A. Analysis

The first section of this appendix fixes some notations for multi indices that are widely used in the main text. Except the distinction into multi and "repeated" indices these are standard notations in multivariate analysis. The next section recalls those properties of analytic functions that are needed in the proof of the Cauchy–Kovalevskaya Theorem. Then some elementary operations with differential equations like reduction to first order or quasi-linearisation are discussed within the context of the formal theory. Finally, a few facts about the modified Stirling numbers that appear in manipulations of Cartan characters and Hilbert functions are presented.

A.1 Multi Indices

We use mostly the standard multi index notation of multivariate analysis. If we are dealing with n variables $x^1, \dots, x^n$, a *multi index* is an ordered n-tuple $\mu = [\mu_1, \dots, \mu_n] \in \mathbb{N}_0^n$ of integers (we always use square brackets to denote multi indices). The sum of its entries is often called its *length* written $|\mu| = \mu_1 + \cdots + \mu_n$. Its factorial is defined by $\mu! = \mu_1! \cdots \mu_n!$. The expression $(x - y)^\mu$ is a short hand for $(x^1 - y^1)^{\mu_1} \cdots (x^n - y^n)^{\mu_n}$. For a function $\phi(x^1, \dots, x^n)$ we write

$$\frac{\partial^{|\mu|}\phi}{\partial x^\mu} = \frac{\partial^{|\mu|}\phi}{\partial (x^1)^{\mu_1} \cdots \partial (x^n)^{\mu_n}} \,. \tag{A.1}$$

For convenience, we define for the special multi index $0 = [0, \dots, 0]$ that $x^0 = 1$ and $\partial^{|0|}\phi/\partial x^0 = \phi$. Obviously, $0! = 1$ and $|0| = 0$. Other special multi indices are

$$\ell_i = [0, \dots, 0, \ell, 0, \dots, 0] \tag{A.2}$$

where $\ell \in \mathbb{N}$ is the ith entry and all other entries vanish.

The addition of multi indices is defined componentwise, i. e.

$$\mu + \nu = [\mu_1 + \nu_1, \dots, \mu_n + \nu_n] \,. \tag{A.3}$$

With this addition the set $\mathbb{N}^n$ of all multi indices acquires the structure of an Abelian monoid. If we want to increase by one the ith entry of a multi index μ, we can now simply write $\mu + 1_i$ using (A.2).

If k is the smallest value such that $\mu_k \neq 0$, we call it the *class* of the multi index μ: $\operatorname{cls}\mu = k$. In (A.1) this means that there are no differentiations with respect to $x^1, \dots, x^{\operatorname{cls}\mu - 1}$. In principle, the class of the multi index $[0,\dots,0]$ is undefined, but in many situations it is convenient to set $\operatorname{cls}[0,\dots,0] = n$.

Some elementary combinatorics leads to the following results:

$$\bigl|\{\mu \in \mathbb{N}_0^n \mid |\mu| = q\}\bigr| = \binom{n+q-1}{q}\,, \tag{A.4a}$$

$$\bigl|\{\mu \in \mathbb{N}_0^n \mid |\mu| = q \wedge \operatorname{cls}\mu = k\}\bigr| = \binom{n-k+q-1}{q-1}\,, \tag{A.4b}$$

$$\bigl|\{\mu \in \mathbb{N}_0^n \mid 0 \leq |\mu| \leq q\}\bigr| = \binom{n+q}{q}\,. \tag{A.4c}$$

A multi index μ of length q and class k is obtained by first setting μ_k to a value $1 \leq q' \leq q$ and then adding an arbitrary multi index in $\mathbb{N}_0^{n-k}$ of length $q - q'$. Thus the second line arises from summing up the first one from 0 to $q-1$ (the possible values of $q - q'$) after replacing n by $n-k$ and using a classical identity for binomial coefficients. In a similar manner, the third line stems from summing up the first one from 0 to q and applying the same identity.

On a few occasions we use an alternative multi index notation, which we call for lack of a better name *repeated index* notation. We distinguish it from multi indices by parentheses instead of brackets. A repeated index ρ of length q always contains exactly q entries with values between 1 and n: $\rho = (\rho_1, \dots, \rho_q)$. The analogue to (A.1) is

$$\frac{\partial^q \phi}{\partial x^\rho} = \frac{\partial^q \phi}{\partial x^{\rho_1} \cdots \partial x^{\rho_q}}\,. \tag{A.5}$$

As we always assume that partial derivatives commute, there is obviously no need to order the entries of a repeated index: any permutation of the entries leaves the index unchanged.

A.2 Real-Analytic Functions

Much of the formal theory is based on formal power series. If such a series converges, it defines an analytic function. In this appendix we collect some basic facts about these functions. Throughout we consider real-valued functions defined in some domain $\Omega \subset \mathbb{R}^n$.

Definition A.2.1. *The function* $f : \Omega \to \mathbb{R}$ *is called* real-analytic *at the point* $x \in \Omega$*, if there exists a neighbourhood* $x \in U \subseteq \Omega$ *such that for all* $y \in U$ *the function* f *can be represented by a power series*

$$f(y) = \sum_{|\mu|=0}^{\infty} a_\mu (y-x)^\mu \,. \tag{A.6}$$

f *is real-analytic in the whole domain* Ω*, if it is real-analytic at every point* $x \in \Omega$*. The set of all functions that are real-analytic in the domain* Ω *is denoted by* $C^\omega(\Omega)$*.*

Obviously, any analytic function is smooth: $C^\omega(\Omega) \subset C^\infty(\Omega)$. The converse is, however, not true. Consider the function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f : x \mapsto \begin{cases} 0 & x \leq 0\,, \\ e^{-\frac{1}{x^2}} & x > 0\,. \end{cases} \tag{A.7}$$

It is easy to see that f is smooth on $\mathbb{R}$. But it is not analytic at the origin. Analyticity is a rather strong property, as the following proposition implies that if we know an analytic function and all its derivatives at one point of Ω, we know it in all Ω. This is sometimes called *unique continuation.*

Proposition A.2.2. *Let* f *and* g *be two functions that are real-analytic in the domain* Ω*. If there exists a point* $x \in \Omega$ *such that*

$$\frac{\partial^{|\mu|} f}{\partial x^\mu}(x) = \frac{\partial^{|\mu|} g}{\partial x^\mu}(x) \qquad \forall \mu \in \mathbb{N}_0^n\,, \tag{A.8}$$

then $f \equiv g$ *in* Ω*.*

Proof. We use a simple topological argument. Consider the set

$$\Sigma = \left\{ y \in \Omega \mid \frac{\partial^{|\mu|} f}{\partial x^\mu}(y) = \frac{\partial^{|\mu|} g}{\partial x^\mu}(y) \quad \forall \mu \in \mathbb{N}_0^n \right\}\,. \tag{A.9}$$

Σ is relatively closed in Ω, as it is defined by equations. On the other hand, Σ must be open, because at any point $x \in S$ the two functions f and g possess identical power series and therefore must be identical in an open neighbourhood. As, by definition, a domain is a connected open set, we must have $\Sigma = \Omega$. □

Another important and characteristic property of analytic functions is the existence of a bound on their derivatives. For a smooth function f defined in a neighbourhood of an arbitrary but fixed point $x \in \Omega$, we say that $f \in C_{M,r}(x)$, if its derivatives at x satisfy *Cauchy's estimate*

$$\left|\frac{\partial^{|\mu|} f}{\partial x^\mu}(x)\right| \leq M \frac{|\mu|!}{r^{|\mu|}} \qquad \forall \mu \in \mathbb{N}^n\,. \tag{A.10}$$

Proposition A.2.3. *Let* $f \in C^\infty(\Omega)$*. Then* $f \in C^\omega(\Omega)$*, if and only if for every compact subset* $\Sigma \subset \Omega$ *there exist two numbers* M *and* r *such that* $f \in \mathcal{C}_{M,r}(x)$ *for all* $x \in \Sigma$*.*

Proof. Let us first assume that f is analytic. Then it is not difficult to show that for all $x \in \Omega$ its derivatives satisfy in a neighbourhood $\mathcal{U}(x)$ the estimate (A.10) with constants $M(x)$ and $r(x)$ (see e.g. [149, Section 2.2.1]). As any compact subset Σ can be covered by a finite number of such neighbourhoods, we can take the maximum of the $M(x)$ and the minimum of the $r(x)$ and obtain the assertion.

For the converse, take a point $x \in \Omega$ and choose as compact set Σ a closed ball of radius R with its centre at x (obviously, the radius R must be chosen small enough so that $\Sigma \subset \Omega$). Then, by assumption, there exist numbers M and r such that $f \in \mathcal{C}_{M,r}(x)$. Consider an arbitrary point $y \in \Sigma$ with $\delta = \sum_{i=1}^n |x_i - y_i| < \min(r, R)$ and introduce for $0 \leq \tau \leq 1$ the function $F(\tau) = f\bigl(x + \tau(y - x)\bigr)$. We apply Taylor's Theorem to $F(\tau)$ and obtain

$$f(y) = F(1) = \sum_{k=0}^{q-1} \frac{1}{k!}\frac{d^kF}{d\tau^k}(1) + \frac{1}{q!}\frac{d^qF}{d\tau^q}(\bar{\tau}) \tag{A.11}$$

for some value $0 \leq \bar{\tau} \leq 1$. A straightforward computation yields for any order $q \geq 0$ the following estimate of the remainder term

$$\Bigl|\frac{1}{q!}\frac{d^qF}{d\tau^q}(\bar{\tau})\Bigr| \leq M\left(\frac{\delta}{r}\right)^q . \tag{A.12}$$

As by construction $\delta < r$, it tends to zero and we find that f can be represented in a neighbourhood of x by a convergent series of the form (A.6). □

Finally, we recall that an analytic function f is *majorised* in a neighbourhood of the point x by another analytic function g, written $f \preceq g$, if the inequality

$$\Bigl|\frac{\partial^{|\mu|}f}{\partial x^\mu}(x)\Bigr| \leq \Bigl|\frac{\partial^{|\mu|}g}{\partial x^\mu}(x)\Bigr| \tag{A.13}$$

holds for all multi indices $\mu \in \mathbb{N}_0^n$. As a simple example we note that any function in $\mathcal{C}_{M,r}(x)$ is majorised by the function

$$g(x) = \frac{Mr}{r - (x_1 + \cdots + x_n)} , \tag{A.14}$$

as the Taylor coefficients of g at the origin are just given by the right hand side of Cauchy's estimate (A.10).

A.3 Elementary Operations with Differential Equations

In this section we study three elementary operations with differential equations: rewriting a higher-order equation as a first-order one, quasi-linearisation and transformation to one dependent variable. In particular, we determine what happens, if we apply the operation to an involutive equation.

Reduction to first order. It is a classical operation to rewrite a differential equation of order q as a first-order equation through the introduction of additional dependent variables representing the derivatives of the original dependent variables. Geometrically, this is based on the identification of $J_q\mathcal{E}$ with a submanifold of $J_1(J_{q-1}\mathcal{E})$ based on the immersion $\iota_{q,1}$ (see Section 1.2). We define intrinsically $\tilde{\mathcal{R}}_1 = \iota_q(\mathcal{R}_q) \subseteq J_1(J_{q-1}\mathcal{E})$.

In local coordinates we get the following picture. Let $(x^i, u^\alpha, p^\alpha_\mu)$ be our usual coordinates on $J_q\mathcal{E}$; on $J_1(J_{q-1}\mathcal{E})$ we use $(x^i, \tilde{u}^\alpha_\nu, \tilde{p}^\alpha_{\nu,i})$ with $0 \leq |\nu| < q$. Then $\iota_{q,1}(x^i, u^\alpha, p^\alpha_\mu) = (x^i, p^\alpha_\nu, p^\alpha_{\nu+1_i})$. If the differential equation $\mathcal{R}_q$ is locally represented by the system $\Phi^\tau(x^i, u^\alpha, p^\alpha_\mu) = 0$, a local representation of $\tilde{\mathcal{R}}_1$ is given by

$$\tilde{\mathcal{R}}_1 : \begin{cases} \tilde{\Phi}^\tau(x^i, \tilde{u}^\alpha_\mu, \tilde{p}^\alpha_{\mu,i}) = 0\ , & \\ \tilde{p}^\alpha_{\mu,i} = \tilde{u}^\alpha_{\mu+1_i}\ , & |\mu| < q\ , \\ \tilde{p}^\alpha_{\mu+1_i,j} = \tilde{p}^\alpha_{\mu+1_j,i}\ , & |\mu| = q-1\ . \end{cases} \tag{A.15}$$

The second and third line represent the equations defining the submanifold $\iota_{q,1}(J_q\mathcal{E}) \subset J_1(J_{q-1}\mathcal{E})$. The function $\tilde{\Phi}^\tau$ is not uniquely defined, as in general several possibilities exist to express a derivative p^α_μ by one of the new coordinates. One way which allows us to compute easily the indices of the symbol $\tilde{\mathcal{M}}_1$ is to use the mapping

$$p^\alpha_\mu \longmapsto \begin{cases} \tilde{u}^\alpha_\mu & \text{for } |\mu| \leq q-1\ , \\ \tilde{p}^\alpha_{\mu-1_k,k} & \text{for } |\mu| = q,\ \operatorname{cls}\mu = k\ . \end{cases} \tag{A.16}$$

Proposition A.3.1. *Let $\tilde{\alpha}^{(k)}_1$ be the Cartan characters of $\tilde{\mathcal{R}}_1$ and $\alpha^{(k)}_q$ those of $\mathcal{R}_q$. Then $\tilde{\alpha}^{(k)}_1 = \alpha^{(k)}_q$ for $1 \leq k \leq n$ and there exists a one-to-one correspondence between the solutions*[1] *of $\tilde{\mathcal{R}}_1$ and $\mathcal{R}_q$. The equation $\mathcal{R}_q$ is involutive, if and only if $\tilde{\mathcal{R}}_1$ is involutive.*

Proof. Due to the substitution rule (A.16) the first line of (A.15) yields $\beta^{(k)}_q$ equations of class k, the same value as in $\mathcal{R}_q$. The second line represents

[1] As usual we are only considering either smooth or formal solutions.

$\dim J_{q-2}\mathcal{E}$ equations of class k and a little combinatorial calculation shows that the last line leads to additionally $\dim\big(S_{q-1}(T^*\mathcal{X})\otimes V\mathcal{E}\big)-\binom{q+n-k-1}{q-1}$ equations of class k. Adding these values we get

$$\tilde{\beta}_1^{(k)}=\dim J_{q-2}\mathcal{E}+\dim\big(S_{q-1}(T^*\mathcal{X})\otimes V\mathcal{E}\big)-\alpha_q^{(k)} \tag{A.17}$$

and thus for the Cartan characters

$$\tilde{\alpha}_1^{(k)}=\dim J_{q-1}\mathcal{E}-\tilde{\beta}_1^{(k)}=\alpha_q^{(k)}\,. \tag{A.18}$$

If x forms a δ-regular coordinate system for $\mathcal{R}_q$, then it does so for $\tilde{\mathcal{R}}_1$, too, and vice versa. The additional equations in (A.15) are in a form that leads to maximal values for the number of equations of class n and so on. Thus only the equations $\tilde{\Phi}^\tau=0$ could make trouble. But again due to our substitution rule (A.16) we see that if we get maximal values for the indices of $\mathcal{M}_q$ we get them also for $\tilde{\mathcal{M}}_1$ and vice versa.

These considerations imply furthermore that involution of the symbol of one of the equations is equivalent to involution of the symbol of the other one. It is easy to see that the second and third line in (A.15) alone form an involutive system and that there arise no integrability conditions between them and the equations $\tilde{\Phi}^\tau=0$. But if the equations $\Phi^\tau=0$ generate integrability conditions, then the same happens with the equations $\tilde{\Phi}^\tau=0$ and vice versa.

The relation between the solution spaces of $\mathcal{R}_q$ and $\tilde{\mathcal{R}}_1$, respectively, is obvious. If a section $\sigma:\mathcal{X}\to\mathcal{E}$ is a solution of $\mathcal{R}_q$, then $j_{q-1}\sigma:\mathcal{X}\to J_{q-1}\mathcal{E}$ is a solution of $\tilde{\mathcal{R}}_1$. Conversely, the second and third line of (A.15) ensure that every solution $\tilde{\sigma}:\mathcal{X}\to J_{q-1}\mathcal{E}$ of $\tilde{\mathcal{R}}_1$ is a section of the form $\tilde{\sigma}=j_{q-1}\sigma$ with $\sigma:\mathcal{X}\to\mathcal{E}$ a solution of $\mathcal{R}_q$. □

Quasi-Linearisation. The above mentioned method for a reduction to first order is not the only possible one. We may apply it not to $\mathcal{R}_q$ itself but to its prolongation $\mathcal{R}_{q+1}$. Recall that a prolongation leads always to a quasi-linear equation. This implies that the arising first-order equation $\bar{\mathcal{R}}_1\subset J_1J_q\mathcal{E}$ is quasi-linear, too. Note that the original equations are now transformed into algebraic constraints.

Proposition A.3.2. *Let $\bar{\alpha}_1^{(k)}$ be the Cartan characters of $\bar{\mathcal{R}}_1$ and $\alpha_q^{(k)}$ those of $\mathcal{R}_q$. Then $\bar{\alpha}_1^{(k)}=\sum_{i=k}^n\alpha_q^{(i)}$ for $1\leq k\leq n$ and there exists a one-to-one correspondence between the solutions of $\bar{\mathcal{R}}_1$ and $\mathcal{R}_q$. If $\mathcal{R}_q$ is involutive, then $\bar{\mathcal{R}}_1$ is involutive, too. If $\bar{\mathcal{R}}_1$ is involutive, then $\mathcal{R}_q$ is formally integrable and $\mathcal{R}_{q+1}$ is involutive.*

Proof. Most of the assertions follow immediately from Proposition A.3.1. According to Proposition 4.2.3, the Cartan characters of $\mathcal{R}_{q+1}$ and $\mathcal{R}_q$ are related by $\alpha_{q+1}^{(k)} = \sum_{i=k}^{n} \alpha_q^{(i)}$. If $\mathcal{R}_q$ is involutive, then $\mathcal{R}_{q+1}$ and hence $\bar{\mathcal{R}}_1$ are involutive, too.

For the final assertion, we first note that if $\bar{\mathcal{R}}_1$ is involutive, then by our previous results $\mathcal{R}_{q+1}$ is involutive. Applying our transformation to $\mathcal{R}_q$ yields a differential equation $\bar{\mathcal{R}}_0 \supseteq \bar{\mathcal{R}}_0^{(1)}$. As the transformation of $\mathcal{R}_{q+1}$ yields $\bar{\mathcal{R}}_1$ and as, because of the involution of $\bar{\mathcal{R}}_1$, the equality $(\bar{\mathcal{R}}_0^{(1)})_{+1} = \bar{\mathcal{R}}_1$ holds, we must in fact have $\bar{\mathcal{R}}_0 = \bar{\mathcal{R}}_0^{(1)}$. But this implies that $\mathcal{R}_q^{(1)} = \mathcal{R}_q$. Hence $\mathcal{R}_q$ is at least formally integrable. □

Example A.3.3. We cannot conclude that involution of $\bar{\mathcal{R}}_1$ entails involution of $\mathcal{R}_q$, as its construction uses a prolongation. Consider again the equation of Example 3.2.2

$$\mathcal{R}_2 : \left\{ u_{yy} = 0 , \quad u_{xx} = 0 . \right. \tag{A.19}$$

In order to make the notation less cumbersome, we introduce the following variable names in $J_2\mathcal{E}$: $r = u_{xx}$, $s = u_{xy}$, $t = u_{yy}$, $v = u_x$ and $w = u_y$. The first order equation $\bar{\mathcal{R}}_1 \subset J_1(J_2\mathcal{E})$ has a local representation of the form

$$\bar{\mathcal{R}}_1 : \begin{cases} t_y = 0 , & t_x = 0 , \\ s_x = 0 , & r_x = 0 , \\ r = 0 , & t = 0 , \\ r_y = s_x , & s_y = t_x , \\ v_y = s , & v_x = r , \\ w_y = t , & w_x = s , \\ u_y = w , & u_x = v . \end{cases} \tag{A.20}$$

The first six equations represent $\mathcal{R}_3$ in the new coordinates; the remaining equations define the submanifold $\iota_{3,1}(J_3\mathcal{E}) \subset J_1(J_2\mathcal{E})$.

