2. Manifolds as Locally Affine Spaces

Recall that an affine subspace of $\mathbb{R}^n$ is a translate of a vector subspace of $\mathbb{R}^n$, i.e., a set of the form $M = x^{(0)} + W = \{x^{(0)} + w \mid w \in W\}$ where $W \subseteq \mathbb{R}^n$ is a vector subspace of $\mathbb{R}^n$. There are two essentially different ways of describing such a space. First, we can specify a d-dimensional affine subspace of $\mathbb{R}^n$ as the solution set of a system of $r := n - d$ independent affine equations, i.e., as a set of the form

$$(2.1) \qquad M = \left\{ \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n \;\middle|\; \begin{array}{r} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{r1}x_1 + \cdots + a_{rn}x_n = b_r \end{array} \right\} = \{x \in \mathbb{R}^n \mid Ax = b\}$$

where $b \in \mathbb{R}^r$ is a given vector and where $A \in \mathbb{R}^{r \times n}$ is a given matrix of rank $\mathrm{rk}(A) = r$. (The condition $\mathrm{rk}(A) = r$ expresses the independence of the r equations in (2.1).) Defining $g : \mathbb{R}^n \to \mathbb{R}^{n-d}$ by $g(x) := Ax - b$, this can be rewritten as

$$(2.2) \qquad M = \{x \in \mathbb{R}^n \mid g(x) = 0\},$$

which means that M can be represented as the zero set of an affine map $g : \mathbb{R}^n \to \mathbb{R}^{n-d}$ whose linear part has full rank. Note that if $x^{(0)}$ is a fixed element of M, we can write

$$(2.3) \qquad M = \{x \in \mathbb{R}^n \mid A(x - x^{(0)}) = 0\} = x^{(0)} + \ker(A)\,.$$

Second, we can describe a d-dimensional affine subspace of $\mathbb{R}^n$ as a set of the form

$$(2.4) \qquad \begin{aligned} M &= \left\{ \begin{bmatrix} x_1^{(0)} \\ \vdots \\ x_n^{(0)} \end{bmatrix} + u_1 \begin{bmatrix} v_1^{(1)} \\ \vdots \\ v_n^{(1)} \end{bmatrix} + \cdots + u_d \begin{bmatrix} v_1^{(d)} \\ \vdots \\ v_n^{(d)} \end{bmatrix} \;\middle|\; u_1, \ldots, u_d \in \mathbb{R} \right\} \\ &= \{x^{(0)} + u_1 v^{(1)} + \cdots + u_d v^{(d)} \mid u \in \mathbb{R}^d\} \end{aligned}$$

where $x^{(0)}$ is a fixed vector and where $(v^{(1)}, \ldots, v^{(d)})$ is a given family of d linearly independent vectors in $\mathbb{R}^n$. (We say that $v^{(1)}, \ldots, v^{(d)}$ "span" or "generate" the vector space W. The linear independence of these vectors ensures that they form a minimal set of generating vectors and hence a basis of W, in which no element is superfluous.) Note that if $B \in \mathbb{R}^{n \times d}$ is the matrix with columns $v^{(1)}, \ldots, v^{(d)}$ and if we define $\varphi : \mathbb{R}^d \to \mathbb{R}^n$ by $\varphi(u) := x^{(0)} + Bu$, then (2.4) becomes

$$(2.5) \qquad M = \{\varphi(u) \mid u \in \mathbb{R}^d\}\,,$$

which means that M can be represented as the image of an affine map $\varphi : \mathbb{R}^d \to \mathbb{R}^n$ whose linear part has full rank. Note that (2.5) can also be written as

$$(2.6) \qquad M = \{x \in \mathbb{R}^n \mid x - x^{(0)} = Bu \text{ for some } u \in \mathbb{R}^d\} = x^{(0)} + \mathrm{im}(B).$$

We call (2.1), (2.2) or (2.3) a description of M in **equation form** and (2.4), (2.5) or (2.6) a description in **parameter form**. One- and two-dimensional affine space are easily visualized.

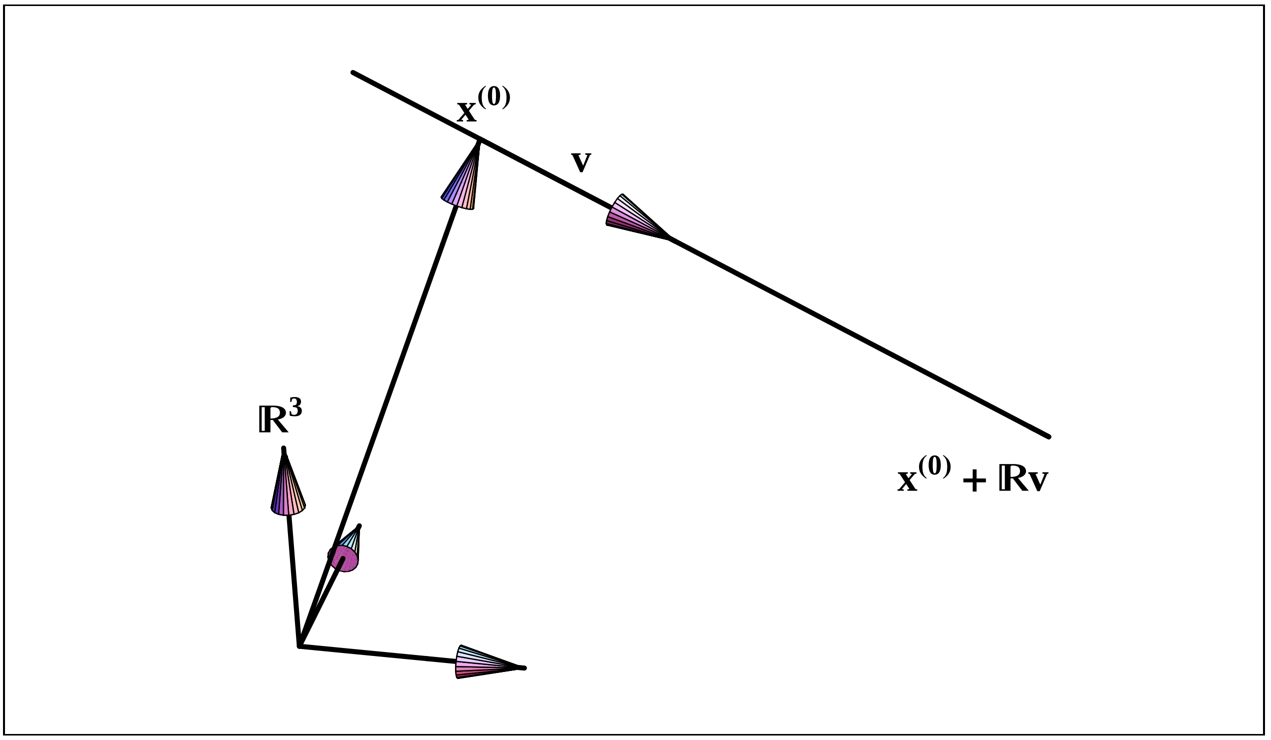

Fig. 2.1: A straight line as a one-dimensional affine subspace of $\mathbb{R}^3$.

