phi A + \psi B = \mathrm{res}(A,B)$.
Step 1 (Initialization) Set $\mathrm{res}(A,B) \leftarrow 1$, $F \leftarrow A$, $G \leftarrow B$, $N \leftarrow \deg(A)$, $M \leftarrow \deg(B)$, $\phi_0 \leftarrow 1$, $\psi_1 \leftarrow 1$, $\phi_1 \leftarrow 0$, $\psi_0 \leftarrow 0$.
Step 2 (Pseudo-division) Set $b_0 \leftarrow lc(G)$ and calculate with (15) polynomials $Q = \sum_{i=0}^{N-M} q_i t^{N-M-i}$, $R \in R[t]$. We set $s \leftarrow N - \deg(R)$ and $\mathrm{res}(A,B) \leftarrow \mathrm{res}(A,B)(b_0^{N-M+1})^{-M} b_0^s (-1)^{MN}$ and also $\phi_2 \leftarrow b_0^{N-M+1}\phi_0 - Q\phi_1$, $\psi_2 \leftarrow b_0^{N-M+1}\psi_0 - Q\psi_1$. If R is constant go to 4., else to 3..
Step 3. (Interchange of F, G) Set $F \leftarrow G$, $G \leftarrow R$, $N \leftarrow M$, $M \leftarrow \deg(R)$ as well as $\phi_0 \leftarrow \phi_1$, $\phi_1 \leftarrow \phi_2$, $\psi_0 \leftarrow \psi_1$, $\psi_1 \leftarrow \psi_2$ and go to 2..
Step 4. (Termination) For $R = 0$ set $\mathrm{res}(A,B) = 0$ and $\phi \leftarrow \phi_2$, $\psi \leftarrow \psi_2$; for $R \neq 0$ set $T \leftarrow \mathrm{res}(A,B)R^{M-1}$ and $\mathrm{res}(A,B) \leftarrow TR$, $\phi \leftarrow \phi_2 T$, $\psi \leftarrow \psi_2 T$. Then terminate.

Remarks

(1) The polynomials ϕ, ψ of 11 satisfy $\deg(\phi) \leq \deg(B)-1$, $\deg(\psi) \leq \deg(A) - 1$. The example $A = t^2 + 1$, $B = t^2 + 4$ shows that equality need not hold:

$$9 = \mathrm{res}(A,B) = -3(t^2+1) + 3(t^2+4) \ .$$

(2) Instead of operating in the quotient field of R we can keep track of the multipliers b_0 and their exponents in each step separately and calculate their product only at the end knowing that $\mathrm{res}(A,B)$ belongs to R.

Example We want to compute $\mathrm{res}(t^3+1, 2t^2-2)$ in $\mathbb{Z}[t]$. In the steps of the algorithm the following data are produced:
1. $F = t^3+1$, $N = 3$, $G = 2t^2-2$, $M = 2$.
2. $2^2(t^3+1) = 2t(2t^2-2) + 4t + 4$, hence $Q = 2t$, $R = 4t+4$ yielding $s = 2$, $\mathrm{res}(A,B) = (2^2)^{-2}2^2 = 2^{-2}$.
3. $F = 2t^2-2$, $N = 2$, $G = 4t+4$, $M = 1$.

2. $4^2(2t^2-2) = (8t-8)(4t+4)+0$, hence $Q = 8t-8$, $R = 0$ yielding $s = 2$, $\mathrm{res}(A,B) = 2^{-2}(4^2)^{-1}4^2 = 2^{-2}$.
4. $\mathrm{res}(A,B) = 0$, $\phi = 32(-t+1)$, $\psi = 16(t^2-t+1)$.