One easily checks that $\bar{\mathcal{R}}_1$ is involutive. We know from Example 3.2.2 that $\mathcal{R}_2$ is only formally integrable but not involutive whereas $\mathcal{R}_3$ is involutive. ◁

Transformation to one dependent variable. It is always possible to transform a differential equation $\mathcal{R}_q$ in m dependent variables u^α and n independent variables x^i into one with only one dependent variable v. However, we must introduce m additional independent variables y^α; the order is raised by one and the relation between the transformed and the original equation is more complicated than in the above reduction to first order.

The basic idea is very simple: we define

$$v = \sum_{\alpha=1}^{m} y^\alpha u^\alpha . \tag{A.21}$$

Then for any multi-index $\mu \in \mathbb{N}_0^n$ the identity

$$\frac{\partial^{|\mu|}u^\alpha}{\partial x^\mu} = \frac{\partial^{|\mu|+1}v}{\partial y^\alpha \partial x^\mu} \tag{A.22}$$

holds and the differential equation $\mathcal{R}_q$ can be rewritten as an equation of order $q+1$ in the single dependent variable v. The definition (A.21) implies that we must add the equations

$$\frac{\partial^2 v}{\partial y^\alpha \partial y^\beta} = 0\,, \qquad 1 \leq \alpha \leq \beta \leq m \tag{A.23}$$

plus all their prolongations up to order $q+1$. This determines a differential equation $\bar{\mathcal{R}}_{q+1}$.

The equations (A.23) imply that any solution of the transformed equation $\bar{\mathcal{R}}_{q+1}$ is of the form

$$v(\mathbf{x}, \mathbf{y}) = \sum_{\alpha=1}^{m} y^\alpha u^\alpha(\mathbf{x}) + \Lambda(\mathbf{x}) \tag{A.24}$$

where $\mathbf{u}(\mathbf{x})$ is a solution of the original equation $\mathcal{R}_q$ and $\Lambda(\mathbf{x})$ is an arbitrary function. Obviously, this arbitrary function is of no interest for us; thus we may consider it as a *gauge symmetry* of $\bar{\mathcal{R}}_{q+1}$.

Proposition A.3.4. *Let* $\bar{\alpha}_{q+1}^{(k)}$ *for* $1 \leq k \leq n+m$ *be the Cartan characters of* $\bar{\mathcal{R}}_{q+1}$ *where we set* $x^k = y^\alpha$ *for* $k = n+\alpha$ *and* $\alpha_q^{(k)}$ *for* $1 \leq k \leq n$ *the Cartan characters of* $\mathcal{R}_q$. *Then*

$$\bar{\alpha}_{q+1}^{(k)} = \begin{cases} 0 & \textit{for } n < k \leq n+m\,, \\ \alpha_q^{(k)} + \binom{q+n-k}{q} & \textit{for } 1 \leq k \leq n\,. \end{cases} \tag{A.25}$$

$\bar{\mathcal{R}}_{q+1}$ *is involutive, if and only if* $\mathcal{R}_q$ *is involutive.*

Proof. Several ways exist to prove this proposition. One may proceed head-on and determine the indices of $\bar{\mathcal{R}}_{q+1}$; they are given by

$$\bar{\beta}_{q+1}^{(k)} = \begin{cases} \binom{q+n+m-k}{q} & \text{for } n < k \leq n+m\,, \\ \beta_q^{(k)} + \frac{m(m+1)}{2} \sum_{i=k}^{n+m} \binom{q+n+m-i-2}{q-2} & \text{for } 1 \leq k \leq n\,. \end{cases} \tag{A.26}$$

The binomial coefficients count the contribution of (A.23) and their prolongations. This yields (A.25) after some computations involving non-trivial identities of binomial coefficients.

A simpler and more direct way exploits (A.24). It shows that the solution of $\bar{\mathcal{R}}_{q+1}$ does not contain any arbitrary functions of the new variables $\mathbf{y}$. Thus no parametric derivatives exist that are pure y-derivatives and $\bar{\alpha}^{(k)}_{q+1} = 0$ for $k > n$. Concerning the x-dependency we see that we recover all arbitrariness of $\mathcal{R}_q$ and have in addition the gauge symmetry. It contributes $\binom{q+n-k}{q}$ parametric derivatives of order $q+1$ and class k for $1 \leq k \leq n$. Hence $\bar{\alpha}^{(k)}_{q+1} = \alpha^{(k)}_q + \binom{q+n-k}{q}$ for $1 \leq k \leq n$.

The assertion about involution is trivial. (A.23) represents obviously an involutive equation; if it was not for the x-dependency, it would be a finite type equation. Cross-derivatives between the original and these added equations vanish trivially. Thus the only source for possible integrability conditions or obstructions to involution is the original equation and there the y-derivatives play no role at all. $\square$

Using Hilbert polynomials, it is much easier to express (A.25). Let $\bar{H}(s)$ be the Hilbert polynomial of $\bar{\mathcal{R}}_{q+1}$, $H(s)$ the one of $\mathcal{R}_q$ and $G(s)$ the one of the Lie pseudogroup defined by the Lie equation

$$\mathcal{G}_1 : \left\{ \frac{\partial \bar{v}}{\partial v} = 1 \;, \quad \frac{\partial \bar{v}}{\partial y^\alpha} = 0 \;, \qquad 1 \leq \alpha \leq m \;. \right. \tag{A.27}$$

It describes the gauge symmetry $\bar{v} = v + \Lambda(x)$ of $\bar{\mathcal{R}}_{q+1}$. Then we get

$$\bar{H}(s) = H(s-1) + G(s) \tag{A.28}$$

with $G(s) = \binom{q+n-1}{q}$.

A.4 Modified Stirling Numbers

The modified Stirling numbers appear in some manipulations of the Cartan characters and the Hilbert polynomial. Thus they are of a combinatorial origin. We introduce them in a fashion similar to the classical Stirling numbers of the first kind [64]. Let us denote by $x^{\overline{n}} = x(x+1)\cdots(x+n-1)$ the nth rising factorial powers (sometimes also called Pochhammer's symbol). Then we can define the Stirling numbers of the first kind[2] $S^{(k)}_n$ for $n \geq k$ by the identity

$$x^{\overline{n}} = \sum_{k=0}^{n} S^{(k)}_n x^k \;. \tag{A.29}$$

The modified Stirling numbers $s^{(n)}_k(q)$ arise, if we write $x = q + r + 1$ for some non-negative integer q in (A.29) and then express the right hand side as a polynomial in r. Table A.1 on page 242 contains some explicit values.

[2] We follow here the notation of Abramowitz and Stegun [3] for the Stirling numbers of the first kind. Graham et al. [64] propose the notation $\left[{n \atop k}\right]$ which we, however, use already in a different context.

Definition A.4.1. *The* modified Stirling numbers $s_k^{(n)}(q)$ *are defined for all non-negative integers* n, k, q *with* $n \geq k$ *by the identity*

$$(q+r+1)^{\overline{n}} = \sum_{k=0}^{n} s_{n-k}^{(n)}(q) r^k \tag{A.30}$$

(note the different position of the scripts compared with (A.29) *and the use of* $n-k$ *instead of* k*).*

We are mainly interested in the fact that the modified Stirling numbers appear in the expansion of binomial coefficients as polynomials. Dividing (A.30) by $n!$ yields immediately

$$\binom{q+r+n}{q+r} = \frac{1}{n!} \sum_{k=0}^{n} s_{n-k}^{(n)}(q)\, r^k \,. \tag{A.31}$$

This also explains why we have used $q+r+1$ in (A.30) instead of the perhaps more natural $q+r$. Using the identity $S_{n+1}^{(k+1)} = \sum_{i=0}^{k} \binom{i}{k} S_n^{(i)}$ [64] one can show that $s_k^{(n)}(0) = S_{n+1}^{(n-k+1)}$.

With the help of the elementary symmetric polynomials one can easily derive a closed form expression for them. The elementary symmetric polynomial $\sigma_k^{(n)}$ is a polynomial of degree k in n variables $x^1, \dots, x^n$ that remains invariant under arbitrary permutations of the variables [40, Chapt. 7]. Thus $\sigma_1^{(n)} = x^1 + \cdots + x^n$, $\sigma_2^{(n)} = x^1x^2 + x^1x^3 + \cdots + x^2x^3 + \cdots + x^{n-1}x^n$ and so on until $\sigma_n^{(n)} = x^1x^2\cdots x^n$. Now we can write

$$s_k^{(n)}(q) = \begin{cases} 1 & \text{for } k = 0\,, \\ \sigma_k^{(n)}(q+1, q+2, \dots, q+n) & \text{for } 0 < k \leq n\,. \end{cases} \tag{A.32}$$

The modified Stirling numbers inherit from the elementary symmetric polynomials the recursion relation

$$s_k^{(n)}(q) = s_k^{(n-1)}(q) + (q+n)\, s_{k-1}^{(n-1)}(q)\,. \tag{A.33}$$

One can use this relation to construct explicit expressions for $s_k^{(n)}(q)$. They become fairly lengthy for larger k.

Lemma A.4.2.

$$s_n^{(n)}(q) = (q+n)!/q!\,, \tag{A.34a}$$

$$s_1^{(n)}(q) = \frac{1}{2} n(n+2q+1)\,, \tag{A.34b}$$

$$s_2^{(n)}(q) = \frac{1}{24} n(n-1)\left(3n^2 + 12qn + 5n + 12q^2 + 12q + 2\right)\,. \tag{A.34c}$$

Proof. (A.34a) and (A.34b) follow immediately from Definition A.4.1. For (A.34c) we apply (A.33)

$$\begin{aligned} s_2^{(n)}(q) &= s_2^{(n-1)}(q) + (q+n)\, s_1^{(n-1)}(q) \\ &= \cdots = \sum_{l=2}^{n-1} (q+l+1)\, s_1^{(l)}(q) + s_2^{(2)}(q) \\ &= \frac{1}{2} \sum_{l=1}^{n-1} (q+l+1) \Big[l^2 + (2q+1) l \Big] . \end{aligned}$$

Evaluation of this sum yields the claim. □

One can also deduce recursion formulae relating the modified Stirling numbers for different values of q or express $s_k^{(n)}(q)$ directly through $s_l^{(m)}(0)$ with $0 \le m \le n$. Like most properties of these numbers they stem from a well-known formula for binomial coefficients.

Lemma A.4.3.

$$s_{n-k+1}^{(n)}(q) = n\, s_{n-k}^{(n-1)}(q) + s_{n-k+1}^{(n)}(q-1) . \tag{A.35}$$

Proof. The basis of this lemma is the following identity coming directly from the Pascal triangle

$$\binom{q+r+n}{q+r} = \sum_{k=0}^{n} \binom{q+r+k-1}{q+r-1} . \tag{A.36}$$

Plugging in (A.31) yields after a few simple index manipulations yet another recursion relation for the modified Stirling numbers

$$s_k^{(n)}(q) = \sum_{j=0}^{k} \frac{n!}{(n-j)!} s_{k-j}^{(n-j)}(q-1) . \tag{A.37}$$

Applying this relation in (A.35) proves at once the lemma. □

In order to express $s_k^{(n)}(q)$ through the modified Stirling numbers with $q = 0$, we must generalise (A.37). Each binomial coefficient $\binom{r+k}{k}$ represents by (A.31) a polynomial in r of degree k. Thus every polynomial in r of degree n can be written as a linear combination of $\binom{r}{0}, \binom{r+1}{1}, \ldots, \binom{r+n}{n}$. Especially there exist numbers $d_k^{(n)}(q)$ such that

$$\binom{q+r+n}{q+r} = \sum_{k=0}^{n} d_k^{(n)}(q) \binom{r+k}{k} . \tag{A.38}$$

Applying repeatedly (A.36) one can express these coefficients as nested sums

$$d_k^{(n)}(q) = \sum_{k_1=k}^{n} \sum_{k_2=k_1}^{n} \cdots \sum_{k_{q-1}=k_{q-2}}^{n} 1 \,. \tag{A.39}$$

For lower values of q these sums are easily evaluated. We omit, however, explicit expressions. Now we obtain similarly as above

$$s_k^{(n)}(q) = \sum_{j=0}^{k} \frac{n!}{(n-j)!} d_k^{(n)}(q) s_{k-j}^{(n-j)}(0) \,. \tag{A.40}$$

Finally, we list in Table A.1 some values of the modified Stirling numbers; the rows correspond to different values of n, the columns of k. We remark that many such integer sequences can be found in the encyclopedia [168] of which also an extended electronic version exists on the web at `www.research.att.com/~njas/sequences`. It contains a number of other generalisations of the classical Stirling numbers (with references).

	q=0				q=1				q=2				q=3			
$n \backslash k$	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1	1				2				3				4			
2	3	2			5	6			7	12			9	20		
3	6	11	6		9	26	24		12	47	60		15	74	120	
4	10	35	50	24	14	71	154	120	18	119	342	360	22	179	638	840

Table A.1. Some values of the modified Stirling numbers

B. Algebra

A proper understanding of involution in the form introduced in Chapter 2 requires certain ideas from commutative algebra which are collected in this appendix. Some classical references for all mentioned topics are [40, 41, 47]; older textbooks on commutative algebra are usually less constructive. The introductory textbook [99] puts strong emphasis on computational issues (using the computer algebra system CoCoa).

The first section quickly reviews some basic algebraic structures like monoids, rings, algebras etc. Much of the material is elementary. Special emphasis is put on modules; they are central for the algebraic analysis of the symbol of a differential equation. In particular, we introduce the Hilbert function and polynomial, as similar concepts appear for differential equations in Section 4.2. A simple determination of the Hilbert function may be considered as one motivation for the introduction of involutive bases.

The second section gives an introduction into the theory of Gröbner bases of polynomial ideals and modules. These bases represent a fundamental tool in commutative algebra and algebraic geometry. Like most computations in a Euclidian vector space become significantly easier, if one chooses an orthonormal basis, almost any computation with ideals profits from the knowledge of Gröbner bases. Standard references are [4, 15, 40]; most books on computer algebra also contain at least an introduction.

For us Gröbner bases are so important, as much of the theory of involutive bases is modelled on them and an involutive basis is a Gröbner basis with additional combinatorial properties. We recall the famous Buchberger algorithm for determining a Gröbner basis in order to allow for a comparison with the completion algorithm derived in Section 6.4. However, deeper questions of the optimisation of the algorithm are not discussed.

Similarly, for lack of space applications of Gröbner bases are omitted; many can be found in above references. They include in particular the problem of determining points on the variety defined by an ideal, in other words solving systems of polynomial equations. Other applications are the union or intersection of ideals, computations with cosets, elimination problems, ...

The third section recalls some aspects of syzygies and free resolutions. It discusses the theorem of Schreyer on a Gröbner basis of the syzygy module and the construction of the syzygy resolution.

B.1 Some Basic Algebraic Structures

Definition B.1.1. *A set $\mathcal{S}$ with an operation $\cdot : \mathcal{S} \times \mathcal{S} \to \mathcal{S}$ is a* semigroup*, if $\cdot$ is associative, i. e. if $s_1 \cdot (s_2 \cdot s_3) = (s_1 \cdot s_2) \cdot s_3$ for all $s_1, s_2, s_3 \in \mathcal{S}$. The semigroup is* commutative *or* Abelian*, if $s_1 \cdot s_2 = s_2 \cdot s_1$ for all $s_1, s_2 \in \mathcal{S}$. A semigroup $\mathcal{S}$ is a* monoid*, if there exists a neutral element $e \in \mathcal{S}$ such that $e \cdot s = s \cdot e = s$ for all $s \in \mathcal{S}$. A monoid $\mathcal{S}$ is a* group*, if every $s \in \mathcal{S}$ has an inverse $s^{-1} \in \mathcal{S}$ such that $s \cdot s^{-1} = e$.*

The for us most important example of an Abelian monoid is given by the multi indices $(\mathbb{N}_0^n, +)$ discussed in Appendix A.1. Its neutral element is $[0, \dots, 0]$. The integers $\mathbb{Z}$ are both a monoid and a group: a monoid with respect to multiplication and a group with respect to addition; the respective neutral elements are 1 and 0. A non-commutative monoid $(\mathcal{C}^\infty(\mathbb{R}), \circ)$ is defined by the smooth functions $\mathbb{R} \to \mathbb{R}$ with the composition as operation and the identity as neutral element. The invertible functions form a group.

Let $(\mathcal{S}, \cdot)$ be a monoid and $\mathcal{I} \subseteq \mathcal{S}$ a subset such that $\mathcal{S} \cdot \mathcal{I} \subseteq \mathcal{I}$, i. e. if we "multiply" an element of $\mathcal{I}$ by an arbitrary element of $\mathcal{S}$, we always obtain again an element of $\mathcal{I}$. Then we call $\mathcal{I}$ a *(left) monoid ideal*. In case that the monoid is not Abelian we may analogously define right monoid ideals by the condition $\mathcal{I} \cdot \mathcal{S} \subseteq \mathcal{I}$ or double sided monoid ideals by $\mathcal{S} \cdot \mathcal{I} \cdot \mathcal{S} \subseteq \mathcal{I}$.

Definition B.1.2. *A* (left) monoid module *over a monoid $\mathcal{S}$ with neutral element e is a set $\mathcal{M}$ together with an operation $\star : \mathcal{S} \times \mathcal{M} \to \mathcal{M}$ so that*

- $\forall m \in \mathcal{M} : \; e \star m = m$;
- $\forall s_1, s_2 \in \mathcal{S}, \forall m \in \mathcal{M} : \; (s_1 \cdot s_2) \star m = s_1 \star (s_2 \star m)$.

A trivial example of a monoid module over $\mathcal{S}$ is a monoid ideal $\mathcal{I} \subseteq \mathcal{S}$ with $\star = \cdot$. If $\mathcal{M}$ is a monoid module, a subset $\mathcal{G} \subset \mathcal{M}$ is a *generating set* of $\mathcal{M}$, if $\mathcal{M}$ is the smallest monoid module containing $\mathcal{G}$. In this case, any element of $\mathcal{M}$ can be represented in the form $s \star g$ with some $s \in \mathcal{S}$ and $g \in \mathcal{G}$. We write $\mathcal{M} = \langle \mathcal{G} \rangle$. The monoid module $\mathcal{M}$ is *finitely generated*, if it possesses a finite generating set.

The monoid $(\mathcal{S}, \cdot)$ is *Noetherian*, if every monoid ideal $\mathcal{I} \subseteq \mathcal{S}$ is finitely generated. Equivalently, every ascending chain of monoid ideals $\mathcal{I}_1 \subseteq \mathcal{I}_2 \subseteq \mathcal{I}_3 \subseteq \cdots \subseteq \mathcal{S}$ becomes stationary, i. e. an index $N \geq 1$ exists such that $\mathcal{I}_N = \mathcal{I}_{N+1} = \mathcal{I}_{N+2} = \cdots$. The monoid $(\mathbb{N}_0^n, +)$ is Noetherian. This can be shown by an induction on n. If we identify the multi index $\mu \in \mathbb{N}_0^n$ with the term x^μ, this result is often called *Dickson's Lemma* (it is usually formulated differently but is equivalent to this statement).

Definition B.1.3. *A* ring *is a set* $\mathcal{R}$ *with two operations* $+, \cdot : \mathcal{R} \times \mathcal{R} \to \mathcal{R}$ *such that* $(\mathcal{R}, +)$ *is an Abelian group, the neutral element of which we denote by* 0*, and* $(\mathcal{R}, \cdot)$ *is a semigroup. Furthermore, the following distributive laws must hold:*

- $\forall r_1, r_2, r_3 \in \mathcal{R}: \ r_1 \cdot (r_2 + r_3) = r_1 \cdot r_2 + r_1 \cdot r_3$*;*
- $\forall r_1, r_2, r_3 \in \mathcal{R}: \ (r_1 + r_2) \cdot r_3 = r_1 \cdot r_3 + r_2 \cdot r_3$*.*

$\mathcal{R}$ *is a* ring with unit*, if* $(\mathcal{R}, \cdot)$ *is a monoid whose neutral element is usually denoted by* 1*.*

A classical example of a ring with unit are the integers $\mathbb{Z}$. If $\mathcal{R}$ is any commutative ring with unit, the set $\mathcal{P} = \mathcal{R}[x^1, \dots, x^n]$ of all polynomials in the variables[1] $x^1, \dots, x^n$ with coefficients in $\mathcal{R}$ is again a ring with unit. Examples of non-commutative rings are rings of linear differential operators with variable coefficients. We deal almost exclusively with rings with unit; in fact in many books the definition of a ring requires the existence of a unit.