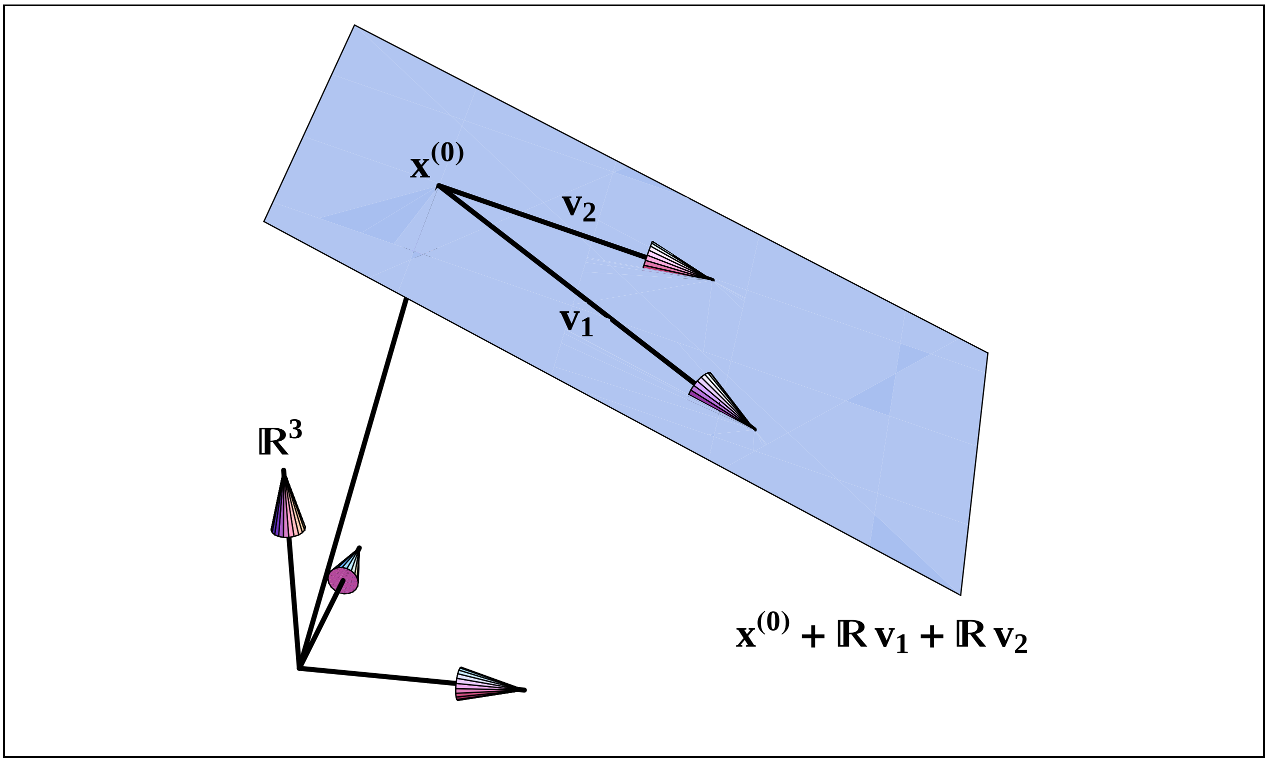

Fig. 2.2: A plane as a two-dimensional affine subspace of $\mathbb{R}^3$.

Using a parameter form yields a construction of M from within: we start with a point in M and then move along straight lines lying in M to get to other points of M; namely, to get from $x^{(0)}$ to $x^{(0)} + u_1 v^{(1)} + \cdots + u_d v^{(d)}$, we need to move u_i units in the direction of $v^{(i)}$, for $1 \leq i \leq d$. Thus the numbers $u_1, \ldots, u_d$ in (2.4), (2.5) and (2.6) form a system of rectilinear coordinates on M which can be used to specify the location of points in M. On the other hand, using an equation form yields a construction of M from the outside: we start with arbitrary points in $\mathbb{R}^n$ and then impose one constraint after the other (each given by an affine equation and each cutting away one possible dimension) until we are reduced to those points which satisfy all the constraints. "Solving" a linear system $Ax = b$ (typically by applying the Gaussian algorithm) means replacing a description of the set $M := \{x \in \mathbb{R}^n \mid Ax = b\}$ in equation form by a description in parameter form. Note, however, that both descriptions have their merits.

• If we want to check whether or not a given point x lies in an affine subspace M, a description in equation form is much more convenient: All we need to do is calculate $Ax - b$ and check if the result is 0. (If M is given in parameter form, we would have to solve the equation $Bu = x - x^{(0)}$ for u and see whether or not a solution $u \in \mathbb{R}^d$ exists.)

• If, on the other hand, we want to explicitly find some points on M, the parameter form is much more suitable: All we need to do is pick a vector $u \in \mathbb{R}^d$ and simply calculate $x := x^{(0)} + Bu$ to specify an element $x \in M$. (If M is given in equation form, we must actually solve the linear system $Ax = b$ to find an element of M.)

We now want to introduce manifolds† as spaces which "locally look like affine spaces". For example, a one-dimensional manifold (a "curve") will be a set which "locally looks like a straight line", and a two-dimensional manifold (a "surface") will be a set which "locally looks like a plane". Of course, we have to define what it means for a set to "locally look like an affine space". In fact, we will give four different definitions, which turn out to be all equivalent. Roughly speaking, a subset $M \subseteq \mathbb{R}^n$ will be called a d-dimensional manifold if it can be locally

- deformed into a d-dimensional affine space;
- represented as the solution set of a system of $n - d$ independent equations;
- parametrized in terms a d independent curvilinear coordinates;
- described as the graph of a mapping $\mathbb{R}^d \to \mathbb{R}^{n-d}$.