Definition B.1.4. *A* field $\Bbbk$ *is a ring with unit such that* $(\Bbbk \setminus \{0\}, \cdot)$ *is an Abelian group with neutral element* 1*.*

The rational numbers $\mathbb{Q}$ are the quotient field of the integers $\mathbb{Z}$. Similarly, the rational functions $\mathcal{R}(x^1, \dots, x^n)$ are the quotient field of the polynomial ring $\mathcal{P}$. The analytic functions do not form a field but only a ring, as the quotient of two analytic functions is no longer analytic at the zeros of the denominator. But the *meromorphic functions*, i.e. the functions that are everywhere analytic except on a discrete set of points where they have a pole, form a field. We consider exclusively fields of *characteristic zero*; any such field contains the rational numbers $\mathbb{Q}$ as a subfield.

Definition B.1.5. *Let* $\mathcal{R}$ *be a ring with unit. An Abelian group* $(\mathcal{M}, +)$ *is called a* (left) $\mathcal{R}$-module*, if a scalar multiplication* $\cdot : \mathcal{R} \times \mathcal{M} \to \mathcal{M}$ *exists with the following properties:*

- $\forall m \in \mathcal{M}: \ 1 \cdot m = m$*;*
- $\forall r_1, r_2 \in \mathcal{R}, \forall m \in \mathcal{M}: \ r_1 \cdot (r_2 \cdot m) = (r_1 r_2) \cdot m$*;*
- $\forall r \in \mathcal{R}, \forall m_1, m_2 \in \mathcal{M}: \ r \cdot (m_1 + m_2) = r \cdot m_1 + r \cdot m_2$*;*
- $\forall r_1, r_2 \in \mathcal{R}, \forall m \in \mathcal{M}: \ (r_1 + r_2) \cdot m = r_1 \cdot m + r_2 \cdot m$*.*

[1] We follow here (as everywhere in this text) the convention in differential geometry and use superscripts to index variables.

This is an obvious extension of the definition of a monoid module. Much of the literature deals only with modules over *commutative* rings. We have not made this restriction, as we want to be able to consider modules over, say, rings of differential operators. One can analogously define *right* modules with a scalar multiplication $\cdot : \mathcal{M} \times \mathcal{R} \to \mathcal{M}$. As we exclusively deal with left modules, we usually omit the "left" and simply speak of modules.

Example B.1.6. The simplest and most familiar examples of modules are *vector spaces*. Indeed, if the ring $\mathcal{R}$ is even a field, then an $\mathcal{R}$-module is nothing but a vector space over $\mathcal{R}$.

Vector spaces are a special case of *free modules*. An $\mathcal{R}$-module $\mathcal{M}$ is called free, if it is of the form $\mathcal{R}^m$ for some positive integer m. Thus the elements of $\mathcal{M}$ can be identified with m-dimensional vectors whose entries are elements of $\mathcal{R}$. Addition and scalar multiplication are defined componentwise and obviously satisfy the conditions above.

Another way to define a free $\mathcal{R}$-module is to take a set $\mathcal{G}$ of *generators* and to consider all formal linear combinations of elements of $\mathcal{G}$ with coefficients in $\mathcal{R}$. An element of the arising module $\mathcal{M}_\mathcal{G}$ has the form $m = \sum_{g\in\mathcal{G}} r_g g$ where only finitely many of the coefficients $r_g \in \mathcal{R}$ do not vanish. Again, addition and scalar multiplication are defined componentwise. $\lhd$

Let $\mathcal{G} \subset \mathcal{M}$ be a subset of $\mathcal{M}$. Then $\mathcal{G}$ is a *generating set* of $\mathcal{M}$, if any element $m \in \mathcal{M}$ can be written as a linear combination $m = \sum_{g\in\mathcal{G}} r_g g$ where only finitely many of coefficients $r_g \in \mathcal{R}$ do not vanish. Note that this representation is in general not unique! A module is called *finitely generated*, if it possesses a finite generating set. A set $\mathcal{G}$ is called *linearly independent*, if $\sum_{g\in\mathcal{G}} r_g g = 0$ implies that all coefficients r_g vanish. A linearly independent generating set is a *basis* of the module $\mathcal{M}$; only in this case the above representation of the elements of $\mathcal{M}$ as linear combinations is unique. Note, however, that most modules do not possess a basis! In fact, a module has a basis, if and only if it is free.

Remark B.1.7. These notions are familiar from linear algebra. One should, however, note that many facts which are true for vector spaces do not hold, in general, for modules. Assume, for example, that we are given a set of elements that are linearly dependent. In a vector space this implies that some of the elements can be expressed as linear combinations of the other ones. In a module this is not necessarily true. Take the free $\mathbb{Z}$-module $\mathbb{Z}^2$ and the two elements $m_1 = \binom{2}{4}$ and $m_2 = \binom{3}{6}$. Obviously, they are linearly dependent, as $3m_1 - 2m_2 = 0$. But over $\mathbb{Z}$ none is a multiple of the other one. $\lhd$

Let $\mathcal{M}$ and $\mathcal{N}$ be two $\mathcal{R}$-modules. A map $\phi : \mathcal{M} \to \mathcal{N}$ is a *module homomorphism*, if it respects the module structures of $\mathcal{M}$ and $\mathcal{N}$, i. e. if for all module elements $m_1, m_2 \in \mathcal{M}$ and for all ring elements $r \in \mathcal{R}$ the equality $\phi(m_1 + r \cdot m_2) = \phi(m_1) + r \cdot \phi(m_2)$ hold. As usual, a bijective homomorphism is called an *isomorphism* and two isomorphic modules can be identified to some extent.

Any finitely generated module is isomorphic to a factor module of a free module; this is called a *presentation by generators and relations*. In principle, the construction of such a presentation for a given $\mathcal{R}$-module $\mathcal{M}$ is straightforward. One takes a finite generating set $\mathcal{G} = \{g_1, \dots, g_m\}$ of $\mathcal{M}$ and the free $\mathcal{R}$-module $\mathcal{R}^m$. If $\mathcal{G}$ is a basis of $\mathcal{M}$, then $\mathcal{M} \cong \mathcal{R}^m$. Otherwise, we define the map

$$\phi : \begin{cases} \mathcal{R}^m & \longrightarrow & \mathcal{M} \\ (r_1, \dots, r_m) & \longmapsto & \sum_{i=1}^m r_i g_i \end{cases} \tag{B.1}$$

and set $\mathcal{N} = \ker \phi$. Then $\mathcal{M} \cong \mathcal{R}^m / N$. In practice, it may be difficult to effectively construct a generating set $\mathcal{G}$ and its relations module $\mathcal{N}$ for an arbitrary module $\mathcal{M}$. Fortunately, most of the modules we are dealing with automatically arise in a presentation by generators and relations, so that we do not bother about this problem.

A finite sequence $(x_1, \dots, x_k)$ of elements of the underlying ring $\mathcal{R}$ is called $\mathcal{M}$-*regular* for the $\mathcal{R}$-module $\mathcal{M}$, if x_1 is not a zero divisor of $\mathcal{M}$ and x_i is not a zero divisor for $\mathcal{M}/\langle x_1, \dots, x_{i-1}\rangle\mathcal{M}$ for $1 < i \leq k$. The maximal length of an $\mathcal{M}$-regular sequence is the *depth* d of the module $\mathcal{M}$. It represents a measure for the size of $\mathcal{M}$.

Definition B.1.8. *Let $\mathcal{R}$ be a ring with unit. An* $\mathcal{R}$-algebra $\mathcal{A}$ *is a (left) $\mathcal{R}$-module together with an $\mathcal{R}$-bilinear multiplication $\cdot : \mathcal{A} \times \mathcal{A} \to \mathcal{A}$.*

The ring $\mathcal{P} = \mathcal{R}[x^1, \dots, x^n]$ of polynomials with coefficients in an Abelian ring $\mathcal{R}$ is a simple example of an Abelian $\mathcal{R}$-algebra. More generally, let $\mathcal{M}$ be a finitely generated free $\mathcal{R}$-module. We introduce the *symmetric algebra* $S\mathcal{M}$ as follows.[2] Let $\{g_1, \dots, g_n\}$ be a basis of $\mathcal{M}$, then $S\mathcal{M}$ is the infinite-dimensional free $\mathcal{R}$-module generated by the basis $\{g^\mu = g_1^{\mu_1} \cdots g_n^{\mu_n} \mid \mu \in \mathbb{N}_0^n\}$. Thus $S\mathcal{M}$ is isomorphic to the polynomial ring $\mathcal{P}$ via the identification $x^i \leftrightarrow g_i$. However, it is important to note that the definition of $S\mathcal{M}$ is independent of the chosen basis of $\mathcal{M}$, in contrast to this isomorphism which is obviously not canonical. Linear differential operators with variable coefficients form a non-commutative algebra over their coefficient ring.

The multiplication in an algebra is not necessarily associative. A very important example of a non-associative algebra is a *Lie algebra*. Here the

[2] A more satisfying definition of the symmetric algebra is as a quotient space of the full tensor algebra over $\mathcal{M}$. However, for our purposes this "pedestrian" approach via a basis suffices.

multiplication is usually written in form of a bracket, $a_1 \cdot a_2 = [a_1, a_2]$, and it must satisfy in addition to the distributive laws the following two axioms: (i) skew-symmetry, $[a_1, a_2] = -[a_2, a_1]$, and (ii) the *Jacobi identity*, $[a_1, [a_2, a_3]] + [a_2, [a_3, a_1]] + [a_3, [a_1, a_2]] = 0$. Note that an associative algebra is always a ring.

Subgroups, subrings, submodules etc. are defined in the obvious way: a subset that is closed under the given operations, i. e. the addition respectively the multiplication of two elements of the subset yields again an element of the subset. Thus in the case of a subring $\mathcal{I}$ one requires for example that $\mathcal{I} + \mathcal{I} \subseteq \mathcal{I}$ and $\mathcal{I} \cdot \mathcal{I} \subseteq \mathcal{I}$. There exists a special class of subrings that are particularly important, namely those which are even a monoid ideal with respect to the ring multiplication.

Definition B.1.9. *Let $\mathcal{R}$ be an arbitrary ring. A subring $\mathcal{I} \subseteq \mathcal{R}$ is called a* (left) ideal, *if $\mathcal{R} \cdot \mathcal{I} \subseteq \mathcal{I}$.*

In a *(left) Noetherian ring* $\mathcal{R}$ any (left) ideal $\mathcal{I} \subseteq \mathcal{R}$ is finitely generated or, equivalently, any ascending chain of (left) ideals $\mathcal{I}_1 \subseteq \mathcal{I}_2 \subseteq \mathcal{I}_3 \subseteq \cdots \subseteq \mathcal{R}$ becomes eventually stationary, i. e. there exists an $n \in \mathbb{N}$ such that $\mathcal{I}_{n+k} = \mathcal{I}_n$ for all k. By the *Hilbert Basis Theorem* the polynomial ring $\mathcal{R}[x^1, \cdots, x^n]$ is Noetherian, if the coefficient ring $\mathcal{R}$ is Noetherian.[3]

Definition B.1.10. *The ring $\mathcal{R}$ is* (non-negatively) graded, *if it has as additive Abelian group $(\mathcal{R}, +)$ a direct sum decomposition $\mathcal{R} = \bigoplus_{i \in \mathbb{N}_0} \mathcal{R}_i$ such that $\mathcal{R}_i \mathcal{R}_j \subseteq \mathcal{R}_{i+j}$. A* graded $\mathcal{R}$-module $\mathcal{M}$ *is an $\mathcal{R}$-module with a direct sum decomposition $\mathcal{M} = \bigoplus_{i \in \mathbb{N}_0} \mathcal{M}_i$ such that $\mathcal{R}_i \cdot \mathcal{M}_j \subseteq \mathcal{M}_{i+j}$. The elements of $\mathcal{R}_i$ or $\mathcal{M}_i$, respectively, are* homogeneous *of degree i. For a given degree q the* truncation $\mathcal{M}_{\geq q}$ *is defined as the module $\bigoplus_{p \geq q} \mathcal{M}_p$, i. e. we throw away all elements of degree less than q.*

If $\mathcal{R}$ is a graded ring, $\mathcal{R}_0$ is a subring and each component $\mathcal{R}_i$ is an $\mathcal{R}_0$-module. The graded ring we are mostly interested in is the polynomial ring $\mathcal{P} = \Bbbk[x^1, \ldots, x^n]$ where a natural grading is given by the total degree. Here $\mathcal{P}_0 = \Bbbk$ is even a field and the components $\mathcal{P}_i$ consisting of all homogeneous polynomials of degree i are vector spaces over $\Bbbk$. A submodule $\mathcal{N} \subseteq \mathcal{M}$ is graded, if it can be decomposed in the form $\mathcal{N} = \bigoplus_{i \in \mathbb{N}_0} (\mathcal{N} \cap \mathcal{M}_i)$. A graded submodule can always be generated by homogeneous elements.

Definition B.1.11. *Let $\mathcal{R}$ be a graded ring with $\mathcal{R}_0$ a field and $\mathcal{M}$ a finitely generated graded $\mathcal{R}$-module. Then the homogeneous components $\mathcal{M}_r$ are*

[3] The proof of Hilbert's Basis Theorem requires the *axiom of choice* (cf. [15, Section 4.1]). If you cannot accept this axiom, you should stop reading now, as most of our algebraic results depend on it!

finite-dimensional vector spaces over $\mathcal{R}_0$. The Hilbert function *of $\mathcal{M}$ is the numerical function*[4] *defined by*

$$h_{\mathcal{M}}(r) = \dim \mathcal{M}_r \; . \tag{B.2}$$

The Hilbert function represents a measure for the size of the graded module $\mathcal{M}$. As a vector space over $\mathcal{R}_0$ the whole module is generally infinite-dimensional, so that this dimension is not a useful measure. Instead we must use a more complicated object like the Hilbert function. Robbiano [152] gives a nice introduction into the theory of Hilbert functions (and numerical functions in general).

Let the ring $\mathcal{R}$ be generated over $\mathcal{R}_0$ by n elements of degree 1; this is for example the case, if $\mathcal{R} = \mathcal{P}$. It is a classical result in commutative algebra that at some degree r_0 the Hilbert function becomes polynomial, i. e. there exists a polynomial $H_{\mathcal{M}}(r)$, the *Hilbert polynomial* of $\mathcal{M}$, such that for all integers $r \geq r_0$ the identity $H_{\mathcal{M}}(r) = h_{\mathcal{M}}(r)$ holds. One can show that $\deg h_{\mathcal{M}} = D \leq n$. The number D is the *(Krull) dimension* of the module $\mathcal{M}$. If the coefficient of r^D in $H_{\mathcal{M}}(r)$ is a_D, then the *multiplicity* of $\mathcal{M}$ is $a_D D!$.

Above we introduced the depth d as an alternative measure for the size of the module $\mathcal{M}$. One can show that always $D \geq d$. Modules such that $D = d$ are called *Cohen–Macaulay*.

Sometimes one prefers instead of the Hilbert function the *Hilbert series*.[5] It is the generating function of $h_{\mathcal{M}}(r)$, i. e. the univariate series in some variable λ defined by

$$\mathcal{H}_{\mathcal{M}}(\lambda) = \sum_{r=0}^{\infty} h_{\mathcal{M}}(r)\lambda^r \; . \tag{B.3}$$

One can show (see Proposition 2.5.2) that this series can always be summed in closed form yielding a rational function $\mathcal{H}_{\mathcal{M}}(\lambda) = f(\lambda)/(1-\lambda)^n$ where $f(\lambda)$ is a polynomial with integer coefficients. For the free module $\mathcal{M} = \mathcal{R}^m$ we get $\mathcal{H}_{\mathcal{M}}(\lambda) = m/(1-\lambda)^n$. The dimension D of $\mathcal{M}$ arises now as the order of the pole of $\mathcal{H}_{\mathcal{M}}$ at $\lambda = 1$.

These notions were originally introduced by Hilbert [77] for the case of polynomial ideals.[6] Here one must distinguish between the *affine* and the *projective* Hilbert function. Let $\mathcal{I}$ be a homogeneous ideal in the polynomial ring $\mathcal{P}$, i. e. an ideal generated by homogeneous elements. Then the quotient $\mathcal{P}/\mathcal{I}$ is a graded $\mathcal{P}$-module and we define the projective Hilbert function $h_{\mathcal{I}}$

[4] A *numerical function* $f(r)$ is a function returning for every integer argument $r \in \mathbb{Z}$ an integer value.

[5] The names *Poincaré* or *Hilbert-Poincaré series* are also used frequently.

[6] Hilbert used the term *characteristic function*.

as the Hilbert function of the graded module $\mathcal{P}/\mathcal{I}$. A particular case of a homogeneous ideal is a *monomial ideal*, i. e. an ideal with a basis consisting only of monomials. Then $h_{\mathcal{I}}(r)$ counts the terms of degree r that are *not* contained in $\mathcal{I}$. As we will see in Remark B.2.9, the general case can always be reduced to the monomial one with the help of Gröbner bases.

For non-homogeneous ideals $\mathcal{I}$, the quotient $\mathcal{P}/\mathcal{I}$ has no obvious grading. There exist two different ways to define nevertheless a Hilbert function which both lead to the same result. We may homogenise the ideal $\mathcal{I}$: we introduce an additional new variable x^0 and consider the ideal

$$\mathcal{I}^h = \langle f^h \mid f \in \mathcal{I} \rangle \subseteq \Bbbk[x^0, x^1, \dots, x^n] \tag{B.4}$$

where the homogeneous polynomial f^h is obtained from f by multiplying each term $x^\mu \in \operatorname{supp} f$ by $(x^0)^{\deg f - |\mu|}$. Then we define the affine Hilbert function of $\mathcal{I}$ as the projective Hilbert function of $\mathcal{I}^h$: $h^a_{\mathcal{I}} = h_{\mathcal{I}^h}$. Alternatively, we may use the following approach. Denote by $\mathcal{P}_{\leq r}$ the vector space of all polynomials f with $\deg f \leq r$ and set $\mathcal{I}_{\leq r} = \mathcal{I} \cap \mathcal{P}_{\leq r}$. Now we define the affine Hilbert function by $h^a_{\mathcal{I}}(r) = \dim \mathcal{P}_{\leq r} - \dim \mathcal{I}_{\leq r}$. For a homogeneous ideal we have the obvious relation $h_{\mathcal{I}}(r) = h^a_{\mathcal{I}}(r) - h^a_{\mathcal{I}}(r-1)$.

The degree of the affine Hilbert polynomial $H^a_{\mathcal{I}}(r)$ is the *(Krull) dimension* D of the ideal $\mathcal{I}$. For an algebraically closed coefficient field $\Bbbk$, the affine variety defined by the ideal $\mathcal{I}$, i. e. the set of points in $\Bbbk^n$ that are common zeros of all polynomials in $\mathcal{I}$, has also the dimension D.

The Hilbert polynomial $H_{\mathcal{I}}(r)$ is an example of a *numerical polynomial*, i. e. $H_{\mathcal{I}} : \mathbb{Z} \to \mathbb{Z}$. These polynomials form a $\mathbb{Z}$-module. One can show with a simple argument from interpolation theory that any numerical polynomial of degree d can be written in the form

$$H_{\mathcal{I}}(r) = \sum_{i=0}^{d} a_i r^i = \sum_{i=0}^{d} b_i \binom{r}{i} . \tag{B.5}$$

The transformation between the coefficients a_i and b_i is a combinatorial operation leading to the modified Stirling numbers discussed in Appendix A.4. The appearance of binomial coefficients in (B.5) is not surprising, as according to (A.4a) the affine Hilbert polynomial of the monomial ideal in $\mathcal{P}$ generated by the single term x^μ is $H^a(r) = \binom{n+r-|\mu|}{n}$.