Since we want manifolds to be "smooth" objects, we will insist that the deformations, equations, coordinates and mappings occurring in the various definitions will be smooth, i.e., given in terms of differentiable functions (whose degree of differentiability will determine the degree of smoothness of the manifold). We now proceed to precisely define the above characterizations of manifolds and then to show that all four characterizations are equivalent (which will turn out to be a consequence of the Implicit Function Theorem).

† We only discuss finite-dimensional manifolds here. What we call a d-dimensional manifold is sometimes also called a manifold modelled after $\mathbb{R}^d$. Replacing $\mathbb{R}^d$ by an arbitrary Banach space X, we could also (practically verbatim) introduce the concept of a manifold modelled after a Banach space X. Such a manifold is called a **Banach manifold**.

Since the precise formulations of the four possible definitions of a manifold will be somewhat technical, each will be preceded by an informal discussion to form the right intuition. To understand what we mean by "locally deforming a set into an affine space", imagine a curve as a piece of string or wire which is "straightened out" to look like a piece of a straight line, or imagine a surface as a piece of cloth or as part of an orange peel which is "flattened out" to look like a part of a plane. That this can only be done locally is clear if you consider a circle as an example of a curve. Clearly, such a circle does not resemble a straight line globally; the topology of a circle differs markedly from that of a straight line, and it is not possible to twist and turn a circle to make it into a piece of a straight line. (A circle is compact, a line is not. A circle with a point removed is connected, i.e., consists of one piece, whereas a line with a point removed consists of two connected components. A circle has a "hole" which can be technically expressed by saying that a circle is not simply connected, whereas a straight line is simply connected.) However, each point of a circle possesses a neighborhood which looks like the neighborhood of a point on a straight line, so that "locally" a circle and a line are really "the same". Thus while the circle as a whole cannot be deformed into a straight line, any sufficiently small part of a circle can be deformed into a part of a straight line. Similarly, a sufficiently small part of a sphere (i.e., the surface of a ball) or of a torus (the surface of a doughnut) can be deformed ("flattened out") into a piece of a plane, while a sphere or torus as a whole cannot. Finally, since we want manifolds to be "smooth" objects, we insist that the mappings which qualify as "deformations" be not just homeomorphisms, but in fact diffeomorphisms. We are now ready to give our first definition which will be used to characterize manifolds.

(2.1) Definition. *We say that a subset $M \subseteq \mathbb{R}^n$ can be locally deformed into a d-dimensional affine space using a C^k-diffeomorphism, if for each point $p \in M$ there are an open neighborhood $U \subseteq \mathbb{R}^n$, an open set $V \subseteq \mathbb{R}^n$ and a C^k-diffeomorphism $F : U \to V$ such that*

$$F(M \cap U) \;=\; V \cap \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_i = 0 \text{ for } i > d\}.$$

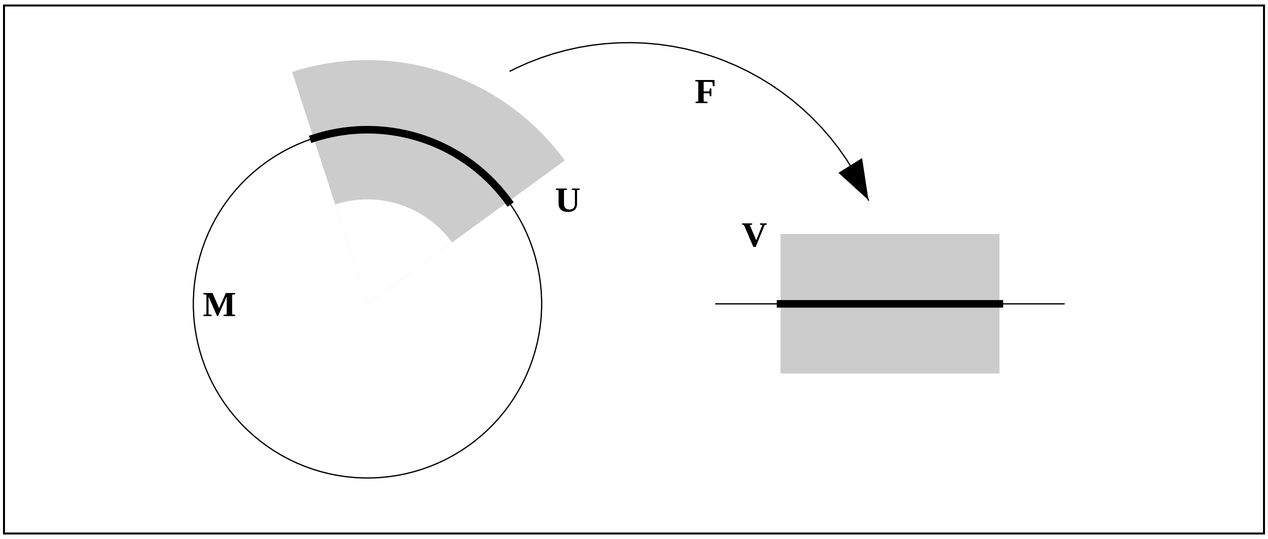

Fig. 2.3: A circle M looks locally like a straight line; i.e., a sufficiently small part of M is smoothly deformable into a part of a straight line. Hence M is a one-dimensional manifold.

In the above definition, we think of the part $M \cap U$ of M as being deformed into a set $\Omega \subseteq \mathbb{R}^d$. However, it does not *a priori* make sense to speak about a deformation of $M \cap U$ without referring to the ambient space, because we would then have to explain when a mapping defined only on M (and not on an open subset of $\mathbb{R}^n$) should be called of class C^k. This is why a deformation is always a deformation between open subsets of the ambient space (even though only the effect of such a deformation on parts of the set M is of interest). Note that it may be possible to deform two sets into each other topologically (i.e., via a homeomorphism), but not smoothly (i.e., via a diffeomorphism).

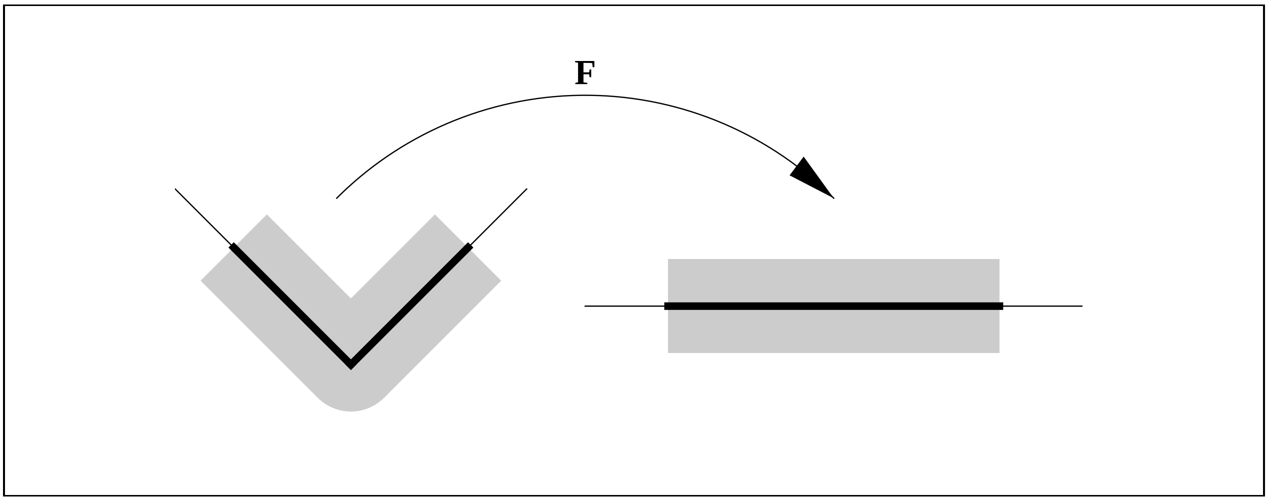

Fig. 2.4: The graph M of the function $y = |x|$ and a straight line can be deformed into each other continuously, but not smoothly. Hence M is not a one-dimensional manifold.

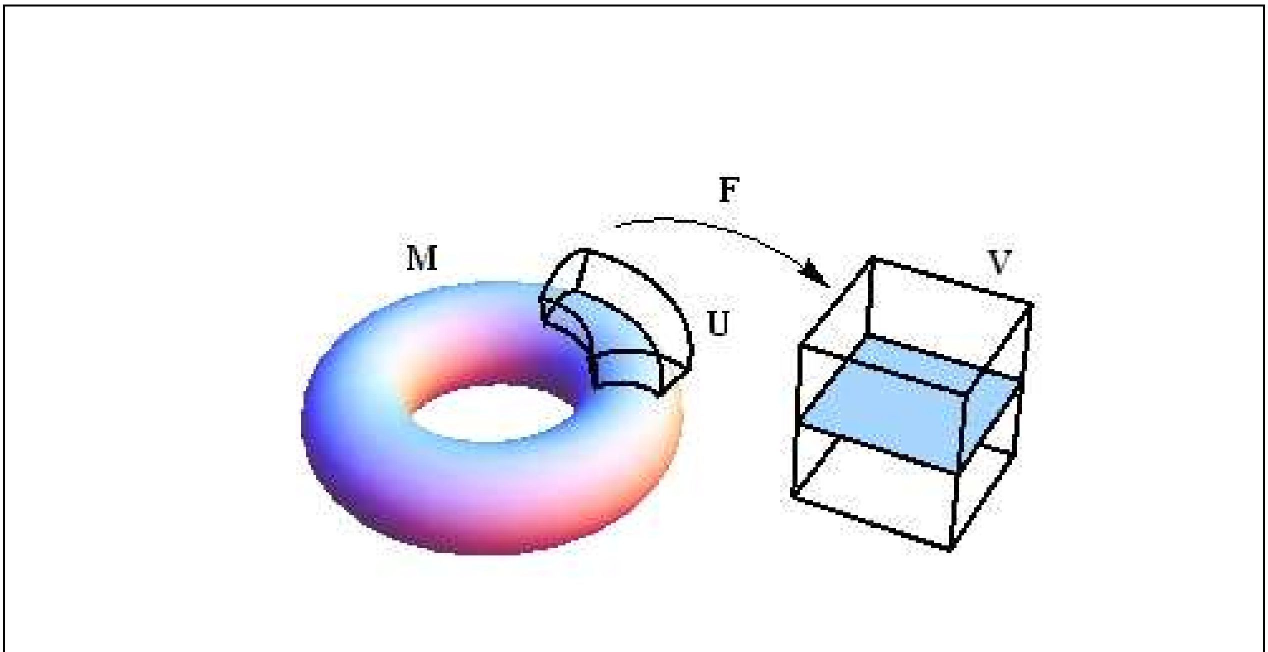

Fig. 2.5: A torus M looks locally like a surface; i.e., a sufficiently small part of M is smoothly deformable into a part of a plane. Hence M is a two-dimensional manifold.

Under deformations, relations between affine spaces turn into relations between manifolds.