B.2 Gröbner Bases for Polynomial Ideals and Modules

Let $\Bbbk$ be a field of characteristic zero and $\mathcal{P} = \Bbbk[x^1, \dots, x^n]$ a polynomial ring. We denote by $\mathbb{T} = \{x^\mu \mid \mu \in \mathbb{N}_0^n\}$ the set of all *terms*.[7] $(\mathbb{T}, \cdot)$ where $\cdot$ is the usual multiplication defines an Abelian monoid which is obviously isomorphic to the monoid $(\mathbb{N}_0^n, +)$ of multi indices introduced in Appendix A.1 via the identification $\mu \leftrightarrow x^\mu$.

A *term order* $\prec$ is a total order on $\mathbb{T}$ compatible with the monoid structure. This means that it satisfies the following two conditions: (i) $1 \preceq t$ for all terms $t \in \mathbb{T}$ and (ii) $s \preceq t$ implies $r \cdot s \preceq r \cdot t$ for all terms $r, s, t \in \mathbb{T}$. These two conditions ensure that $\prec$ is a *well-order*, i. e. that a strictly decreasing sequence $t_1 \succ t_2 \succ \cdots$ cannot be infinite. If a term order fulfils in addition the condition that $s \prec t$ whenever $\deg s < \deg t$, it is called *degree compatible*. As $\mathbb{T}$ and $\mathbb{N}_0^n$ are isomorphic as monoids, we can also speak of term orders on $\mathbb{N}_0^n$. In fact, most term orders are defined via multi indices.

Example B.2.1. Some important term orders are the following ones. The *lexicographic* order is defined by $x^\mu \prec_{\text{lex}} x^\nu$, if the last non-vanishing entry of $\mu - \nu$ is negative. Thus $(x^2)^2 x^3 \prec_{\text{lex}} x^1(x^3)^2$. With respect to the *reverse lexicographic* order $x^\mu \prec_{\text{revlex}} x^\nu$, if the first non-vanishing entry of $\mu - \nu$ is positive. Now we have $x^1(x^3)^2 \prec_{\text{revlex}} (x^2)^2 x^3$. However, $\prec_{\text{revlex}}$ is *not* a term order, as it violates the first condition: $x^1 \prec_{\text{revlex}} 1$. The reverse lexicographic order should not be confused with the *inverse lexicographic* order $\prec_{\text{invlex}}$ which arises from $\prec_{\text{lex}}$ by inverting the order of the variables, i. e. $x^\mu \prec_{\text{invlex}} x^\nu$, if the first non-vanishing entry of $\mu - \nu$ is negative.

Degree compatible versions exist of all these orders. $x^\mu \prec_{\text{deglex}} x^\nu$, if $|\mu| < |\nu|$ or if $|\mu| = |\nu|$ and $x^\mu \prec_{\text{lex}} x^\nu$. Similarly, $x^\mu \prec_{\text{degrevlex}} x^\nu$, if $|\mu| < |\nu|$ or if $|\mu| = |\nu|$ and $x^\mu \prec_{\text{revlex}} x^\nu$. Note that $\prec_{\text{degrevlex}}$ *is* a term order, in fact a very important one!

Here we have defined the orders *inverse* to the usual convention in most texts on Gröbner bases. The classical forms arises, if one inverts the order of the variables: $x^1, \dots, x^n \to x^n, \dots, x^1$. Our version fits better to the conventions used in differential equations theory, in particular to our definition of the class of a multi index in Appendix A.1 (see below). $\triangleleft$

Any polynomial $f \in \mathcal{P}$ can be written in the form $f = \sum_{t \in \mathbb{T}} c_t t$ where only finitely many of the coefficients $c_t \in \Bbbk$ are different from 0. The set

$$\operatorname{supp} f = \{t \in \mathbb{T} \mid c_t \neq 0\} \tag{B.6}$$

[7] There is a notorious dispute in the commutative algebra community whether the elements of $\mathbb{T}$ should be called terms or monomials. We call them terms; a monomial is the product of a term and a coefficient.

is the *support* of f. For a given term order $\prec$, $\operatorname{supp} f$ contains a largest term, the *leading term* $\operatorname{lt}_\prec f$ of the polynomial f. Its coefficient is the *leading coefficient* $\operatorname{lc}_\prec f$ and if $\operatorname{lt}_\prec f = x^\mu$, we call μ the *leading exponent* $\operatorname{le}_\prec f$. Finally, the *leading monomial* is defined by $\operatorname{lm}_\prec f = \operatorname{lc}_\prec f \cdot \operatorname{lt}_\prec f$. If $\mathcal{F}$ is a finite set of polynomials, we write $\operatorname{lt}_\prec \mathcal{F}$ for the set $\{\operatorname{lt}_\prec f \mid f \in \mathcal{F}\}$.

Let $\mathcal{F} = \{f_1, \dots, f_r\}$ be a finite set of polynomials; it generates the following ideal in the ring $\mathcal{P}$

$$\langle \mathcal{F} \rangle = \left\{ h \in \mathcal{P} \mid h = \sum_{i=1}^{r} g_i f_i \;\; g_i \in \mathcal{P} \right\} . \tag{B.7}$$

If $\mathcal{I} = \langle \mathcal{F} \rangle$ for some ideal $\mathcal{I} \subseteq \mathcal{P}$, then $\mathcal{F}$ is called a *basis*[8] of $\mathcal{I}$. By *Hilbert's Basis Theorem* any ideal in $\mathcal{P}$ has a finite basis (which is of course not unique), as any field is trivially a Noetherian ring.

Lemma B.2.2. *Let for a term order* $\prec$ *the condition* $\operatorname{lt}_\prec f \in \mathbb{k}[x^1, \dots, x^k]$ *be equivalent to* $f \in \mathbb{k}[x^1, \dots, x^k]$ *for any polynomial* $f \in \mathcal{P}$. *Then* $\prec$ *is the lexicographic order* $\prec_{\text{lex}}$. *If the condition* $\operatorname{lt}_\prec f \in \langle x^1, \dots, x^k \rangle$ *is equivalent to* $f \in \langle x^1, \dots, x^k \rangle$ *for any homogeneous polynomial* $f \in \mathcal{P}$, *then* $\prec$ *is the degree reverse lexicographic order* $\prec_{\text{degrevlex}}$.

Proof. We leave the case of the lexicographic order as an exercise; it goes completely analogously to the for us more important case of the degree reverse lexicographic order. Let x^μ and x^ν be two arbitrary monomials of the same degree such that the first non-vanishing entry of $\mu - \nu$ is $\mu_k - \nu_k$. Without loss of generality, we assume that $\mu_k > \nu_k$. Set $\rho = [\nu_1, \dots, \nu_k, 0, \dots, 0]$ and consider the multi indices $\bar\mu = \mu - \rho$ and $\bar\nu = \nu - \rho$. Obviously, $x^{\bar\mu} \in \langle x^1, \dots, x^k \rangle$ whereas $x^{\bar\nu} \notin \langle x^1, \dots, x^k \rangle$. Considering the homogeneous polynomial $f = x^{\bar\mu} + x^{\bar\nu}$, the assumption of the lemma implies that the leading term of f must be $x^{\bar\nu}$. By the monotonicity of term orders, we conclude that $x^\mu \prec x^\nu$. But by definition, we also have $x^\mu \prec_{\text{degrevlex}} x^\nu$. ◻

We say that a term order *respects classes*, if for multi indices μ, ν of the same length $\operatorname{cls} \mu < \operatorname{cls} \nu$ implies $x^\mu \prec x^\nu$. It is now easy to see that by Lemma B.2.2 only one class respecting term order exists: the degree reverse lexicographic order.

If $\mathcal{I}$ is any ideal in $\mathcal{P}$, its *leading ideal* $\operatorname{lt}_\prec \mathcal{I}$ for a given term order $\prec$ is the monomial ideal generated by the leading terms of all polynomials in $\mathcal{I}$

$$\operatorname{lt}_\prec \mathcal{I} = \langle \{\operatorname{lt}_\prec h \mid h \in \mathcal{I}\} \rangle . \tag{B.8}$$

[8] Obviously, every polynomial ideal is a module over the polynomial ring. One should note that strictly speaking these "ideal bases" are only generating sets and not bases, as a proper ideal never defines a free module. But this abuse of language has become standard in commutative algebra.

Definition B.2.3. *Let $\mathcal{I} \subseteq \mathcal{P}$ be a non-zero ideal. A finite set $\mathcal{G} \subset \mathcal{P}$ is a* Gröbner basis[9] *of $\mathcal{I}$ for the term order $\prec$, if $\mathrm{lt}_\prec \mathcal{I} = \langle \mathrm{lt}_\prec \mathcal{G} \rangle$.*

Example B.2.4. Consider the set $\mathcal{F} = \{f_1 = \underline{xy} - y,\ f_2 = \underline{y^2} - x\} \subset \mathbb{k}[x, y]$ and let $\prec$ be a degree compatible order. The underlined terms are the leading terms of the elements of $\mathcal{F}$. Set $\mathcal{I} = \langle \mathcal{F} \rangle$; then $\mathcal{F}$ is *not* a Gröbner basis of $\mathcal{I}$ with respect to $\prec$. The polynomial $f_3 = \underline{x^2} - x = y \cdot f_1 + (1 - x) \cdot f_2$ is obviously an element of $\mathcal{I}$, but its leading term x^2 is not contained in the monomial ideal generated by $\mathrm{lt}_\prec \mathcal{F} = \{xy, y^2\}$. We have not yet a way to prove it, but the augmented set $\mathcal{G} = \{f_1, f_2, f_3\}$ is a Gröbner basis of $\mathcal{I}$. $\triangleleft$

Note that Definition B.2.3 does not require $\mathcal{G} \subset \mathcal{I}$; but we will later see that this is indeed always the case. A set $\mathcal{G}$ may well be a Gröbner basis of an ideal for one term order but not for another one; the chosen term order has a great influence on the basis. We sometimes simply say that a set $\mathcal{G}$ is a Gröbner basis. This means that it is a Gröbner basis of the ideal $\langle \mathcal{G} \rangle$ for some term order $\prec$ which is obvious from the context.

While Definition B.2.3 is probably the simplest and most direct way to define a Gröbner basis, it does not make obvious why Gröbner bases are so important or how one can check whether a given set $\mathcal{G}$ is a Gröbner basis. Furthermore, it is not clear whether Gröbner bases exist at all. In order to answer such questions we need a few additional concepts.

Definition B.2.5. *For a given term order $\prec$, the polynomial $g \in \mathcal{P}$ is* reducible *with respect to the finite subset $\mathcal{F} \subset \mathcal{P}$, if there exists a term $t \in \operatorname{supp} g$ and a polynomial $f \in \mathcal{F}$ such that $\mathrm{lt}_\prec f \mid t$. A polynomial $g' \in \mathcal{P}$ that is not reducible with respect to $\mathcal{F}$ and such that $g' - g \in \langle \mathcal{F} \rangle$ is called a* normal form *of g with respect to F: $g' = \mathrm{NF}_{\mathcal{F},\prec}(g)$.*

Algorithm B.1 allows us to determine a normal form of a polynomial g with respect to a set $\mathcal{F}$. If g is not in normal form with respect to $\mathcal{F}$, we take the largest term $t \in \operatorname{supp} g$ divisible by the leading term of an element $f \in \mathcal{F}$ and compute the difference $g' = g - (c_t t / \mathrm{lm}_\prec f) f$; obviously, this eliminates the monomial $c_t t$ in g. If g' is still reducible, we iterate. As after each step the largest divisible term in $\operatorname{supp} g'$ becomes smaller and $\prec$ is a well-order, the algorithm terminates after a finite number of steps.

The normal form of a polynomial g with respect to an arbitrary set $\mathcal{F}$ is not unique. This can already be seen from Algorithm B.1. In Line /5/ there might exist several elements $f \in \mathcal{F}$ such that $\mathrm{lt}_\prec f \mid t$ and any of them can be used for eliminating the monomial $c_t t$. For different choices we must expect to obtain different normal forms.

[9] The terminology was introduced by Buchberger in honour of his thesis supervisor Gröbner. Some authors prefer to speak of *standard bases*.

Algorithm B.1 A simple normal form algorithm

Input: finite set $\mathcal{F} \subset \mathcal{P}$, polynomial $g \in \mathcal{P}$, term order $\prec$
Output: a normal form $g' = \mathrm{NF}_{\mathcal{F},\prec}(g)$
/1/ $\mathcal{S} \leftarrow \{t \in \operatorname{supp} g \mid \exists f \in \mathcal{F} : \mathrm{lt}_\prec f \mid t\}$
/2/ $g' \leftarrow g$
/3/ **while** $\mathcal{S} \neq \emptyset$ **do**
/4/ $\quad t = \max_\prec \mathcal{S}$
/5/ $\quad$ choose $f \in \mathcal{F}$ such that $\mathrm{lt}_\prec f \mid t$
/6/ $\quad g' \leftarrow g' - (c_t t/\mathrm{lm}_\prec f)\, f$
/7/ $\quad \mathcal{S} \leftarrow \{t \in \operatorname{supp} g' \mid \exists f \in \mathcal{F} : \mathrm{lt}_\prec f \mid t\}$
/8/ **end_while**
/9/ **return** g'

Example B.2.6. We continue with the set $\mathcal{F} = \{f_1, f_2\}$ of Example B.2.4 and a degree compatible term order. If we reduce the polynomial $\underline{xy^2} - x$ first with respect to f_1, we obtain the normal form $\underline{x^2} - x$. If we start with a reduction with respect to f_2, the normal form 0 arises after one further reduction step. ◁

Closely related to the normal form algorithm is the so-called *division algorithm* which is used in many textbooks for the introduction of Gröbner bases. By a slight modification of Algorithm B.1, namely recording the individual reduction steps, one can determine a representation $g = \sum_{f \in \mathcal{F}} h_f f + g'$ with coefficients $h_f \in \mathcal{P}$. In Line /2/ we initialise $h_f \leftarrow 0$ for every $f \in \mathcal{F}$; in Line /6/ we compute in addition $h_f \leftarrow h_f + c_t t/\mathrm{lm}_\prec f$ for the f chosen in the previous line. This may be considered as a multivariate generalisation of polynomial division and g' is sometimes called the *remainder* of the reduction. By construction, the "quotients" h_f satisfy $\mathrm{lt}_\prec(h_f f) \preceq \mathrm{lt}_\prec g$.

Proposition B.2.7. *Let $\mathcal{I} \subseteq \mathcal{P}$ be a non-zero ideal and $\mathcal{G} \subset \mathcal{P}$ a finite set. Then the following statements are equivalent.*

(i) *The set $\mathcal{G}$ is a Gröbner basis of $\mathcal{I}$ for the term order $\prec$.*
(ii) *Every polynomial $f \in \mathcal{P}$ possesses a unique normal form with respect to the set $\mathcal{G}$.*
(iii) *Every polynomial $f \in \mathcal{I}$ can be written in the form*

$$f = \sum_{g \in \mathcal{G}} h_g g \tag{B.9}$$

where the coefficients $h_g \in \mathcal{P}$ satisfy $\mathrm{lt}_\prec(h_g g) \preceq \mathrm{lt}_\prec f$ for all $g \in \mathcal{G}$.

Proof. Assume that $\mathcal{G}$ is a Gröbner basis of $\mathcal{I}$ with respect to the given term order $\prec$. Let f_1 and f_2 be two normal forms of an arbitrary polynomial

$f \in \mathcal{P}$; then obviously $f_1 - f_2 \in \mathcal{I}$. The leading term of $f_1 - f_2$ is an element of $\operatorname{supp} f_1 \cup \operatorname{supp} f_2$. By definition of a normal form this set cannot contain any term that is divisible by the leading term of an element of $\mathcal{G}$. On the other hand, by definition of a Gröbner basis, $\operatorname{lt}_\prec(f_1 - f_2)$ must be divisible by the leading term of an element of $\mathcal{G}$. This contradiction can only be resolved, if $f_1 - f_2 = 0$ and thus (i) implies (ii).

If every polynomial in $\mathcal{P}$ possesses a unique normal form, then the normal form of all polynomials $f \in \mathcal{I}$ is zero, as zero is trivially a normal form of any element of the ideal. If we apply Algorithm B.1 in its extended form, i. e. the division algorithm, to such a polynomial f, we obtain a representation of the form (B.9). As already remarked above, the coefficients h_g satisfy by construction the required condition. Thus (iii) follows from (ii).

Finally, we assume that (iii) holds. The condition on the coefficients obviously implies that for any $f \in \mathcal{I}$ the leading term $\operatorname{lt}_\prec f$ is divisible by $\operatorname{lt}_\prec g$ for some $g \in \mathcal{G}$. Thus $\operatorname{lt}_\prec \mathcal{I} = \langle \operatorname{lt}_\prec \mathcal{G} \rangle$ and consequently (i) is true. $\square$

Example B.2.8. The right hand side of (B.9) is called a *standard representation* of f. It is important to note that it is *not* unique. While the proposition above asserts that all possible normal form algorithms return the same result, they may achieve it in different ways and thus the associated division algorithms yield different standard representations. Consider again the Gröbner basis $\mathcal{G} = \{f_1, f_2, f_3\}$ of Example B.2.4. The polynomial $f = xy^2 - x \in \langle \mathcal{G} \rangle$ has two different standard representations: $f = yf_1 + f_2 = xf_2 + f_3$. $\lhd$

Remark B.2.9. Proposition B.2.7 implies that any Gröbner basis $\mathcal{G}$ of an ideal $\mathcal{I}$ provides us with a method to determine unique representatives for the cosets in $\mathcal{P}/\mathcal{I}$, as all polynomials in a coset $[f]$ have the same normal form with respect to $\mathcal{G}$. The normal form even defines an isomorphism between $\mathcal{P}/\mathcal{I}$ considered as a vector space over $\Bbbk$ and the vector space generated by all terms *not* contained in $\operatorname{lt}_\prec \mathcal{I} = \langle \operatorname{lt}_\prec \mathcal{G} \rangle$, as any normal form contains only such terms. This fact was already observed by Macaulay [110] who concluded that every ideal $\mathcal{I}$ has the same Hilbert function as the *monomial* ideal $\operatorname{lt}_\prec \mathcal{I}$.[10]

An effective mean to perform arithmetics with cosets was the original motivation for Buchberger to introduce Gröbner bases in his thesis [28]. It also allows us to decide algorithmically the *ideal membership problem*: given a Gröbner basis $\mathcal{G}$, we have $f \in \mathcal{I}$, if and only if $\operatorname{NF}_{\mathcal{G},\prec}(f) = 0$. $\lhd$

Note that by the definition of a basis, it is obvious that all elements of $\mathcal{I}$ can be written in the form (B.9) (compare with (B.7)). The particular property of a Gröbner basis is the additional condition on the coefficients.

[10] In the case of the *affine* Hilbert function, i. e. for non-homogeneous ideals, this holds only for degree compatible term orders.

Definition B.2.10. *Let $\mathcal{G}$ be a Gröbner basis with respect to the term order $\prec$. The basis $\mathcal{G}$ is* minimal, *if it does not contain two distinct elements $g, g' \in \mathcal{G}$ such that $\mathrm{lt}_\prec g \mid \mathrm{lt}_\prec g'$. The basis $\mathcal{G}$ is called* reduced, *if every element $g \in \mathcal{G}$ is in normal form with respect to the set $\mathcal{G}\setminus\{g\}$ and if $\mathrm{lc}_\prec g = 1$ for all $g \in \mathcal{G}$.*

Obviously, any reduced Gröbner basis is minimal. The terminology "minimal basis" has the following origin. For an ideal $\mathcal{I}$ one calls a basis $\mathcal{F}$ minimal, if one cannot remove an element of $\mathcal{F}$ without losing the property that $\mathcal{F}$ is a basis of $\mathcal{I}$. A general ideal has many minimal bases (even with different numbers of elements); but every *monomial* ideal possesses a unique minimal *monomial* basis. If $\mathcal{G}$ is a minimal Gröbner basis of $\mathcal{I}$ with respect to the term order $\prec$, then the set $\mathrm{lt}_\prec \mathcal{G}$ is the unique minimal monomial basis of $\mathrm{lt}_\prec \mathcal{I}$.