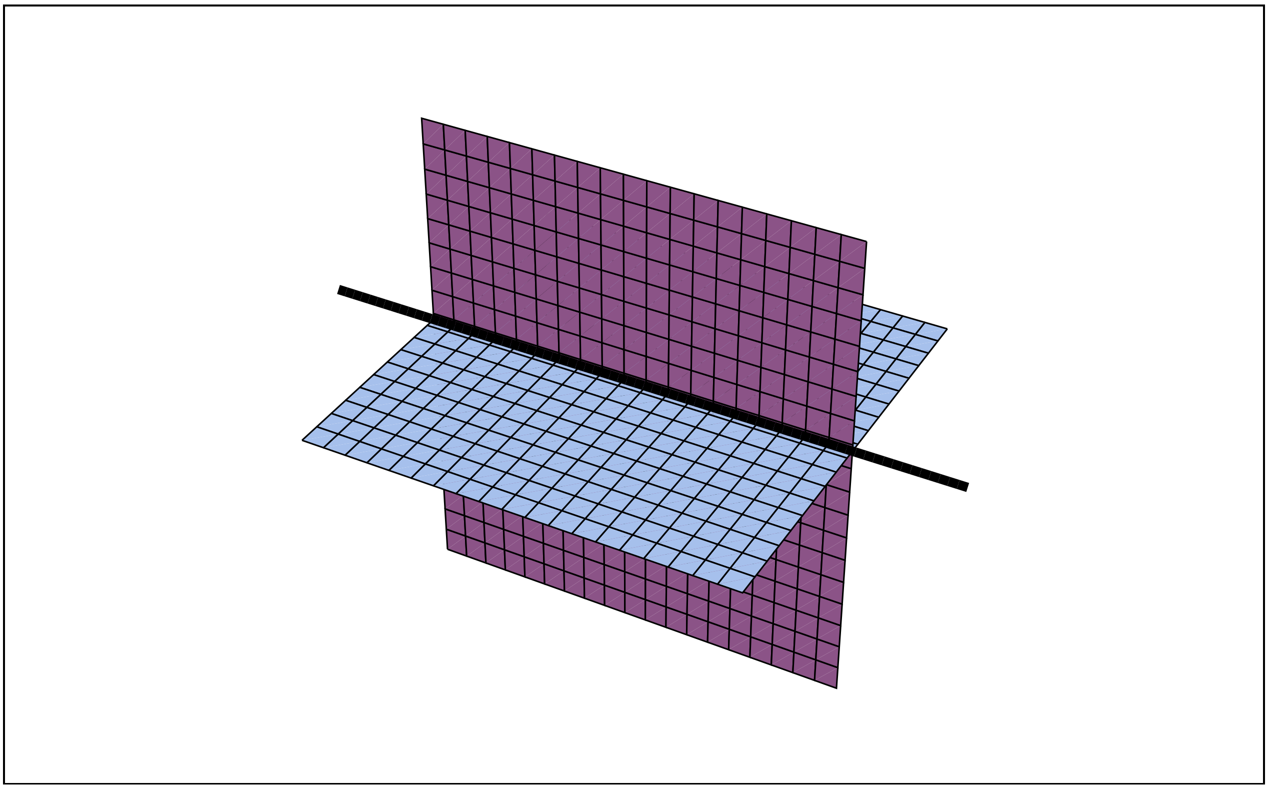

Fig. 2.6: A line as an intersection of two planes.

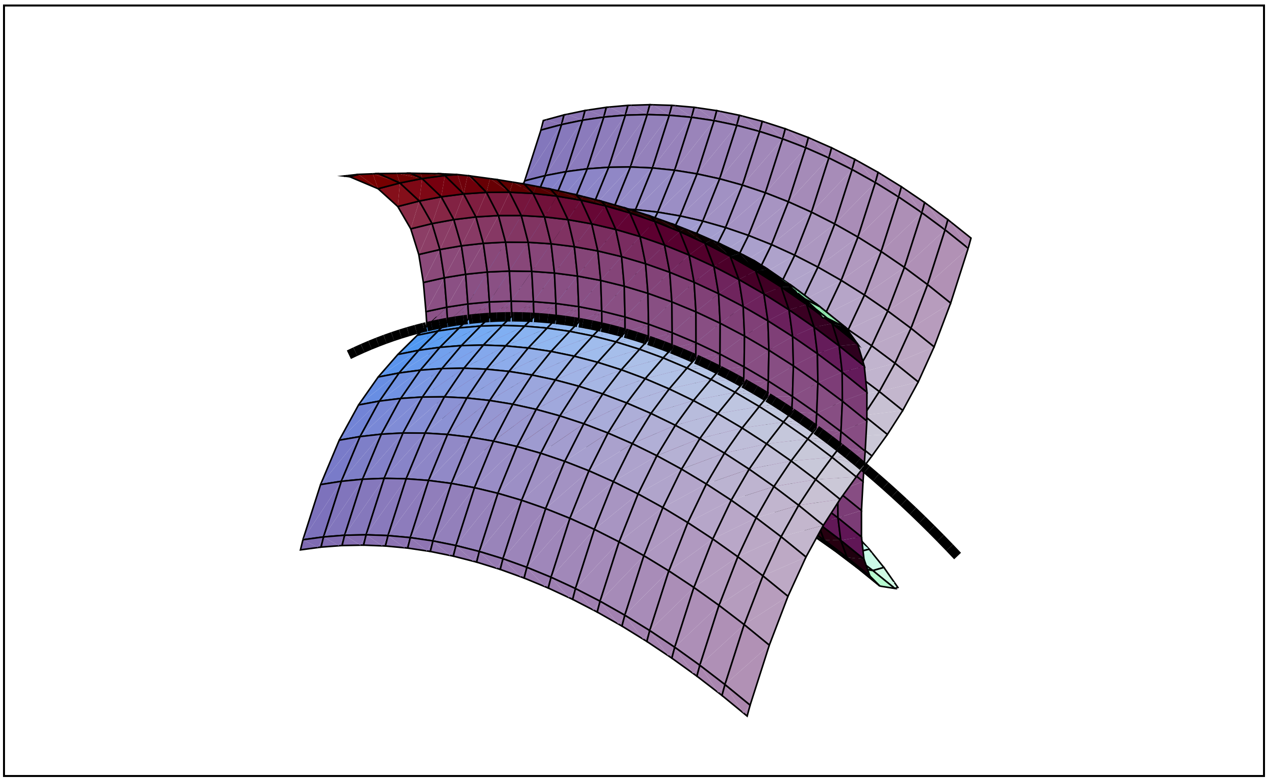

Fig. 2.7: A curve ("deformed line") as an intersection of two surfaces ("deformed planes").

The definition of a d-dimensional manifold as a subset of $\mathbb{R}^n$ which can locally be smoothly deformed into a d-dimensional affine space is geometrically appealing, but it is not the most practical way to check whether or not a given set is a manifold and to study such manifolds. Hence let us turn to the possibility of defining a manifold as the solution set of a system of (generally nonlinear) equations. A d-dimensional affine subspace can be defined as the solution set of a family of $n-d$ equations $g_i(x_1,\ldots,x_n)=0$ where each g_i is an affine function and where the functions g_i are independent in the sense that their linear parts are independent. We now relax this requirement by allowing the functions g_i to be rather arbitrary nonlinear functions; however, we insist that these functions g_i are "smooth" (i.e., have some degree of differentiability) and are independent (which can be made precise by requiring that their linearizations at any point are independent). Since we want manifolds to be spaces which *locally* look like affine space, we do not require that the functions g_i be defined on all of $\mathbb{R}^n$; we require their existence only locally, i.e., in the vicinity of any given point of a manifold. This motivates the following definition.

(2.2) Definition. *Let $M \subseteq \mathbb{R}^n$. We say that M is locally the solution set of $n-d$ independent equations of class C^k if for each point $p \in M$ there exist an open neighborhood $U \subseteq \mathbb{R}^n$ and $n-d$ functions $g_1,\ldots,g_{n-d}: U \to \mathbb{R}$ of class C^k, whose derivatives $g_1'(x),\ldots,g_{n-d}'(x)$ are linearly independent as linear forms on $\mathbb{R}^n$ at each point $x \in U$, such that*

$$M \cap U \;=\; \{x \in U \mid g_1(x) = \cdots = g_{n-d}(x) = 0\}\,.$$

Equivalently (by treating the real-valued functions g_i as component functions of a vector-valued mapping g) this can be rephrased by saying that for each point $p \in M$ there exist an open neighborhood $U \subseteq \mathbb{R}^n$ and a function $g: U \to \mathbb{R}^{n-d}$ of class C^k such that $\operatorname{rk} g'(x) = n-d$ *at each point $x \in U$ and such that*

$$M \cap U \;=\; \{x \in U \mid g(x) = 0\}\,.$$

Thus the independence of the functions g_i is defined as the independence of their linearizations $g_i'(x)$ or, equivalently, of their gradients $(\nabla g_i)(x)$, at any point $x \in U$. Points x at which this condition is not satisfied for a given family of functions g_i (i.e, points at which $\operatorname{rk} g'(x) < n-d$) are called **singular points** of g.

Similarly, it is possible to define (and study) manifolds via parametrizations. Recall that a d-dimensional affine subspace of $\mathbb{R}^n$ can be defined as the set M of all vectors $\varphi(u_1,\ldots,u_d)$ where the parametrization $\varphi: \mathbb{R}^d \to \mathbb{R}^n$ is an injective affine function. (The injectivity is equivalent to the condition that the rank of the linear part of φ is d, and this condition ensures that each point on M can be uniquely specified by numbers $u_1,\ldots,u_d$, which thus serve as rectilinear coordinates on M.) We now relax the requirement that φ be an affine map and allow rather arbitrary nonlinear parametrizations; however, we insist that the parametrizations used are smooth and that their linearizations at each point of their domain are injective. Since we want manifolds to be spaces which *locally* look like affine spaces, we only require parametrizations which are defined on some open subset of $\mathbb{R}^d$ (not necessarily all of $\mathbb{R}^d$) and which cover some portion of the manifold M in question (not necessarily all of M). Thus we arrive at the following formal definition.

(2.3) Definition. *Let $M \subseteq \mathbb{R}^n$. We say that M can be locally C^k-parametrized by d parameters if for each point $p \in M$ there are an open set $\Omega \subseteq \mathbb{R}^d$ (called the* **parameter domain***), an open neighborhood $U \subseteq \mathbb{R}^n$ of p and a C^k-mapping $\varphi : \Omega \to \mathbb{R}^n$ (called* **parametrization***), which maps Ω homeomorphically onto $\varphi(\Omega)$ and for which, at each point $\xi \in \Omega$, the partial derivatives $(\partial_1\varphi)(\xi), \ldots, (\partial_d\varphi)(\xi) \in \mathbb{R}^n$ are linearly independent, such that*

$$M \cap U \;=\; \{\varphi(\xi_1, \ldots, \xi_d) \mid (\xi_1, \ldots, \xi_d) \in \Omega\}.$$

The numbers $\xi_1, \ldots, \xi_d$ are considered as curvilinear coordinates on $M \cap V$. The condition that, at any given point $\xi \in \Omega$, the partial derivatives $(\partial_i\varphi)(\xi)$ (where $1 \leq i \leq d$) are linearly dependent means geometrically that the images of the coordinate lines in Ω really form a grid of coordinate curves on M which point into d different directions at any point $\varphi(\xi)$. This condition can be restated by saying that the rank of $\varphi'(\xi)$ is d, i.e., that $\varphi'(\xi)$ is injective. Points where this condition is not satisfied for a given mapping φ are called **singular points** of φ.

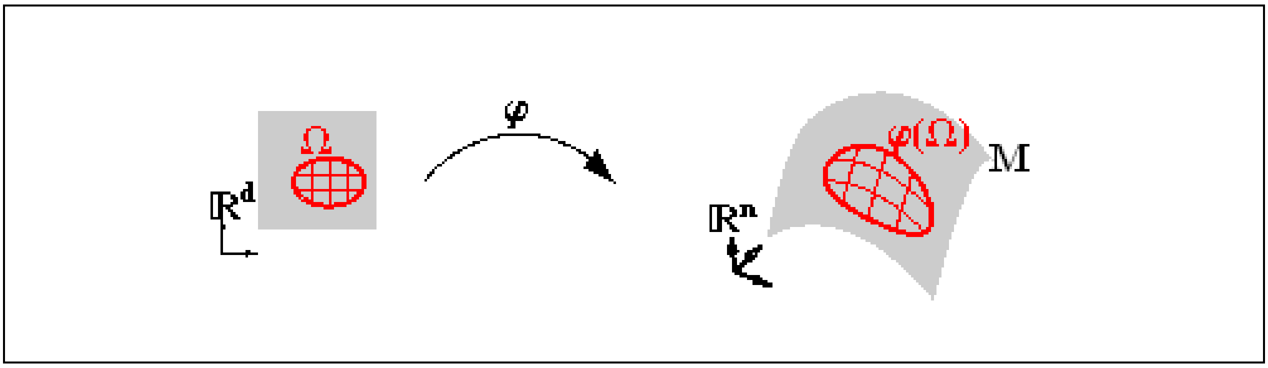

Fig. 2.8: A parametrization deforms a rectilinear system of coordinates in the parameter domain into a system of curvilinear coordinates on a manifold.

The concepts of deformation and parametrization are closely related to each other.

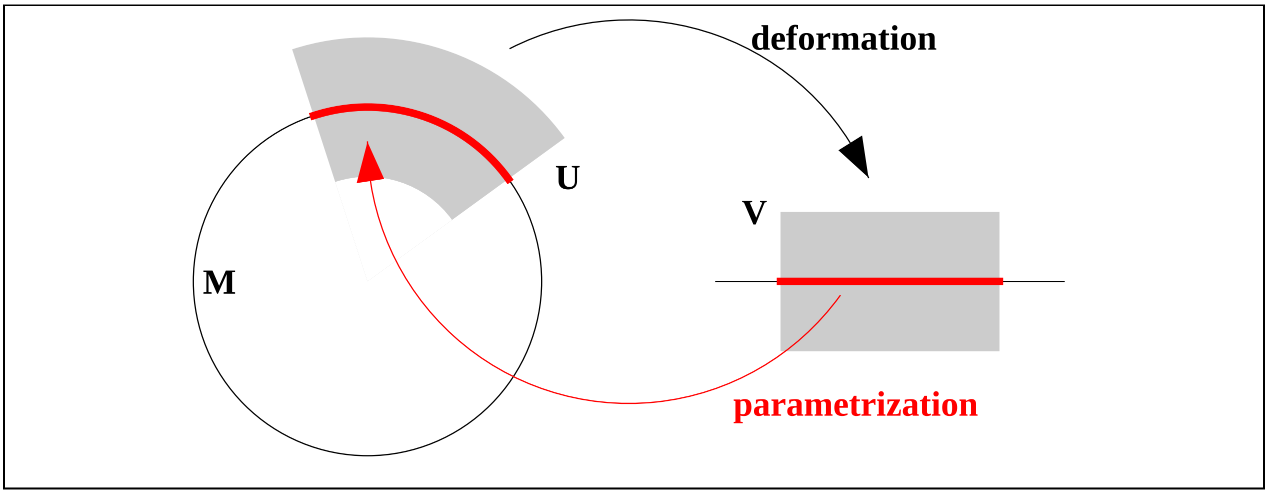

Fig. 2.9: Relation between deformations and parametrizations.