Theorem B.2.11. *Let $\mathcal{I} \subseteq \mathcal{P}$ be an arbitrary non-zero ideal. Then $\mathcal{I}$ possesses a unique reduced Gröbner basis for any term order $\prec$.*

All the characterisations of Gröbner bases given so far are not very constructive, as they do not tell us how to find a Gröbner basis for a given ideal. This will change now with the help of the *S-polynomial*[11] $S_\prec(f, g)$ of two polynomials f, g. It is defined as

$$S_\prec(f,g) = \frac{\mathrm{lcm}(\mathrm{lm}_\prec f, \mathrm{lm}_\prec g)}{\mathrm{lm}_\prec f} \cdot f - \frac{\mathrm{lcm}(\mathrm{lm}_\prec f, \mathrm{lm}_\prec g)}{\mathrm{lm}_\prec g} \cdot g \tag{B.10}$$

where lcm denotes as usual the least common multiple. Note that the factors are chosen such that the leading monomials cancel in the subtraction.

Proposition B.2.12. *A finite set $\mathcal{G}$ is a Gröbner basis of the ideal $\langle \mathcal{G} \rangle$, if and only if for every pair $f, g \in \mathcal{G}$ the S-polynomial $S_\prec(f, g)$ reduces to zero with respect to $\mathcal{G}$, i. e. $\mathrm{NF}_{\mathcal{G},\prec}(S_\prec(f, g)) = 0$.*

Based on this result, we describe a concrete algorithm for the construction of Gröbner bases: the *Buchberger algorithm* (see Algorithm B.2). Given a finite set $\mathcal{G}$, we form the set $\mathcal{S}$ of all pairs $\{g_1, g_2\}$ of different elements of $\mathcal{G}$. Then we take one pair and compute the corresponding S-polynomial. If its normal form $\bar{g}$ with respect to $\mathcal{G}$ vanishes, we choose the next pair. Otherwise we add $\bar{g}$ to $\mathcal{G}$ and start anew. The algorithm stops with a Gröbner basis, when all S-polynomials reduce to zero.[12]

[11] The S stands for *syzygy*; see Appendix B.3.

[12] Normal form computations may also be considered in the more abstract framework of *term rewriting*. Buchberger's algorithm is then an instance of a famous critical pair completion procedure, namely the *Knuth–Bendix algorithm* [90]. More on this relation can be found in [193].

Algorithm B.2 Buchberger algorithm for determining Gröbner bases

Input: finite set $\mathcal{F} \subset \mathcal{P}$, term order $\prec$
Output: Gröbner basis $\mathcal{G}$ of $\langle \mathcal{F} \rangle$

/1/ $\mathcal{G} \leftarrow \mathcal{F}$
/2/ $\mathcal{S} \leftarrow \{\{g_1, g_2\} \mid g_1, g_2 \in \mathcal{G}, g_1 \neq g_2\}$
/3/ **while** $\mathcal{S} \neq \emptyset$ **do**
/4/ choose $\{g_1, g_2\} \in \mathcal{S}$
/5/ $\mathcal{S} \leftarrow \mathcal{S} \setminus \{\{g_1, g_2\}\}$
/6/ $\bar{g} = \mathrm{NF}_{\mathcal{G},\prec}(S_\prec(g_1, g_2))$
/7/ **if** $\bar{g} \neq 0$ **then**
/8/ $\mathcal{S} \leftarrow \mathcal{S} \cup \{\{\bar{g}, g\} \mid g \in \mathcal{G}\}$
/9/ $\mathcal{G} \leftarrow \mathcal{G} \cup \{\bar{g}\}$
/10/ **end_if**
/11/ **end_while**
/12/ **return** $\mathcal{G}$

Theorem B.2.13. *Algorithm B.2 terminates for any finite set $\mathcal{F} \subset \mathcal{P}$ and any term order $\prec$ with a Gröbner basis of the ideal $\langle \mathcal{F} \rangle$.*

Proof. The main point is to show the termination of the algorithm. If it terminates, the outcome is clearly a Gröbner basis of $\langle \mathcal{F} \rangle$: we only add elements of this ideal so that the span does not change and the termination criterion is that all S-polynomials reduce to zero which by Proposition B.2.12 characterises Gröbner bases.

The termination follows from a simple Noetherian argument. The algorithm determines a sequence $\mathcal{G}_1 \subsetneq \mathcal{G}_2 \subsetneq \mathcal{G}_3 \subsetneq \cdots$ of finite sets. Consider the new element $\bar{g}$ added in Line /9/ to $\mathcal{G}_i$ in order to obtain $\mathcal{G}_{i+1}$. It results from a normal form computation with respect to $\mathcal{G}_i$ in Line /6/. This implies that $\mathrm{lt}_\prec \bar{g} \notin \langle \mathrm{lt}_\prec \mathcal{G}_i \rangle$ and thus we obtain a strictly ascending chain of monomial ideals: $\langle \mathrm{lt}_\prec \mathcal{G}_1 \rangle \subsetneq \langle \mathrm{lt}_\prec \mathcal{G}_2 \rangle \subsetneq \langle \mathrm{lt}_\prec \mathcal{G}_3 \rangle \subsetneq \cdots$. By Dickson's Lemma it is not possible that such a chain becomes infinite. Hence after a finite number of steps we must have $\mathcal{G}_i = \mathcal{G}_{i+1}$ and the algorithm terminates with a Gröbner basis of the ideal $\langle \mathcal{F} \rangle$. □

Example B.2.14. We apply Buchberger's algorithm to the set $\mathcal{F} = \{f_1, f_2\}$ of Example B.2.4 and some degree compatible term order $\prec$. The S-polynomial of the two generators is $S_\prec(f_1, f_2) = x^2 - xy$ and a normal form of it with respect to $\mathcal{F}$ is given by $f_3 = \underline{x^2} - x$. As one can easily check by explicitly computing the normal forms of the two new S-polynomials $S_\prec(f_1, f_3)$ and $S_\prec(f_2, f_3)$, the set $\mathcal{G} = \{f_1, f_2, f_3\}$ is a Gröbner basis (even a reduced one) of the ideal $\langle \mathcal{F} \rangle$ for any degree compatible term order. ◁

Note that it is *not* recommended to use the algorithm in the described form for concrete computations! The form above makes the basic ideas fairly transparent, but it is not efficient. Furthermore, Algorithm B.2 does not determine a reduced Gröbner basis; this would require additional normal form computations each time a new element has been added to the basis.

The choice of the pairs $\{g_1, g_2\}$ in Line /4/ is important for the efficency. A number of strategies have been developed for it. The classical *normal selection strategy* chooses that pair for which the term $\operatorname{lcm}(\operatorname{lt}_\prec g_1, \operatorname{lt}_\prec g_2)$ is minimal with respect to the used term order $\prec$. It works very well for degree compatible term orders. Another popular approach is the "sugar cube" strategy [60] which is based on a simulated homogenisation of the input. In contrast to the normal selection strategy, this approach works well also for term orders like the purely lexicographic one.

Another crucial observation is that the algorithm spends most of its time computing normal forms. Buchberger found two criteria for predicting that certain S-polynomials reduce to zero. Their use often leads to drastic gains in efficiency, as the corresponding pairs $\{g_1, g_2\}$ can be thrown away without any computation.

Proposition B.2.15 (Buchberger's first criterion). *Let $g_1, g_2 \in \mathcal{P}$ be two polynomials such that $\gcd(\operatorname{lt}_\prec g_1, \operatorname{lt}_\prec g_2) = 1$. Then $\operatorname{NF}(S_\prec(g_1, g_2), \{g_1, g_2\}) = 0$, i. e. the S-polynomial $S_\prec(g_1, g_2)$ reduces to zero modulo the set $\{g_1, g_2\}$.*

Whereas the first criterion gives an immediate reducibility, the second one predicts that a certain S-polynomial will become reducible at some stage of the algorithm, possibly much later. Thus it is a considerably deeper result.

Proposition B.2.16 (Buchberger's second criterion). *Let $\mathcal{G} \subset \mathcal{P}$ be a finite set and $g_1, g_2, h \in \mathcal{G}$ such that $\operatorname{lt}_\prec h \mid \operatorname{lcm}(\operatorname{lt}_\prec g_1, \operatorname{lt}_\prec g_2)$. Then it suffices to treat in Algorithm B.2 the S-polynomials $S_\prec(g_1, h)$ and $S_\prec(g_2, h)$; the S-polynomial $S_\prec(g_1, g_2)$ may be ignored.*

Various ways exist to incorporate these criteria in Algorithm B.2, but we omit here a discussion. The proof of the first criterion is not very difficult; it follows from a direct computation. The second one is based on syzygy theory and we will briefly mention the underlying theory in the next section.

Optimisations of the Buchberger algorithm (and the design of alternative algorithms for the construction of Gröbner bases) is still a major research topic in computer algebra and we must refer to the literature (cited in the above mentioned textbooks). The complexity of computing Gröbner bases and deciding the ideal membership problem was thoroughly analysed by Mayr and collaborators in a number of publications; a survey with many references is given in [119]. A by now classical result is that the ideal membership problem for polynomials with rational coefficients is in the worst case exponential in space [118, 120]. The same holds for determining a reduced Gröbner basis.

One should stress that this complexity result is less a question of designing clever algorithms but ideals exist where the *size* of the basis is so huge, as Mayr and Meyer [120] demonstrated with an explicit example. Thus simply storing the basis requires already exponential space. At first sight, this seems to imply that it is hopeless to compute Gröbner bases. But one should always keep in mind that all these results concern the *worst case* complexity. There are some indications that many examples appearing in applications have a considerably lower complexity, as they are of a more geometric origin whereas all the known worst case examples are of a combinatorial origin. Some discussions of this and related issues can be found in [13].

All the results in this section apply to polynomial rings over a *field* $\mathbb{k}$. This restriction is important; we have nowhere payed particular attention to the coefficients, as we could always assume that we can divide by them, i. e. divisibility of monomials was solely a matter of the terms involved. This is no longer the case, if we consider polynomials with coefficients out of some ring. There exists a theory of Gröbner bases for this more general situation (see e. g. [4, Chapt. 4]), but things are getting more complicated.

Gröbner bases are defined for submodules of free modules in the same manner as for ideals; one only needs a bit more notation. We represent the elements of the free module $\mathcal{P}^m$ as vectors $\mathbf{f} = (f_1, \dots, f_m)$ with $f_\alpha \in \mathcal{P}$. The unit vectors $\boldsymbol{e}_\alpha$ with $e_{\alpha\beta} = \delta_{\alpha\beta}$ and $1 \leq \alpha \leq m$ form the standard basis of the module $\mathcal{P}^m$; thus $\mathbf{f} = f_1\boldsymbol{e}_1 + \cdots + f_m\boldsymbol{e}_m$.

First, we must extend the concept of a term order. A term $\mathbf{t}$ is now a vector of the form $\mathbf{t} = t\boldsymbol{e}_\alpha$ for some α and with $t \in \mathbb{T}$ a term in $\mathcal{P}$. We denote the set of all terms by $\mathbb{T}_m$; it is a monoid module over $\mathbb{T}$. By a slight abuse of notation, we will sometimes "divide" terms: if $\mathbf{t} = r\mathbf{s}$, then we set $\mathbf{t}/\mathbf{s} = r\boldsymbol{e}_\alpha$; thus the result is again a vector. Terms belonging to different unit vectors cannot be divided. A *term order* $\prec$ on $\mathbb{T}_m$ must satisfy the following two conditions: if $\mathbf{s}, \mathbf{t} \in \mathbb{T}_m$ are two arbitrary terms, then (i) $\mathbf{t} \preceq r\mathbf{t}$ and (ii) $\mathbf{s} \prec \mathbf{t}$ implies $r\mathbf{s} \prec r\mathbf{t}$ for all $r \in \mathbb{T}$. Obviously, these are simple generalisations of the conditions for a term order on $\mathbb{T}$ and assert that $\prec$ is compatible with the structure of $\mathbb{T}_m$ as a monoid module over $\mathbb{T}$.

In practice one uses on $\mathbb{T}_m$ mostly term orders derived from orders on $\mathbb{T}$. Two natural ways to "lift" an order are *TOP* and *POT*. These abbreviations stand for *Term Over Position* and *Position Over Term*, respectively. Let $\prec$ be a term order on $\mathbb{T}$. Then $s\boldsymbol{e}_\alpha \prec_{TOP} t\boldsymbol{e}_\beta$, if either $s \prec t$ or $s = t$ and $\alpha < \beta$, i. e. we rank the comparison of the terms s, t over the comparison of their position in the vector. We have $s\boldsymbol{e}_\alpha \prec_{POT} t\boldsymbol{e}_\beta$, if either $\alpha < \beta$ or

$\alpha = \beta$ and $s \prec t$, i. e. the position in the vector is more important than the terms. It is easy to show that both $\prec_{\mathrm{TOP}}$ and $\prec_{\mathrm{POT}}$ are term orders on $\mathbb{T}_m$.

The notion of a class respecting term order can trivially be extended to orders on $\mathbb{T}_m$. In a similar manner to the proof of Lemma B.2.2 one shows that the only class respecting term orders are the TOP lifts of the degree reverse lexicographic order for arbitrary labellings of the unit vectors $\mathbf{e}_\alpha$.

Given this definition of term orders on $\mathbb{T}_m$, we can straightforwardly extend all the notions introduced in the last section to submodules $\mathcal{M} \subseteq \mathcal{P}^m$. The leading term, coefficient, and monomial are defined as before; again we simply write $\mathrm{lt}_\prec \mathcal{M}$ for the monomial submodule generated by the leading terms of the elements of $\mathcal{M}$. Now we can easily generalise Definition B.2.3.

Definition B.2.17. *Let $\mathcal{M}$ be a submodule of the free module $\mathcal{P}^m$. A finite set $\mathcal{G} \subset \mathcal{P}^m$ is called a* Gröbner basis *of $\mathcal{M}$ for a given term order $\prec$, if $\mathrm{lt}_\prec \mathcal{M} = \langle \mathrm{lt}_\prec \mathcal{G} \rangle$.*

We may continue to use Algorithm B.1 for computing normal forms. In order to apply Buchberger's Algorithm B.2, we still have to generalise the S-polynomials. First, we introduce the least common multiple of two terms $\mathbf{s} = s\mathbf{e}_\alpha$ and $\mathbf{t} = t\mathbf{e}_\beta$ in $\mathbb{T}_m$: it is zero, if $\alpha \neq \beta$, and $\mathrm{lcm}(s,t)\mathbf{e}_\alpha$ otherwise. Then we define the S-"polynomial" of two elements $\mathbf{f}, \mathbf{g} \in \mathcal{P}^m$ by simply copying (B.10):

$$\mathbf{S}_\prec(\mathbf{f}, \mathbf{g}) = \frac{\mathrm{lcm}(\mathrm{lm}_\prec \mathbf{f}, \mathrm{lm}_\prec \mathbf{g})}{\mathrm{lm}_\prec \mathbf{f}} \cdot \mathbf{f} - \frac{\mathrm{lcm}(\mathrm{lm}_\prec \mathbf{f}, \mathrm{lm}_\prec \mathbf{g})}{\mathrm{lm}_\prec \mathbf{g}} \cdot \mathbf{g}\ . \tag{B.11}$$

Now the results and algorithms of Appendix B.2 can be straightforwardly extended to the module case. In particular, it turns out that all mentioned propositions and theorems remain valid.

B.3 Syzygies and Free Resolutions

Like many other things in commutative algebra, the term *syzygy* goes back to Hilbert. It comes from the Greek word for yoke and indeed a syzygy defines a "yoke" connecting some module elements.

Definition B.3.1. *Let $\mathcal{F} = \{\mathbf{f}_1, \dots, \mathbf{f}_s\}$ be a finite set of elements of the free module $\mathcal{P}^m$ where again $\mathcal{P} = \mathbb{k}[x^1, \dots, x^n]$. A* syzygy *of $\mathcal{F}$ is an element $\mathbf{h} = (h_1, \dots, h_s) \in \mathcal{P}^s$ such that*

$$\sum_{\tau=1}^{s} h_\tau \mathbf{f}_\tau = 0\ . \tag{B.12}$$

The syzygies of $\mathcal{F}$ form a submodule of $\mathcal{P}^s$, the syzygy module $\mathrm{Syz}(\mathcal{F})$.

The components of a syzygy can be understood as solutions of a linear system of equations over the ring $\mathcal{P}$. Its $m \times s$ matrix is easily obtained, if we write the module elements $\mathbf{f}_\tau$ as m-dimensional column vectors. In other words, $\mathrm{Syz}(\mathcal{F})$ is the kernel of the linear map

$$\phi_{\mathcal{F}} : \begin{cases} \mathcal{P}^s & \longrightarrow & \mathcal{P}^m \\ (h_1, \dots, h_s) & \longmapsto & h_1\mathbf{f}_1 + \cdots + h_s\mathbf{f}_s \end{cases} . \tag{B.13}$$

Thus $\mathrm{Syz}(\mathcal{F}) = \{0\}$, if and only if $\mathcal{F}$ is the basis of a free module.

A generating set of the module $\mathrm{Syz}(\mathcal{F})$ can be easily determined, if $\mathcal{F}$ is a Gröbner basis of $\langle\mathcal{F}\rangle$ (for some term order $\prec$). For an arbitrary set $\mathcal{F}$ one first computes a Gröbner basis $\mathcal{G}$ of the ideal $\langle\mathcal{F}\rangle$ and then determines $\mathrm{Syz}(\mathcal{F})$ from $\mathrm{Syz}(\mathcal{G})$ (see e. g. [4]). So let $\mathcal{G} = \{\mathbf{g}_1, \dots, \mathbf{g}_s\}$ be a Gröbner basis for some term order $\prec$; without loss of generality we may assume that $\mathcal{G}$ is reduced. Furthermore, let $\mathbf{t}_\sigma = \mathrm{lt}_\prec \mathbf{g}_\sigma$ and $\mathbf{t}_{\sigma\tau} = \mathrm{lcm}(\mathbf{t}_\sigma, \mathbf{t}_\tau)$. It follows from Proposition B.2.7 (iii) that we can write

$$\mathbf{S}_\prec(\mathbf{g}_\rho, \mathbf{g}_\sigma) = \sum_{\tau=1}^{s} h_{\rho\sigma\tau}\mathbf{g}_\tau \tag{B.14}$$

where the leading terms of the polynomials $h_{\rho\sigma\tau} \in \mathcal{P}$ satisfy $\mathrm{lt}_\prec \mathbf{S}_\prec(\mathbf{g}_\rho, \mathbf{g}_\sigma) = \max_\prec\bigl\{\mathrm{lt}_\prec(h_{\rho\sigma\tau}\mathbf{g}_\tau) \mid 1 \le \tau \le s\bigr\}$. Now we define for $\rho \neq \sigma$

$$\mathbf{S}_{\rho\sigma} = \frac{\mathbf{t}_{\rho\sigma}}{\mathbf{t}_\rho}\boldsymbol{e}_\rho - \frac{\mathbf{t}_{\rho\sigma}}{\mathbf{t}_\sigma}\boldsymbol{e}_\sigma - \mathbf{h}_{\rho\sigma} \in \mathcal{P}^s \tag{B.15}$$

with $\mathbf{h}_{\rho\sigma} = (h_{\rho\sigma 1}, \dots, h_{\rho\sigma s})$. By the definition of the S-"polynomial" and (B.14), this is a syzygy in $\mathrm{Syz}(\mathcal{G})$.

Theorem B.3.2. *Let $\mathcal{G}$ be a Gröbner basis. The syzygy module* $\mathrm{Syz}(\mathcal{G})$ *is generated by the set* $\mathcal{S}_{\mathcal{G}} = \{\mathbf{S}_{\rho\sigma} \mid 1 \le \rho < \sigma \le s\}$.

Thus we automatically obtain a generating set of the module $\mathrm{Syz}(\mathcal{G})$, if we compute the Gröbner basis $\mathcal{G}$ with Buchberger's algorithm and use the extended form of Algorithm B.1, the division algorithm discussed after Example B.2.6. It delivers us the coefficients $h_{\tau\sigma\rho}$ of the representation (B.14).