Given a diffeomorphism $F : U \to V$ between open subsets $U, V \subseteq \mathbb{R}^n$ which deforms $M \cap U$ into $\Omega \subseteq V \cap (\mathbb{R}^d \times \{0\}) \subseteq \mathbb{R}^d \times \mathbb{R}^{n-d} = \mathbb{R}^n$, we can identify Ω with a subset of $\mathbb{R}^d$ and then obtain a parametrization

$$\varphi \;:=\; F^{-1}\big|_{\Omega} : \Omega \to M \cap U\,.$$

Conversely, a given parametrization φ can be extended to a diffeomorphism

$$\Phi : \Omega \times V_0 \to U_{\text{open}} \subseteq \mathbb{R}^n$$

where V_0 is a sufficiently small open neighborhood of 0 in $\mathbb{R}^{n-d}$; then the inverse of Φ is a diffeomorphism which deforms $M \cap U$ into Ω. Hence we can (somewhat sloppily) say that parametrizations and deformations of part of a manifold are inverses of each other.

There is one instance where a set can be parametrized in a particularly simple way, namely, if this set (say $M \subseteq \mathbb{R}^n$) is the graph of some function $f : \Omega \to \mathbb{R}^{n-d}$ where $\Omega \subseteq \mathbb{R}^d$. In this case, a parametrization $\varphi : \Omega \to \mathbb{R}^d \times \mathbb{R}^{n-d} = \mathbb{R}^n$ if M is given by

$$\varphi(\xi_1, \ldots, \xi_d) \;:=\; \big(\xi_1,\, \ldots,\, \xi_d,\, f(\xi_1, \ldots, \xi_d)\big).$$

(The condition $\operatorname{rk} \varphi'(\xi) = d$ is trivially satisfied for such a parametrization.) The special nature of the set M is that, given a point $x \in M$, the last $n-d$ coordinates of x can be expressed as functions of the first d coordinates of x. We now want to allow permutations of the coordinates; i.e., we want to study sets M with the property that, in a neighborhood U of any given point $p \in M$, we can specify a set of d coordinates such that the remaining $n-d$ coordinates of $x \in U$ can be expressed in terms of these chosen d coordinates by specified functions. The precise definition is as follows. (Recall that the symmetric group Sym_n consists of all permutations of the set $\{1, \ldots, n\}$, i.e., has all bijective mappings $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ as its elements.)

(2.4) Definition. *Let $M \subseteq \mathbb{R}^n$. We say that M is locally the graph of a function $f : \mathbb{R}^n \to \mathbb{R}^{n-d}$ of class C^k if for each point $p = (p_1, \ldots, p_n) \in M$ there are an index permutation $\sigma \in \mathrm{Sym}_n$, open neighborhoods $U' \subseteq \mathbb{R}^d$ of $p' := (p_{\sigma(1)}, \ldots, p_{\sigma(d)})$ and $U'' \subseteq \mathbb{R}^{n-d}$ of $p'' := (p_{\sigma(d+1)}, \ldots, p_{\sigma(n)})$ and a C^k-mapping $f : U' \to U''$ such that, writing $x' := (x_{\sigma(1)}, \ldots, x_{\sigma(d)})$ and $x'' := (x_{\sigma(d+1)}, \ldots, x_{\sigma(n)})$, we have*

$$M \cap (U' \times U'') \;=\; \{(x', x'') \in U' \times U'' \mid x'' = f(x')\}.$$

We will now show that, as a consequence of the Implicit Function Theorem and of the Inverse Function Theorem, the four definitions (2.1), (2.2), (2.3) and (2.4) are all equivalent.

(2.5) Theorem and Definition. *For a set $M \subseteq \mathbb{R}^n$ the following four conditions are equivalent:*

(1) *M is locally the solution set of $n-d$ independent equations of class C^k; i.e., for each point $p \in M$ there exist an open neighborhood $U \subseteq \mathbb{R}^n$ and a function $g : U \to \mathbb{R}^{n-d}$ of class C^k of full rank $n-d$ such that*

$$M \cap U \;=\; \{x \in U \mid g(x) = 0\}\,; \tag{C1}$$

(2) *M is locally the graph of a function $f : \mathbb{R}^d \to \mathbb{R}^{n-d}$ of class C^k; i.e., for each $p \in M$ there are an index permutation $\sigma \in \mathrm{Sym}_n$, open neighborhoods $U' \subseteq \mathbb{R}^d$ of $p' := (p_{\sigma(1)}, \ldots, p_{\sigma(d)})$ and $U'' \subseteq \mathbb{R}^{n-d}$ of $p'' := (p_{\sigma(d+1)}, \ldots, p_{\sigma(n)})$ and a C^k-mapping $f : U' \to U''$ such that, writing $x' := (x_{\sigma(1)}, \ldots, x_{\sigma(d)})$ and $x'' := (x_{\sigma(d+1)}, \ldots, x_{\sigma(n)})$, we have*

$$M \cap (U' \times U'') \;=\; \{(x', x'') \in U' \times U'' \mid x'' = f(x')\}; \tag{C2}$$

(3) *M can locally be deformed into a d-dimensional affine space using a C^k-diffeomorphism; i.e., for each $p \in M$ there are an open neighborhood $U \subseteq \mathbb{R}^n$, an open set $V \subseteq \mathbb{R}^n$ and a C^k-diffeomorphism $F : U \to V$ such that*

$$F(M \cap U) \;=\; V \cap \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_i = 0 \textit{ for } i > d\}; \tag{C3}$$