An interesting result given by Schreyer [156] in his diploma thesis is that this generating set is even a Gröbner basis for a certain term order. Let $\mathcal{F} = \{\mathbf{f}_1, \dots, \mathbf{f}_s\}$ be a subset of $\mathcal{P}^m$ and $\prec$ a term order on $\mathbb{T}_m$. We define an order $\prec_{\mathcal{F}}$ on $\mathbb{T}_s$ by setting $s\boldsymbol{e}_\sigma \prec_{\mathcal{F}} t\boldsymbol{e}_\tau$ if either $\mathrm{lt}_\prec(s\mathbf{f}_\sigma) \prec \mathrm{lt}_\prec(t\mathbf{f}_\tau)$ or $\mathrm{lt}_\prec(s\mathbf{f}_\sigma) = \mathrm{lt}_\prec(t\mathbf{f}_\tau)$ and $\tau < \sigma$ (note that we reverse the order of the indices). One can show that for arbitrary sets $\mathcal{F}$ (not necessarily Gröbner bases) $\prec_{\mathcal{F}}$ is a term order on $\mathbb{T}_s$.

Theorem B.3.3 (Schreyer). *The set* $\mathcal{S}_\mathcal{G}$ *is a Gröbner basis of the syzygy module* $\mathrm{Syz}(\mathcal{G})$ *with respect to the term order* $\prec_\mathcal{G}$.

Syzygy theory is also underlying Buchberger's second criterion (Proposition B.2.16). Let $\mathcal{T} = \{t_1, \dots, t_s\} \subset \mathbb{T}$ be a finite set of terms. As above we set $t_{\rho\sigma} = \mathrm{lcm}(t_\rho, t_\sigma)$ and $\mathbf{S}_{\rho\sigma} = \frac{t_{\rho\sigma}}{t_\rho}\mathbf{e}_\rho - \frac{t_{\rho\sigma}}{t_\sigma}\mathbf{e}_\sigma$. As we are only dealing with terms, no polynomials $h_{\rho\sigma}$ arise. Setting $t_{\rho\sigma\tau} = \mathrm{lcm}(t_\rho, t_\sigma, t_\tau)$, it is a simple exercise to verify by direct computation that

$$\frac{t_{\rho\sigma\tau}}{t_{\rho\sigma}}\mathbf{S}_{\rho\sigma} + \frac{t_{\rho\sigma\tau}}{t_{\sigma\tau}}\mathbf{S}_{\sigma\tau} + \frac{t_{\rho\sigma\tau}}{t_{\tau\rho}}\mathbf{S}_{\tau\rho} = 0\,. \tag{B.16}$$

An obvious consequence is that if $t_\tau \mid t_{\rho\sigma}$, then $\mathbf{S}_{\rho\sigma} \in \langle \mathbf{S}_{\sigma\tau}, \mathbf{S}_{\tau\rho}\rangle$. More generally, we may formulate the following statement.

Lemma B.3.4. *Let* $\mathcal{S} \subseteq \mathcal{S}_\mathcal{T} = \{\mathbf{S}_{\rho\sigma} \mid 1 \leq \rho < \sigma \leq s\}$ *be a generating set of* $\mathrm{Syz}(\mathcal{T})$. *Assume that the three pairwise distinct indices* ρ, σ, τ *are such that*[13] $\mathbf{S}_{\rho\sigma}, \mathbf{S}_{\sigma\tau}, \mathbf{S}_{\rho\tau} \in \mathcal{S}$ *and* $t_\tau \mid t_{\rho\sigma}$. *Then the smaller set* $\mathcal{S} \setminus \{\mathbf{S}_{\rho\sigma}\}$ *still generates* $\mathrm{Syz}(\mathcal{T})$.

Buchberger's second criterion is now a simple corollary to a generalisation of Proposition B.2.12. If $\mathcal{G} \subset \mathcal{P}$ is a finite set, we take $\mathcal{T} = \mathrm{lt}_\prec \mathcal{G}$. Given a subset $\mathcal{S} \subseteq \mathcal{S}_\mathcal{T}$ that generates $\mathrm{Syz}(\mathcal{T})$, one can show that it suffices for $\mathcal{G}$ being a Gröbner basis, if all S-polynomials $S_\prec(g_\rho, g_\sigma)$ with $\mathbf{S}_{\rho\sigma} \in \mathcal{S}$ reduce to zero modulo $\mathcal{G}$.

Definition B.3.5. *Let* $\mathcal{R}$ *be a ring and* $\mathcal{M}$ *an* $\mathcal{R}$*-module. An exact sequence*

$$\cdots \longrightarrow \mathcal{R}^{s_j} \xrightarrow{\psi_j} \mathcal{R}^{s_{j-1}} \longrightarrow \cdots \longrightarrow \mathcal{R}^{s_1} \xrightarrow{\psi_1} \mathcal{R}^{s_0} \xrightarrow{\psi_0} \mathcal{M} \longrightarrow 0\,, \tag{B.17}$$

i. e. a sequence where $\mathrm{im}\,\psi_{j+1} = \ker\psi_j$ *for all* $j \geq 0$ *and* $\mathrm{im}\,\psi_0 = \mathcal{M}$, *is called a* free resolution *of* $\mathcal{M}$. *If there exists an index* ℓ *such that* $s_\ell \neq 0$ *but* $s_j = 0$ *for all* $j > \ell$, *the resolution is* finite *and its* length *is* $\ell + 1$.

For a finitely generated $\mathcal{R}$-module $\mathcal{M}$ we can construct such a free resolution step by step using presentations by generators and relations. Assume we know such a presentation for $\mathcal{M}$, i. e. $\mathcal{M} \cong \mathcal{R}^{s_0}/\mathcal{M}_0$. This is equivalent to the short exact sequence

$$0 \longrightarrow \mathcal{M}_0 \xrightarrow{\iota_0} \mathcal{R}^{s_0} \xrightarrow{\phi_0} \mathcal{M} \longrightarrow 0 \tag{B.18}$$

where ι_0 is the inclusion map and ϕ_0 the map defined by (B.1). As $\mathcal{M}_0$ is again finitely generated, we can find a presentation $\mathcal{M}_0 \cong \mathcal{R}^{s_1}/\mathcal{M}_1$ leading to

[13] If $\rho > \sigma$, then we understand that $\mathbf{S}_{\sigma\rho} \in \mathcal{S}$ etc.

an analogous short exact sequence with maps ι_1 and ϕ_1. If we define $\psi_0 = \phi_0$ and $\psi_1 = \iota_0 \circ \phi_1$, we obtain the exact sequence

$$0 \longrightarrow \mathcal{M}_1 \xrightarrow{\iota_1} \mathcal{R}^{s_1} \xrightarrow{\psi_1} \mathcal{R}^{s_0} \xrightarrow{\psi_0} \mathcal{M} \longrightarrow 0\,. \tag{B.19}$$

We can iterate this construction taking a presentation of $\mathcal{M}_1$ leading to a finitely generated module $\mathcal{M}_2$ and so on. Each presentation leads to a short exact sequence of the form (B.18) with associated maps ι_i and ϕ_i. Defining $\psi_i = \iota_{i-1} \circ \phi_i$ we obtain the following picture where the diagonals come from the presentations and the central row defines a free resolution of $\mathcal{M}$:

$$\begin{tikzcd}
 & & 0 \arrow[dr] & & 0 & & & & \\
 & & & \mathcal{M}_1 \arrow[ur] \arrow[dr] & & & & & \\
\cdots \arrow[rr, "\psi_3"] \arrow[dr] & & \mathcal{R}^{s_2} \arrow[rr, "\psi_2"] \arrow[ur] & & \mathcal{R}^{s_1} \arrow[rr, "\psi_1"] \arrow[dr] & & \mathcal{R}^{s_0} \arrow[r, "\psi_0"] & \mathcal{M} \arrow[r] & 0\,. \\
 & \mathcal{M}_2 \arrow[ur] \arrow[dr] & & & & M_0 \arrow[ur] \arrow[dr] & & & \\
0 \arrow[ur] & & 0 & & 0 \arrow[ur] & & 0 & &
\end{tikzcd} \tag{B.20}$$

We are mainly interested in the case that $\mathcal{R} = \mathcal{P}$. Here explicit presentations are easily obtained. Let the finite set $\mathcal{F}$ generate the $\mathcal{P}$-module $\mathcal{M}$ and take $\operatorname{Syz}(\mathcal{F})$ as $\mathcal{M}_0$. If $\mathcal{F}$ is a Gröbner basis of $\mathcal{M}$, the Schreyer Theorem B.3.3 provides us with a Gröbner basis of $\operatorname{Syz}(\mathcal{F})$, so that we can iterate the construction obtaining a *syzygy resolution* of $\mathcal{M}$. Hilbert's famous Syzygy Theorem states that any finitely generated polynomial module has a finite free resolution.

Theorem B.3.6 (Hilbert). *Let $\mathcal{P} = \Bbbk[x^1, \dots, x^n]$. Every finitely generated $\mathcal{P}$-module has a finite free resolution whose length is at most n.*

Free resolutions are not unique. But if we impose an additional condition, then we obtain a resolution that is unique up to isomorphism. Denote by $\mathcal{P}_+$ the ideal $\langle x^1, \dots, x^n \rangle$.

Definition B.3.7. *A free resolution of a $\mathcal{P}$-module is called* minimal, *if the maps $\psi_j : \mathcal{P}^{s_j} \to \mathcal{P}^{s_{j-1}}$ satisfy $\operatorname{im} \psi_j \subseteq \mathcal{P}_+ \cdot \mathcal{P}^{s_{j-1}}$. In this case, the numbers s_j are the* Betti numbers *of $\mathcal{M}$ and the length of the resolution is the* projective dimension $\operatorname{proj\,dim} \mathcal{M}$.

C. Differential Geometry

As we are concerned with a geometric approach to differential equations, we need some basic notions in differential geometry which are collected in this appendix. We start with a section introducing (differentiable) manifolds and some spaces associated with them: the tangent and the cotangent bundle. The following section extends to a slightly more complicated concept, fibred manifolds, and also briefly introduces bundles, although we hardly need their additional structures. The final section studies in some more detail vector fields and differential forms and the basic operations with them.

Standard references on differential geometry are [2, 21, 91, 92, 103, 191], but they go much further than we need. For most of our purposes a working knowledge of some basic notions around manifolds is sufficient. Much of the material we need can also be found in a very accessible form in [129].

C.1 Manifolds

The jet bundle formalism is based on fibred manifolds, but before we study these we briefly recall the definition of an ordinary manifold. Loosely speaking, it is something that locally looks like an $\mathbb{R}^m$. When we speak of a manifold, we will always mean a finite-dimensional, smooth (C^∞) manifold.

Definition C.1.1. *An* m*-dimensional* manifold *is a set* $\mathcal{M}$ *with a countable family* $\mathcal{A}$ *of* coordinate charts $(\mathcal{U}, \phi_\mathcal{U})$ *where* $\mathcal{U} \subseteq \mathcal{M}$ *is a subset and* $\phi_\mathcal{U} : \mathcal{U} \to \Omega_\mathcal{U}$ *a map onto a connected open subset* $\Omega_\mathcal{U} \subseteq \mathbb{R}^m$ *such that:*

1. *The charts in* $\mathcal{A}$ *cover the whole manifold:* $\mathcal{M} = \bigcup_{(\mathcal{U},\phi_\mathcal{U})\in\mathcal{A}} \mathcal{U}$.
2. *If two charts* $(\mathcal{U}, \phi_\mathcal{U}), (\mathcal{V}, \phi_\mathcal{V}) \in \mathcal{A}$ *overlap, the* transition function
$$\psi_{\mathcal{U}\mathcal{V}} = \phi_\mathcal{V} \circ \phi_\mathcal{U}^{-1} : \phi_\mathcal{U}(\mathcal{U} \cap \mathcal{V}) \subseteq \mathbb{R}^m \longrightarrow \phi_\mathcal{V}(\mathcal{U} \cap \mathcal{V}) \subseteq \mathbb{R}^m \tag{C.1}$$
is a smooth function.
3. *If* $p_1 \in \mathcal{U}_1$ *and* $p_2 \in \mathcal{U}_2$ *are two distinct points of* $\mathcal{M}$*, there exist open neighbourhoods* $\Omega_1 \subseteq \Omega_{\mathcal{U}_1}$ *of* $\phi_{\mathcal{U}_1}(p_1)$ *and* $\Omega_2 \subseteq \Omega_{\mathcal{U}_2}$ *of* $\phi_{\mathcal{U}_2}(p_2)$ *such that* $\phi_{\mathcal{U}_1}^{-1}(\Omega_1) \cap \phi_{\mathcal{U}_2}^{-1}(\Omega_2) = \emptyset$.

$\mathcal{A}$ is called an *atlas* and provides the set $\mathcal{M}$ with a *differentiable structure*. The charts endow the manifold $\mathcal{M}$ with a *topology*: the open sets are the inverse images $\phi_{\mathcal{U}}^{-1}(\Omega)$ of open sets $\Omega \subseteq \Omega_{\mathcal{U}}$ in $\mathbb{R}^m$. The third point in the definition above states that $\mathcal{M}$ with this topology is a Hausdorff space. These topological aspects are not so important for us, as we are mainly working locally, i. e. we stay within a chart $(\mathcal{U}, \phi_{\mathcal{U}}) \in \mathcal{A}$ where we can map everything into $\mathbb{R}^m$ and then calculate in coordinates as usual. Typically, we do not explicitly mention the chart but simply say *"let* $\mathbf{x} = (x^1, \dots, x^m)$ *be local coordinates on* $\mathcal{M}$*"*. This means that we go into a chart $(\mathcal{U}, \phi_{\mathcal{U}}) \in \mathcal{A}$ and identify each point $p \in \mathcal{U}$ with the vector $\mathbf{x} = \phi_{\mathcal{U}}(p) \in \mathbb{R}^m$.

One can also define analytic or C^k or topological manifolds; in this case one requires that the transition functions $\psi_{\mathcal{U}\mathcal{V}}$ describing changes of coordinates are of class C^ω or C^k or just homeomorphisms, respectively. Infinite-dimensional manifolds are usually obtained by modelling them on some Banach space instead of $\mathbb{R}^m$.

Example C.1.2. The simplest example of a manifold is a connected open subset $\Omega \subseteq \mathbb{R}^m$. In this case $\mathcal{A}$ contains only one chart, namely $(\Omega, \mathrm{id}_\Omega)$.

A slightly more complicated example of an m-dimensional manifold is the surface $\mathcal{S}^m$ of the unit sphere in $\mathbb{R}^{m+1}$. Here it is not possible to cover the whole manifold with a single chart, as $\mathcal{S}^m$ is not homeomorphic to $\mathbb{R}^m$; one needs at least two charts using the stereographic projection. $\triangleleft$

The space $C^\infty(\mathcal{M}, \mathbb{R})$ of all smooth real-valued functions defined on the manifold $\mathcal{M}$ is denoted by $\mathcal{F}(\mathcal{M})$. It is a ring, as we can add and multiply such functions simply by adding and multiplying their values. A function $f \in \mathcal{F}(\mathcal{M})$ induces in a chart $(\mathcal{U}, \phi_{\mathcal{U}})$ a function $f_{\mathcal{U}} = f \circ \phi_{\mathcal{U}} : \Omega_{\mathcal{U}} \to \mathbb{R}$. We call f differentiable in $\mathcal{U}$, if $f_{\mathcal{U}}$ is differentiable. It is easy to see that differentiability at a point $p \in \mathcal{M}$ is independent of the used chart, as the transition functions are smooth.

A familiar object from elementary differential geometry is the tangent plane of a surface. Its generalisation to an arbitrary manifold $\mathcal{M}$ leads to the tangent bundle $T\mathcal{M}$. We start by introducing the tangent space $T_p\mathcal{M}$ at an arbitrary but fixed point $p \in \mathcal{M}$. There exist several possibilities; we use a rather geometric one which also stresses the close relation between tangent and jet bundles.

A curve on the manifold $\mathcal{M}$ is a differentiable map $\gamma : \mathcal{I} \subseteq \mathbb{R} \to \mathcal{M}$ from some real interval into $\mathcal{M}$. Without loss of generality we assume that $0 \in \mathcal{I}$. For the construction of $T_p\mathcal{M}$ we take all curves with $\gamma(0) = p$. Two such curves γ_1, γ_2 are *tangent* at p, if in some coordinate chart $(\mathcal{U}, \phi_{\mathcal{U}})$ with $p \in \mathcal{U}$ the tangent vectors $(\phi_{\mathcal{U}} \circ \gamma_1)'(0)$ and $(\phi_{\mathcal{U}} \circ \gamma_2)'(0)$ coincide (note that $\phi_{\mathcal{U}} \circ \gamma_i$ is a curve in $\mathbb{R}^m$ where we know tangent vectors from calculus). One can show that this definition is independent of the used chart.

Let $[\gamma]_p$ be the set of all curves $\bar{\gamma}$ with $\bar{\gamma}(0) = p$ which are tangent to γ at p. An element of $T_p\mathcal{M}$ is such an equivalence class of curves. It is not difficult to see that $T_p\mathcal{M}$ is a real vector space of the same dimension as the manifold $\mathcal{M}$, as a coordinate chart maps the elements of $T_p\mathcal{M}$ bijectively on vectors in $\mathbb{R}^m$.

Definition C.1.3. *Let $p \in \mathcal{M}$ be a point on the manifold $\mathcal{M}$. The* tangent bundle $T\mathcal{M}$ *is the union of all tangent spaces*

$$T\mathcal{M} = \bigcup_{p \in \mathcal{M}} \{p\} \times T_p\mathcal{M}\,. \tag{C.2}$$

The tangent bundle projection *is the map $\tau_{\mathcal{M}} : T\mathcal{M} \to \mathcal{M}$ defined by $\tau_{\mathcal{M}}\bigl(\{p\} \times [\gamma]_p\bigr) = p$ for any $[\gamma]_p \in T_p\mathcal{M}$.*

Let $\mathcal{M}$ and $\mathcal{N}$ be two manifolds and $f : \mathcal{M} \to N$ an arbitrary smooth map. It induces a map $Tf : T\mathcal{M} \to T\mathcal{N}$, the *tangent map*, which can be defined intrinsically as follows. Let $\gamma_{\mathcal{M}} : \mathcal{I} \subseteq \mathbb{R} \to \mathcal{M}$ be an arbitrary curve on $\mathcal{M}$. Together with f it defines a curve $\gamma_{\mathcal{N}} : \mathcal{I} \to \mathcal{N}$ on the manifold $\mathcal{N}$ by $\gamma_{\mathcal{N}} = f \circ \gamma_{\mathcal{M}}$. By definition, a point in $T_p\mathcal{M}$ is an equivalence class of such curves $\gamma_{\mathcal{M}}$ with $\gamma_{\mathcal{M}}(0) = p$. Obviously, all curves in this class are mapped on curves $\gamma_{\mathcal{N}}$ with $\gamma_{\mathcal{N}}(0) = f(p)$ and it is not difficult to see that in fact the class $[\gamma_{\mathcal{M}}]_p$ is mapped into the class $[\gamma_{\mathcal{N}}]_{f(p)}$. Thus we obtain a *linear* map $T_pf : T_p\mathcal{M} \to T_{f(p)}\mathcal{N}$. Putting together all these maps for different p yields the map Tf.

The tangent bundle $T\mathcal{M}$ is again a manifold, as the differentiable structure of $\mathcal{M}$ induces a differentiable structure on $T\mathcal{M}$. Namely, let $\mathcal{A}$ be an atlas of $\mathcal{M}$; then $T\mathcal{A} = \bigl\{(T\mathcal{U}, T\phi_{\mathcal{U}}) \mid (\mathcal{U}, \phi_{\mathcal{U}}) \in \mathcal{A}\bigr\}$ is an atlas of $T\mathcal{M}$. We have $\dim T\mathcal{M} = 2\dim\mathcal{M}$. Let $(\mathcal{U}, \phi_{\mathcal{U}})$ be a coordinate chart with $p \in \mathcal{U}$ and $\phi_{\mathcal{U}}(p) = \boldsymbol{x} = (x^1, \dots, x^m)$. Consider for $1 \leq i \leq m$ the curves $\gamma_i : \mathcal{I} \to \mathcal{M}$ with $(\phi_{\mathcal{U}} \circ \gamma_i)(t) = \boldsymbol{x} + t \cdot \boldsymbol{e}_i$ where $\boldsymbol{e}_i$ is the ith unit vector in $\mathbb{R}^m$. We denote the vector $[\gamma_i]_p$ by $\partial_{x^i}\big|_p$. These vectors form a basis of $T_p\mathcal{M}$. One often denotes coordinates with respect to this basis by $\dot{\boldsymbol{x}} = (\dot{x}^1, \dots, \dot{x}^m)$, so that local coordinates on $T\mathcal{M}$ have the form $(\boldsymbol{x}, \dot{\boldsymbol{x}}) \in \mathbb{R}^m \times \mathbb{R}^m$.

Locally, the tangent map Tf corresponds to the linearisation of f via its Jacobian (sometimes denoted df). Let us introduce coordinates $(\boldsymbol{x}, \dot{\boldsymbol{x}})$ on $T\mathcal{M}$ (with some chart $(\mathcal{U}, \phi_{\mathcal{U}})$ on $\mathcal{M}$) and $(\boldsymbol{y}, \dot{\boldsymbol{y}})$ on $T\mathcal{N}$ (with some chart $(\mathcal{V}, \psi_{\mathcal{V}})$ on $\mathcal{N}$). We write $\boldsymbol{y} = \bar{f}(\boldsymbol{x})$, where $\bar{f} = \phi_{\mathcal{U}}^{-1} \circ f \circ \psi_{\mathcal{V}} : \mathcal{U} \to \mathcal{V}$ is the coordinate representation of f. Then Tf has in these coordinates the following form:

$$\mathsf{T}f : \begin{cases} \mathsf{T}\mathcal{U} & \longrightarrow & \mathsf{T}\mathcal{V} \\ (x, \dot{x}) & \longmapsto & \left(\bar{f}(x), \dfrac{\partial \bar{f}}{\partial x}(x)\dot{x}\right) \end{cases} . \tag{C.3}$$

The *rank* of the map f at a point $p \in \mathcal{U}$ is defined as the rank of the Jacobian $\partial\bar{f}/\partial x$ at the point $\phi_{\mathcal{U}}(p)$. One can show that this definition is independent of the chosen coordinate charts. If $\mathcal{K}$ is a further manifold and $g : \mathcal{N} \to \mathcal{K}$ a further map, the familiar *chain rule* can be expressed as $\mathsf{T}(g \circ f) = \mathsf{T}g \circ \mathsf{T}f$.

Mostly, we are somewhat sloppy with our notations and do not distinguish between the map f and its coordinate representation $\bar{f}$. Thus we simply write $y = f(x)$ omitting the coordinate maps $\phi_{\mathcal{U}}$ and $\psi_{\mathcal{V}}$.

Let $f : \mathcal{M} \to \mathcal{N}$ be a smooth map. We assume for simplicity that rank f is constant on the manifold $\mathcal{M}$. Maps with the maximal possible rank (which is obviously the minimum of $\dim \mathcal{M}$ and $\dim \mathcal{N}$) are of particular importance and have got special names: if $\dim \mathcal{M} \leq \dim \mathcal{N}$, one speaks of an *immersion*; for $\dim \mathcal{M} \geq \dim \mathcal{N}$ of a *submersion*. Locally, immersions look like linear injections, submersions like projections (see [61, Section I.2] for a precise form of this statement). If f is a bijection and f^{-1} is also smooth, then f is called a *diffeomorphism* and the two manifolds $\mathcal{M}$ and $\mathcal{N}$ are *diffeomorphic*. Obviously, this requires that $\mathsf{T}f$ is an isomorphism with $(\mathsf{T}f)^{-1} = \mathsf{T}(f^{-1})$ and thus $\dim \mathcal{M} = \dim \mathcal{N}$.

An n-dimensional *submanifold* of a manifold $\mathcal{M}$ is a subset $\mathcal{N} \subseteq \mathcal{M}$ where for every point $p \in \mathcal{N}$ there exists a coordinate chart $\mathcal{U} \subseteq \mathcal{M}$ with a coordinate map $\phi_{\mathcal{U}} : \mathcal{U} \to \mathbb{R}^n \times \mathbb{R}^{m-n}$ such that $\phi_{\mathcal{U}}(\mathcal{N} \cap \mathcal{U}) = \phi_{\mathcal{U}}(\mathcal{U}) \cap (\mathbb{R}^n \times \{0\})$. This implies that a submanifold is in particular a manifold, as it inherits a differentiable structure from $\mathcal{M}$. The natural inclusion map $\iota : \mathcal{N} \hookrightarrow \mathcal{M}$ mapping each point on itself is an injective immersion. The *codimension* of the submanifold $\mathcal{N}$ in $\mathcal{M}$ is defined as $\operatorname{codim} \mathcal{N} = \dim \mathcal{M} - \dim \mathcal{N}$.

If $\mathcal{N}_1$ and $\mathcal{N}_2$ are two submanifolds of $\mathcal{M}$, we say that they are *transversal*, if for all $p \in \mathcal{N}_1 \cap \mathcal{N}_2$ the tangent spaces satisfy $\mathsf{T}_p(\mathcal{N}_1 \cap \mathcal{N}_2) = \mathsf{T}_p\mathcal{N}_1 \cap \mathsf{T}_p\mathcal{N}_2$. This condition can be interpreted as follows. Let γ_i be curves on $\mathcal{N}_i$ with $\gamma_1(0) = p = \gamma_2(0)$. The two curves can be tangent at p, if and only if there exists an interval $\mathcal{I} \subset \mathbb{R}$ such that $\gamma_i(\mathcal{I}) \subseteq \mathcal{N}_1 \cap \mathcal{N}_2$.

Definition C.1.4. *The dual space of the tangent space* $\mathsf{T}_p\mathcal{M}$*, denoted by* $\mathsf{T}_p^*\mathcal{M}$*, is called the* cotangent space *of the manifold* $\mathcal{M}$ *at the point* $p \in \mathcal{M}$*. Thus its elements are linear maps (functionals)* $\lambda_p : \mathsf{T}_p\mathcal{M} \to \mathbb{R}$*. The* cotangent bundle $\mathsf{T}^*\mathcal{M}$ *is the union of all cotangent spaces*

$$\mathsf{T}^*\mathcal{M} = \bigcup_{p \in M} \{p\} \times \mathsf{T}_p^*\mathcal{M} . \tag{C.4}$$

The cotangent bundle projection *is the map* $\tau^*_{\mathcal{M}} : \mathsf{T}\mathcal{M} \to \mathcal{M}$ *defined by* $\tau^*_{\mathcal{M}}\big(\{p\} \times \lambda_p\big) = p$ *for any* $\lambda_p \in \mathsf{T}_p^*\mathcal{M}$*.*

Above we have mentioned that the vectors $\partial_{x^i}\big|_p$ form a basis of $T_p\mathcal{M}$. The elements of the dual basis in $T_p^*\mathcal{M}$ are denoted by $dx^i\big|_p$, i. e. we have the relations $dx^i\big|_p(\partial_{x^j}\big|_p) = \delta^i_j$.

C.2 Fibred Manifolds and Bundles

Fibred manifolds can be seen as a generalisation of the Cartesian product of two manifolds. As we will see, in the neighbourhood of a point a fibred manifold is in fact always of this form.

Definition C.2.1. *A* fibred manifold *is a triple* $(\mathcal{E}, \pi, \mathcal{B})$ *where* $\mathcal{E}$ *and* $\mathcal{B}$ *are manifolds and* $\pi : \mathcal{E} \to \mathcal{B}$ *is a surjective submersion.* $\mathcal{E}$ *is called the* total space *of the fibred manifold,* $\mathcal{B}$ *the* base space *and* π *the* projection. *The subset* $\pi^{-1}(b) \subset \mathcal{E}$ *is the* fibre *over the point* $b \in \mathcal{B}$ *and written* $\mathcal{E}_b$.

We usually speak simply of the fibred manifold $\mathcal{E}$, if the base space $\mathcal{B}$ and the projection π are obvious from the context. If $\dim \mathcal{E} = n + m$ and $\dim \mathcal{B} = n$, then m is called the *fibre dimension* of the fibred manifold.

Example C.2.2. The simplest example of a fibred manifold is a *trivial fibred manifold.* Its total space is of the form $\mathcal{E} = \mathcal{B} \times \mathcal{F}$ for some manifold $\mathcal{F}$ and the map π is the projection pr_1 on the first factor. In this case, each fibre can be identified with the manifold $\mathcal{F}$. ◁

Locally, any fibred manifold looks like a trivial fibred manifold. With the help of the Implicit Function Theorem one can show that, because of the properties of a submersion, to each point $\xi \in \mathcal{E}$ a neighbourhood $\mathcal{U}_\xi \subset \mathcal{E}$, a manifold $\mathcal{F}_\xi$ and a diffeomorphism $\phi_\xi : \mathcal{U}_\xi \to \pi(\mathcal{U}_\xi) \times \mathcal{F}_\xi$ exist such that the diagram

$$\begin{array}{ccc} \mathcal{U}_\xi & \xrightarrow{\phi_\xi} & \pi(\mathcal{U}_\xi) \times \mathcal{F}_\xi \\ {\scriptstyle \pi|_{\mathcal{U}_\xi}}\big\downarrow & & \big\downarrow{\scriptstyle \mathrm{pr}_1} \\ \pi(\mathcal{U}_\xi) & \xrightarrow[\mathrm{id}]{} & \pi(\mathcal{U}_\xi) \end{array} \tag{C.5}$$

commutes. This means that $\mathrm{pr}_1(\phi_\xi(\zeta)) = \pi(\zeta)$ for any point $\zeta \in \mathcal{U}_\xi$. Thus any fibred manifold can be covered with subsets $\mathcal{U}_\xi$ which are diffeomorphic to products $\pi(\mathcal{U}_\xi) \times \mathcal{F}_\xi$. It is a natural thought to exploit this for the introduction of special coordinate charts on $\mathcal{E}$, so-called *adapted coordinates.* Namely, we first map with ϕ_ξ on $\pi(\mathcal{U}_\xi)\times\mathcal{F}_\xi$ and then use charts on $\pi(\mathcal{U}_\xi)$ and $\mathcal{F}_\xi$, respectively. The arising coordinate maps are of the form $\mathcal{U}_\xi \to \mathbb{R}^n \times \mathbb{R}^m$

and if two points $\xi_1, \xi_2 \in \mathcal{U}_\xi$ lie in the same fibre, i.e. $\pi(\xi_1) = \pi(\xi_2)$, then their first n coordinates are identical. We mostly denote adapted coordinates by $(\mathbf{x}, \mathbf{u}) = (x^1, \dots, x^n, u^1, \dots, u^m)$.

A submanifold $\mathcal{F} \subseteq \mathcal{E}$ of the total space is a *fibred submanifold*, if the restriction $\pi|_{\mathcal{F}} : \mathcal{F} \to \mathcal{B}$ is still a surjective submersion. In other words, $(\mathcal{F}, \pi|_{\mathcal{F}}, \mathcal{B})$ must again be a fibred manifold.

Fibre bundles are a special kind of fibred manifolds that look even stronger like a trivial bundle. Here we require that we may choose the subsets $\mathcal{U}_\xi$ in the form $\pi^{-1}(\mathcal{W}_b)$ where $\mathcal{W}_b$ is a neighbourhood in $\mathcal{B}$ of the point $b = \pi(\xi)$, i.e. $\mathcal{U}_\xi$ consists of complete fibres. A covering of $\mathcal{E}$ with such subsets is a *local trivialisation* of the bundle. One can show that to every local trivialisation a manifold $\mathcal{F}$ exists such that all the manifolds $\mathcal{F}_\xi$ are diffeomorphic to it. $\mathcal{F}$ is called the *typical fibre* of the bundle.

Special kinds of bundles arise, if the typical fibre $\mathcal{F}$ has additional structures. For us important are *vector bundles* where $\mathcal{F}$ is a vector space $\mathcal{V}$ and *affine bundles* where $\mathcal{F}$ is an affine space $\mathcal{A}$. Recall that an affine space $\mathcal{A}$ is modelled on a vector space $\mathcal{V}$: $\mathcal{A}$ is a set on which $\mathcal{V}$ acts free and transitively as additive group; in other words, we identify the "difference" $\alpha - \beta$ of any two points $\alpha, \beta \in \mathcal{A}$ with a vector $v \in \mathcal{V}$. As we can do this over every point of the base space, a vector bundle is associated with every affine bundle.

Example C.2.3. Simple examples of bundles are both tangent and cotangent bundle of a manifold $\mathcal{M}$. Their typical fibre is the vector space $\mathbb{R}^m$ where $m = \dim \mathcal{M}$; thus they are even vector bundles. The coordinates $(\mathbf{x}, \dot{\mathbf{x}})$ on the tangent bundle $T\mathcal{M}$ are adapted. $\triangleleft$

This special form of the sets $\mathcal{U}_\xi$ seems to be a rather weak condition. In fact, it has a number of consequences. For example, the differentiable structure of $\mathcal{E}$ is completely determined by those of $\mathcal{B}$ and $\mathcal{F}$. Bundles also carry an additional structure, namely a group acting on the fibres. As we do not need this, we stop here and refer to the literature for more details.

Definition C.2.4. *A* (local) section σ *of the fibred manifold* $\pi : \mathcal{E} \to \mathcal{B}$ *is a map* $\sigma : \mathcal{U} \subseteq \mathcal{B} \to \mathcal{E}$ *such that* $\pi \circ \sigma = \mathrm{id}_\mathcal{U}$. *The space of all local sections is denoted by* $\Gamma_{\mathrm{loc}}(\mathcal{B}, \mathcal{E})$ *or, if the base space is obvious from the context, shortly by* $\Gamma_{\mathrm{loc}}(\mathcal{E})$.

In adapted coordinates $(\mathbf{x}, \mathbf{u})$ on the fibred manifold $\mathcal{E}$, a section σ has the local form $\sigma(\mathbf{x}) = (\mathbf{x}, s(\mathbf{x}))$, as the condition $\pi \circ \sigma = \mathrm{id}_\mathcal{U}$ implies that every point $b \in \mathcal{U}$ is mapped to a point in the fibre $\mathcal{E}_b$ above it. The notion of a section can be understood as a generalisation of the graph of a function. Indeed, if we are dealing with a trivial fibred manifold $\mathcal{E} = \mathcal{B} \times \mathcal{F}$, every function $s : \mathcal{B} \to \mathcal{F}$ defines a section $\sigma(b) = (b, s(b))$ and the set $\sigma(\mathcal{B}) \subset \mathcal{B}\times\mathcal{F}$ is the graph of s.

Let $(\mathcal{E}, \pi, \mathcal{B})$ and $(\mathcal{F}, \rho, \mathcal{B})$ be two fibred manifolds over the same base space. Their *fibred product* is the fibred manifold $(\mathcal{E} \times_{\mathcal{B}} \mathcal{F}, \pi \times_{\mathcal{B}} \rho, \mathcal{B})$ where the total space $\mathcal{E} \times_{\mathcal{B}} \mathcal{F}$ consists of all pairs $(\xi, \zeta) \in \mathcal{E} \times \mathcal{F}$ with $\pi(\xi) = \rho(\zeta)$ and the projection is defined by $(\pi \times_{\mathcal{B}} \rho)(\xi, \zeta) = \pi(\xi) = \rho(\zeta)$. It is straightforward to show that this indeed yields again a fibred manifold.

The total space $\mathcal{E} \times_{\mathcal{B}} \mathcal{F}$ may also be considered as a fibred manifold over either $\mathcal{E}$ or $\mathcal{F}$, respectively, with obvious projections on the first or second factor, respectively. In this case one speaks of a *pull-back* and uses the notation $\rho^*(\pi) : \rho^*(\mathcal{E}) \to \mathcal{F}$ and $\pi^*(\rho) : \pi^*(\mathcal{F}) \to \mathcal{E}$, respectively.

Let $(\mathcal{E}, \pi, \mathcal{B})$ and $(\mathcal{F}, \rho, \mathcal{C})$ be two arbitrary fibred manifolds. A smooth map $F : \mathcal{E} \to \mathcal{F}$ is a *fibred morphism*, if there exists a map $f : \mathcal{B} \to \mathcal{C}$ (which is in fact uniquely determined and smooth) such that $\rho \circ F = f \circ \pi$. It maps fibres into fibres: if $\xi_1, \xi_2 \in \mathcal{E}_b$, then $F(\xi_1), F(\xi_2) \in \mathcal{F}_{f(b)}$. An example of a fibred morphism is the tangent map $Tf : T\mathcal{M} \to T\mathcal{N}$ of a map $f : \mathcal{M} \to \mathcal{N}$.

The projection π introduces in a natural way the notion of a vertical vector. Vertical vectors are tangential to the fibres. As π is a map between the two manifolds $\mathcal{E}$ and $\mathcal{B}$, it induces a linear map $T\pi$ between the tangent spaces $T\mathcal{E}$ and $T\mathcal{B}$ locally given by (C.3).

Definition C.2.5. *The* vertical bundle *of the fibred manifold* $\pi : \mathcal{E} \to \mathcal{B}$ *is the subbundle* $V\mathcal{E} = \ker T\pi \subset T\mathcal{E}$.

In adapted coordinates $(\mathbf{x}, \mathbf{u})$ on the manifold $\mathcal{E}$, vertical vectors at a point $\xi \in \mathcal{E}$ are of the form $v = v^\alpha \partial_{u^\alpha}\big|_\xi$, i.e. all their components in the directions $\partial_{x^i}\big|_\xi$ corresponding to the coordinates in the base space $\mathcal{B}$ vanish.

Example C.2.6. In the case of the trivial fibred manifold $\mathcal{E} = \mathcal{B} \times \mathcal{F}$ we have the decomposition $T\mathcal{E} = T\mathcal{B} \times T\mathcal{F}$ and we can identify $V\mathcal{E} = \mathcal{E} \times T\mathcal{F}$. ◁

If vertical vectors exists, one expects also horizontal vectors. However, this requires an additional structure on the fibred manifold $\mathcal{E}$, namely a *connection*, as it is not possible to give an intrinsic definition of a horizontal vector for arbitrary fibred manifolds. The simple idea that in adapted coordinates such vectors should have the form $v = v^i \partial_{x^i}\big|_\xi$ does not work, as this form does not remain invariant under changes to other adapted coordinates.

One way to define a connection is to choose (in a smooth manner) at each point $\xi \in \mathcal{E}$ a subspace $H_\xi\mathcal{E} \subset T_\xi\mathcal{E}$, a *horizontal space*, such that $T_\xi\mathcal{E} = H_\xi\mathcal{E} \oplus V_\xi\mathcal{E}$; i.e. each vector $v_\xi \in T_\xi\mathcal{E}$ can be decomposed in a sum $v_\xi = v_\xi^H + v_\xi^V$ with $v_\xi^H \in H_\xi\mathcal{E}$ and $v_\xi^V \in V_\xi\mathcal{E}$. This induces two natural maps: the *horizontal projector* $P^H : T\mathcal{E} \to H\mathcal{E}$ with $P^H(\xi, v_\xi) = (\xi, v_\xi^H)$ and the *vertical projector* $P^V : T\mathcal{E} \to V\mathcal{E}$ with $P^V(\xi, v_\xi) = (\xi, v_\xi^V)$. Conversely, each of these projectors uniquely defines the connection.

Alternatively, one may define a connection with a *horizontal lift*. This is a map $A : \mathcal{E} \underset{\mathcal{B}}{\times} T\mathcal{B} \to T\mathcal{E}$ which is linear in the second factor and a right inverse to the map $\rho = \tau_\mathcal{E} \times T\pi : T\mathcal{E} \to \mathcal{E} \underset{\mathcal{B}}{\times} T\mathcal{B}$, i.e. $\rho \circ A = \mathrm{id}_{\mathcal{E} \underset{\mathcal{B}}{\times} T\mathcal{B}}$. In adapted coordinates ρ maps a point $\big((x, u); (\dot{x}, \dot{u})\big) \in T\mathcal{E}$ into the point $\big((x, u); (x, \dot{x})\big) \in \mathcal{E} \underset{\mathcal{B}}{\times} T\mathcal{B}$. This implies that A must be locally of the form $\big((x, u); (x, \dot{x})\big) \mapsto \big((x, u); (\dot{x}, \bar{A}(x, u)\dot{x})\big)$ where $\bar{A}(x, u)$ is an $m \times n$ matrix. The horizontal spaces of the connection defined by the map A are given by $H_\xi\mathcal{E} = A(\xi, T_{\pi(\xi)}\mathcal{B})$. This explains the name horizontal lift: A takes a vector in $T\mathcal{B}$ and lifts it to a horizontal vector in $T\mathcal{E}$. Each connection uniquely determines such a horizontal lift and conversely each horizontal lift defines uniquely a connection.

C.3 Vector Fields and Differential Forms

In this section we want to discuss in some more detail sections of the tangent and cotangent bundle of a manifold. They are fundamental objects in differential geometry.

Definition C.3.1. *Sections of the bundle* $T\mathcal{M}$ *are called* vector fields. *We write* $\mathfrak{X}(\mathcal{M})$ *for* $\Gamma_{loc}(T\mathcal{M})$.

$\mathfrak{X}(\mathcal{M})$ possesses some algebraic structures (see Appendix B.1). In particular, it is an $\mathcal{F}(\mathcal{M})$-module: we can add vector fields and multiply them by functions in $\mathcal{F}(\mathcal{M})$. In local coordinates $(x^1, \dots, x^m)$, the module $\mathfrak{X}(\mathcal{M})$ is generated by the fields $\partial_{x^i} : p \mapsto \partial_{x^i}\big|_p$. Thus every vector field can be written locally as $X = v^i\partial_{x^i}$ with some functions $v^i \in \mathcal{F}(\mathcal{M})$.

The notation ∂_{x^i} already indicates that vector fields can also be interpreted as linear first-order differential operators acting on $\mathcal{F}(\mathcal{M})$. This is a more algebraic alternative to our geometric introduction of tangent vectors in the last section: a vector field is an $\mathbb{R}$-linear map $X : \mathcal{F}(\mathcal{M}) \to \mathcal{F}(\mathcal{M})$ that satisfies the Leibniz rule $X(fg) = X(f)g + fX(g)$ (such maps are also called *derivations*). The relation between the two approaches is provided by the tangent map. If $f \in \mathcal{F}(\mathcal{M})$, then $Tf : T\mathcal{M} \to T\mathbb{R} = \mathbb{R} \oplus \mathbb{R}$ and for any vector field $X \in \mathfrak{X}(\mathcal{M})$ we define the function $X(f) \in \mathcal{F}(\mathcal{M})$ by $X(f)(p) = T_pf(X_p)$ for any point $p \in \mathcal{M}$. In local coordinates, the vector field ∂_{x^i} has then the meaning one expects: $\partial_{x^i}(f) = \partial f/\partial x^i$.

With this interpretation, $\mathfrak{X}(\mathcal{M})$ can be easily given the structure of a real Lie algebra. As multiplication we take the *Lie bracket*: $[X, Y] = X \circ Y - Y \circ X$

for any two fields $X, Y \in \mathfrak{X}(\mathcal{M})$. It is easy to verify that this yields again a derivation on $\mathcal{F}(\mathcal{M})$, i.e. a vector field in $\mathfrak{X}(\mathcal{M})$. If the two vector fields are given in local coordinates by $X = v^i \partial_{x^i}$ and $Y = w^i \partial_{x^i}$, then their Lie bracket is the vector field

$$[X, Y] = \left(v^j \frac{\partial w^i}{\partial x^j} - w^j \frac{\partial v^i}{\partial x^j} \right) \partial_{x^i} \ . \tag{C.6}$$

Because of these two linear structures, module over $\mathcal{F}(\mathcal{M})$ and Lie algebra over $\mathbb{R}$, one must be careful when speaking about linear combinations of vector fields. In order to avoid misunderstandings, one should always state where the coefficients live, if it is not evident from the context.

A d-dimensional *distribution* $\mathcal{D}$ on a manifold $\mathcal{M}$ assigns in a smooth manner to each point $p \in \mathcal{M}$ a vector subspace $\mathcal{D}_p \subseteq T_p\mathcal{M}$ with $\dim \mathcal{D}_p = d$. Typically, it is given as the span of some linearly independent vector fields $X^1, \dots, X^d$, i.e. $\mathcal{D}_p = \langle X^1_p, \dots, X^d_p \rangle$. A distribution is *involutive*, if it is closed under the Lie bracket. This means, if X and Y are two vector fields in $\mathcal{D}$ (thus $X_p, Y_p \in \mathcal{D}_p$ for all $p \in \mathcal{M}$), then $[X, Y]$ also lives in $\mathcal{D}$. The distribution is called *integrable*, if every point $p \in \mathcal{M}$ lies on a *maximal integral manifold* of $\mathcal{D}$, i.e. on a submanifold $\mathcal{N} \subseteq \mathcal{M}$ of dimension d such that $\mathcal{D}\big|_{\mathcal{N}} = T\mathcal{N}$. The famous *Frobenius Theorem* states that a distribution is integrable, if and only if it is involutive.

A simple example of a distribution is the vertical bundle of a fibred manifold. It is not difficult to see that it is even an involutive distribution; the fibres are integral manifolds of maximal dimension. If there is a connection defined on the fibred manifold, its horizontal spaces also form a distribution. In contrast to the vertical bundle, it is in general not involutive. Connections with an involutive horizontal distribution are called *flat*.

Definition C.3.2. *Sections of the bundle* $T^*\mathcal{M}$ *are called* differential (one-) forms. *We write* $\Omega^1(\mathcal{M})$ *for* $\Gamma_{loc}(T^*\mathcal{M})$.

Just like $\mathfrak{X}(\mathcal{M})$, the bundle $\Omega^1(\mathcal{M})$ is an $\mathcal{F}(\mathcal{M})$-module. In a chart $\mathcal{U}$ with local coordinates $x^1, \dots, x^n$, a module basis is given by the one-forms $dx^i : p \mapsto dx^i\big|_p$. Obviously, this is the dual basis to $\{\partial_{x^i}\}$.

A d-dimensional *codistribution* $\mathcal{D}^*$ on a manifold $\mathcal{M}$ maps each point $p \in \mathcal{M}$ smoothly to a vector subspace $\mathcal{D}^*_p \subseteq T^*_p\mathcal{M}$ with $\dim \mathcal{D}^*_p = d$. This notion is dual to distributions. If $\mathcal{D} \subseteq T\mathcal{M}$ is a distribution, its *annihilator* $\mathcal{D}^0 = \{\omega_p \in T^*_p\mathcal{M} \mid p \in \mathcal{M},\ \omega_p(\mathcal{D}_p) = 0\}$ is a codistribution (of dimension $\dim \mathcal{M} - \dim \mathcal{D} = \operatorname{codim} \mathcal{D}$). If a codistribution is spanned by some one-forms $\omega^1, \dots, \omega^d$, these are classically called a *Pfaffian system*.

A codistribution $\mathcal{D}^*$ is called *involutive*, if the exterior derivative of any form $\omega \in \mathcal{D}^*$ vanishes modulo $\mathcal{D}^*$, i.e. we can write $d\omega = \sum_{i=1}^{r} \eta_i \wedge \omega_i$

with $\omega_i \in \mathcal{D}^*$. With the help of (C.15) one easily checks that this condition is dual to the condition for an involutive distribution: $\mathcal{D}$ is involutive, if and only if $\mathcal{D}^0$ is involutive. An *integral manifold* of a codistribution $\mathcal{D}^*$ is a submanifold $\mathcal{N} \subseteq \mathcal{M}$ such that the pull-back $\iota^*(\mathcal{D}^*)$ with respect to the inclusion map $\iota : \mathcal{N} \hookrightarrow \mathcal{M}$ vanishes: $\iota^*\omega = 0$ for all $\omega \in \mathcal{D}^*$. The dual version of the *Frobenius Theorem* asserts that a codistribution is involutive, if and only if it is *integrable*, i.e. $\mathcal{D}^*$ possesses at any point $p \in \mathcal{M}$ integral manifolds of dimension d.

If there exist one-forms, one expects that there are also two-forms etc. Indeed, like over any finite-dimensional vector space, we can build at each point $p \in \mathcal{M}$ the *exterior algebra* $\Lambda(T_p^*\mathcal{M})$. It is a graded algebra (see Appendix B.1) with $\Lambda^0(T_p^*\mathcal{M}) = \mathbb{R}$ and $\Lambda^1(T_p^*\mathcal{M}) = T_p^*\mathcal{M}$. The forms of higher degree can be obtained with the help of the *wedge product*: $\wedge : \Lambda^k(T_p^*\mathcal{M}) \times \Lambda^\ell(T_p^*\mathcal{M}) \to \Lambda^{k+\ell}(T_p^*\mathcal{M})$. It is associative and distributive but non-commutative. Instead we have for $\omega \in \Lambda^k(T_p^*\mathcal{M})$ and $\eta \in \Lambda^\ell(T_p^*\mathcal{M})$ that $\omega \wedge \eta = (-1)^{k\ell}\eta \wedge \omega$. The precise definition is obtained by antisymmetrisation of the tensor product and requires the choice of a combinatorial factor. We will use the following convention

$$(\omega \wedge \eta)(v_1, \dots, v_{k+\ell}) = \frac{1}{k!\ell!} \sum_{\pi \in S_{k+\ell}} \epsilon(\pi)\omega(v_{\pi(1)}, \dots, v_{\pi(k)})\eta(v_{\pi(k+1)}, \dots, v_{\pi(k+\ell)}) \,. \tag{C.7}$$

Here $S_{k+\ell}$ is the symmetric group containing all permutations of the numbers $1, \dots, k+\ell$ and $\epsilon(\pi) = \pm 1$ depending on whether we are dealing with an even or an odd permutation. Some authors use a factor $1/(k+\ell)!$, but we will stick to the convention above.

We can also form vector bundles with the spaces $\Lambda^k(T_p^*\mathcal{M})$. Their local sections are contained in $\Omega^k(\mathcal{M})$ and are called *differential k-forms*.Note that zero-forms are just (local) functions $\mathcal{M} \to \mathbb{R}$. The direct sum of all these vector bundles yields the bundle $\Omega(\mathcal{M})$. In some local coordinates on the manifold $\mathcal{M}$, a basis of $\Omega^k(\mathcal{M})$ is given by the forms $dx^{i_1} \wedge \cdots \wedge dx^{i_k}$ with $i_1 < i_2 < \cdots < i_k$. However, typically on writes a k-form locally as $\omega = \omega_{i_1 \dots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}$ where the smooth functions $\omega_{i_1 \dots i_k}$ are totally antisymmetric in their indices.

If $f : \mathcal{N} \to \mathcal{M}$ is a smooth map, we can *pull-back* forms. If $\omega \in \Omega^k(\mathcal{M})$, we define via the tangent map Tf a form $f^*\omega \in \Omega^k(\mathcal{N})$ by setting

$$(f^*\omega)_p(v_1, \dots, v_k) = \omega_{f(p)}\bigl(T_p f(v_1), \dots, T_p f(v_k)\bigr) \,. \tag{C.8}$$

Note that this can in particular be applied to the inclusion map ι, if $\mathcal{N}$ is a submanifold of $\mathcal{M}$. In local coordinates we get the following picture. Let

$x^1, \dots, x^n$ be coordinates in some chart $\mathcal{U} \subseteq \mathcal{N}$ and $y^1, \dots, y^m$ coordinates in a chart $\mathcal{V} \subseteq \mathcal{M}$ where the charts are chosen such that $f(\mathcal{U}) \subseteq \mathcal{V}$ and we can write $y^j = f^j(\mathbf{x})$. If $\omega = \omega_{j_1 \dots j_k} dy^{j_1} \wedge \cdots \wedge dy^{j_k}$ is a k-form defined on $\mathcal{V}$, then its pull-back takes on $\mathcal{U}$ the local form

$$f^*\omega = \frac{\partial f^{j_1}}{\partial x^{i_1}} \cdots \frac{\partial f^{j_k}}{\partial x^{i_k}} \omega_{j_1 \dots j_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k} . \tag{C.9}$$

Given an algebra $\mathcal{A}$, a *derivation* on it is a map $\delta : \mathcal{A} \to \mathcal{A}$ that is linear with respect to the vector space structure and that satisfies a (possibly graded) *Leibniz rule* $\delta(fg) = \delta(f)g \pm f\delta(g)$. There exist several important derivations on $\Omega(\mathcal{M})$.

Definition C.3.3. *The* interior derivative *with respect to a given vector field* $X \in \mathfrak{X}(\mathcal{M})$ *is a map* $\iota_X : \Omega^{k+1}(\mathcal{M}) \to \Omega^k(\mathcal{M})$ *defined by*

$$(\iota_X\omega)\big(X_1, \dots, X_k\big) = \omega\big(X, X_1, \dots, X_k\big) \tag{C.10}$$

for all forms $\omega \in \Omega^{k+1}(\mathcal{M})$ *and vector fields* $X_1, \dots, X_k \in \mathfrak{X}(\mathcal{M})$.

Thus the interior derivative reduces the form degree. Its graded Leibniz rule is easily derived from the properties of the wedge product of forms: $\iota_X(\omega_1 \wedge \omega_2) = (\iota_X\omega_1) \wedge \omega_2 + (-1)^{k_1}\omega_1 \wedge (\iota_X\omega_2)$, if $\omega_i \in \Omega^{k_i}(\mathcal{M})$. In local coordinates, the interior derivative corresponds to what is classically called a "contraction": if $X = v^i\partial_{x^i}$ and $\omega = \omega_{i_1 \cdots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}$, then

$$\iota_X\omega = \omega_{i_1 i_2 \cdots i_k} v^{i_1} dx^{i_2} \wedge \cdots \wedge dx^{i_k} . \tag{C.11}$$

The next derivative increases the form degree, but it is less easy to introduce. We take here an algebraic approach and define it via some characteristic properties. One can show that there exists one and only one map satisfying all these properties.

Definition C.3.4. *The* exterior derivative *of differential forms is the unique map* $d : \Omega^k(\mathcal{M}) \to \Omega^{k+1}(\mathcal{M})$ *satisfying the following properties.*

- d *is an* $\mathbb{R}$*-linear map: for all* $a_1, a_2 \in \mathbb{R}$ *and* $\omega_1, \omega_2 \in \Omega^k(\mathcal{M})$

$$d(a_1\omega_1 + a_2\omega_2) = a_1 d\omega_1 + a_2 d\omega_2 . \tag{C.12}$$

- d *satisfies the Leibniz rule: for all* $\omega_1 \in \Omega^{k_1}(\mathcal{M})$ *and* $\omega_2 \in \Omega^{k_2}(\mathcal{M})$

$$d(\omega_1 \wedge \omega_2) = (d\omega_1) \wedge \omega_2 + (-1)^{k_1}\omega_1 \wedge d\omega_2 . \tag{C.13}$$

- d *is nilpotent:* $d \circ d = 0$.

• *If* $f : M \to \mathbb{R}$ *is a zero-form, the one-form* df *is locally given by*

$$df(x) = \frac{\partial f}{\partial x^i}(x)\,dx^i \,. \tag{C.14}$$

Given a k-form ω, an explicit coordinate-free expression for $d\omega$ is

$$\begin{aligned} d\omega(X_1, \dots, X_{k+1}) = &\sum_{i=1}^{k+1} (-1)^{i+1} X_i\big(\omega(X_1, \dots, \widehat{X}_i, \dots, X_{k+1})\big) + \\ &+ \sum_{i<j} \omega\big([X_i, X_j], X_1, \dots, \widehat{X}_i, \dots, \widehat{X}_j, \dots, X_{k+1}\big) \end{aligned} \tag{C.15}$$

where the hat indicates that the corresponding vector is omitted. In particular, for a function $f \in \mathcal{F}(\mathcal{M})$ we obtain $df(X) = X(f)$; thus the one-form df is just the fibre component of the tangent map Tf. An important property of the exterior derivative is that it commutes with the pull-back: if $f : \mathcal{M} \to \mathcal{N}$ is a smooth map, then $d \circ f^* = f^* \circ d$. In local coordinates, one obtains for $\omega = \omega_{i_1 \cdots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}$ that

$$d\omega = \frac{\partial \omega_{i_1 \cdots i_k}}{\partial x^i} dx^i \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} \,. \tag{C.16}$$

Remark C.3.5. If $f : \mathcal{M} \to \mathbb{R}$ is a smooth function on $\mathcal{M}$ such that zero is a regular value of it, then it is a classical result in differential geometry that its zero set $\mathcal{N} = \{p \in \mathcal{M} \mid f(p) = 0\}$ is a submanifold of $\mathcal{M}$. The tangent spaces $T_p\mathcal{N}$ may be identified with vector subspaces of $T_p\mathcal{M}$. These vector subspaces are easily determined with the help of the one-form $df \in \Omega^1(\mathcal{M})$: $T_p\mathcal{N} = \{v \in T_p\mathcal{M} \mid df_p(v) = 0\}$. As zero is a regular value, by definition df_p does not vanish for any point $p \in \mathcal{N}$ so that this condition is always non-trivial.

This follows directly from our geometric definition of tangent vectors as equivalence class of curves. If $\gamma : \mathcal{I} \to \mathcal{N}$ is a curve with $\gamma(0) = p$, then obviously $f\big(\gamma(t)\big) \equiv 0$. Differentiating this relation with respect to t yields $df\big(\dot{\gamma}(t)\big) \equiv 0$. Evaluation at $t = 0$ yields our result.

More generally, if the submanifold $\mathcal{N}$ is the common zero set of the functions $f^1, \dots, f^r$, then any vector in $T_p\mathcal{N}$ satisfies $df^1_p(v) = \cdots = df^r_p(v) = 0$ and vice versa. The vector subspace $N^*_p\mathcal{N} \subset T^*_p\mathcal{N}$ spanned by $df^1_p, \dots, df^r_p$ is sometimes called the *conormal space* of $\mathcal{N}$ at the point p. ◁

Remark C.3.6. We call a k-form $\omega \in \Omega^k(\mathcal{M})$ *exact*, if a $(k-1)$-form $\eta \in \Omega^{k-1}(\mathcal{M})$ exists such that $\omega = d\eta$. A form ω such that $d\omega = 0$ is *closed.* Obviously, any exact form is closed, as $d^2 = 0$. The *Poincaré Lemma* asserts that locally (more precisely, in star-shaped domains), the converse also holds: every closed form is exact. Globally, however, this is in general not true. Those closed forms that are not exact generate the *de Rham cohomology* of the manifold $\mathcal{M}$, a topological invariant. ◁

The last derivative is again with respect to a vector field and can be defined for arbitrary tensor fields, although we will introduce it only for differential forms. The general definition uses a limit process taking place along the flow lines of X. Our definition is then a theorem and (C.17) usually referred to as *Cartan's magic formula.*

Definition C.3.7. *The* Lie derivative *with respect to a vector field* $X \in \mathfrak{X}(\mathcal{M})$ *is the map* $\mathcal{L}_X : \Omega^k(\mathcal{M}) \to \Omega^k(\mathcal{M})$ *defined by*

$$\mathcal{L}_X = \iota_X \circ d + d \circ \iota_X \, . \tag{C.17}$$

For zero-forms, i. e. functions, this yields $\mathcal{L}_X f = Xf$. For a k-form ω we obtain the following coordinate-free expression

$$\begin{aligned} \mathcal{L}_X \omega(X_1, \dots, X_k) &= X\big(\omega(X_1, \dots, X_k)\big) + \\ &\quad + \sum_{i=1}^{k} \omega(X_1, \dots, [X_i, X], \dots, X_k) \, . \end{aligned} \tag{C.18}$$

Locally, we find for $X = v^i \partial_{x^i}$ and $\omega = \omega_{i_1 \cdots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}$ that

$$\begin{aligned} \mathcal{L}_X \omega = \Bigg(& v^i \frac{\partial \omega_{i_1 \cdots i_k}}{\partial x^i} + \\ & \frac{\partial v^i}{\partial x^{i_1}} \omega_{i i_2 \cdots i_k} + \cdots + \frac{\partial v^i}{\partial x^{i_k}} \omega_{i_1 \cdots i_{k-1} i} \Bigg) dx^{i_1} \wedge \cdots \wedge dx^{i_k} \, . \end{aligned} \tag{C.19}$$

The Lie derivative commutes with the exterior derivative, $d \circ \mathcal{L}_X = \mathcal{L}_X \circ d$, and satisfies for arbitrary differential forms $\omega_1, \omega_2 \in \Omega(\mathcal{M})$ the *ungraded* Leibniz rule $\mathcal{L}_X(\omega_1 \wedge \omega_2) = \mathcal{L}_X(\omega_1) \wedge \omega_2 + \omega_1 \wedge \mathcal{L}_X(\omega_2)$.