(4) *M can be locally C^k-parametrized by d parameters; i.e., for each $p \in M$ there are an open set $\Omega \subseteq \mathbb{R}^d$, an open neighborhood $U \subseteq \mathbb{R}^n$ of p and a C^k-mapping $\varphi : \Omega \to \mathbb{R}^n$ of full rank d which is a homeomorphism onto its image such that*

$$M \cap U \;=\; \{\varphi(u_1, \ldots, u_d) \mid (u_1, \ldots, u_d) \in \Omega\}. \tag{C4}$$

If these conditions are satisfied, we call M a d-dimensional **manifold** *of class C^k or, more specifically, a d-dimensional* **embedded submanifold** *of $\mathbb{R}^n$ of class C^k.* †

Proof. (1)⟹(2). Let $p \in M$ and let U and g as in (C1). Since $\mathrm{rk}\, g'(p) = n-d$, there are indices $i_1 < \cdots < i_{n-d}$ such that $\det\big(\partial(g_1, \ldots, g_{n-d})/\partial(x_{i_1}, \ldots, x_{i_{n-d}})\big)(p) \neq 0$. Then the Implicit Function Theorem states that there are open neighborhoods $U' \subseteq \mathbb{R}^d$ of p' and $U'' \subseteq \mathbb{R}^{n-d}$ of p'' with $U' \times U'' \subseteq U$ and a C^k-function $f : U' \to U''$ such that (C2) holds.

(2)⟹(3). Let $p \in M$ and let U', U'' and f as in (C2). Let $U := U' \times U''$ and define $F : U \to \mathbb{R}^n$ by $F(x', x'') := \big(x', x'' - g(x')\big)$. Then $V := F(U)$ is open in $\mathbb{R}^n$, and $F : U \to V$ is a diffeomorphism such $F(M \cap U) = V \cap \{(x', x'') \mid x'' = 0\}$, so that (C3) holds.

(3)⟹(4). Let $p \in M$ and let $F : U \to V$ as in (C3). Identifying $\mathbb{R}^d$ with $\{x \in \mathbb{R}^n \mid x_{d+1} = \cdots = x_n = 0\}$, we let $\Phi := F^{-1} : V \to U$ and $\Omega := U \cap \mathbb{R}^d$; then $\varphi := \Phi\,|_\Omega$ is a parametrization satisfying (C4).

† For the moment we will use the expression “manifold” synonymously with “embedded submanifold of some space $\mathbb{R}^n$”. Later on, we will study abstract manifolds which are not *a priori* given as subsets of some ambient space $\mathbb{R}^n$.

(4)$\Longrightarrow$(3). Let $\xi \in \Omega$. Since $\operatorname{rk} \varphi'(\xi) = d$ we may assume (after an index permutation, if necessary) that $\det\bigl(\partial(\varphi_1, \ldots, \varphi_d)/\partial(\xi_1, \ldots, \xi_d)\bigr)(\xi) \neq 0$. Then the Inverse Function Theorem guarantees the existence of an open neighborhood $\Omega_0 \subseteq \Omega$ of ξ and of an open set $W \subseteq \mathbb{R}^d$ such that $\widehat{\varphi} := (\varphi_1, \ldots, \varphi_d) : \Omega_0 \to W$ is a C^k-diffeomorphism. For $\xi \in \mathbb{R}^n$ we let $\xi' := (\xi_1, \ldots, \xi_d)$ and $\xi'' := (\xi_{d+1}, \ldots, \xi_n)$ and write $\xi = (\xi', \xi'')$. Then we define $\Phi : \Omega_0 \times \mathbb{R}^{n-d} \to W \times \mathbb{R}^{n-d}$ by $\Phi(\xi', \xi'') := \varphi(\xi') + (0, \xi'')$, i.e.,

$$\Phi(\xi_1, \ldots, \xi_n) \;:=\; \begin{bmatrix} \varphi_1(\xi_1, \ldots, \xi_d) \\ \vdots \\ \varphi_d(\xi_1, \ldots, \xi_d) \\ \varphi_{d+1}(\xi_1, \ldots, \xi_d) + \xi_{d+1} \\ \vdots \\ \varphi_n(\xi_1, \ldots, \xi_d) + \xi_n \end{bmatrix}.$$

Then Φ is a C^k-diffeomorphism from $V := \Omega_0 \times \mathbb{R}^{n-d}$ onto $W \times \mathbb{R}^{n-d}$ with $\Phi(\Omega_0 \times \{0\}) = \varphi(\Omega_0)$; hence $F := \Phi^{-1}$ satisfies $\Omega_0 \times \{0\} = F\bigl(\varphi(\Omega_0)\bigr)$. Since $\varphi(\Omega_0)$ is open in $\varphi(\Omega)$ and hence in M, there is a neighborhood $U \subseteq \mathbb{R}^n$ of p with $\varphi(\Omega_0) = M \cap U$. It follows that $F(M \cap U) = F\bigl(\varphi(\Omega_0)\bigr) = \Omega_0 \times \{0\} = V \cap (\mathbb{R}^d \times \{0\})$.

(3)$\Longrightarrow$(1). Condition (C3) implies $F(M \cap U) = F(U) \cap (\mathbb{R}^d \times \{0\})$; hence $M \cap U = \{x \in U \mid F_{d+1}(x) = \cdots = F_n(x) = 0\}$, so that (C1) holds with $g_i := F_{d+i}$ for $1 \leq i \leq n-d$. ■

(2.6) Example. The **unit sphere** in $\mathbb{R}^n$ is the set

$$\mathbb{S}^{n-1} \;:=\; \{x \in \mathbb{R}^n \mid \|x\| = 1\} \;=\; \{x \in \mathbb{R}^n \mid x_1^2 + \cdots + x_n^2 = 1\},$$

i.e., the zero set of the function $g : \mathbb{R}^n \to \mathbb{R}$ given by $g(x_1, \ldots, x_n) := x_1^2 + \cdots + x_n^2 - 1$. Since $g'(x) = (2x_1, \ldots, 2x_n)$ is never zero for $x \in \mathbb{S}^{n-1}$, the rank of $g'(x)$ is one for each $x \in \mathbb{S}^{n-1}$, which shows that $\mathbb{S}^{n-1}$ is a manifold of dimension $n-1$. (Since g is analytic, even polynomial, this manifold is of class C^ω.)

By its very definition, $\mathbb{S}^{n-1}$ is given as the zero set of a function $g : \mathbb{R}^n \to \mathbb{R}$ of constant rank 1. The easiest way to verify the other (equivalent) characterizations of a manifold is by studying the intersections of $\mathbb{S}^{n-1}$ with the open half-spaces

$$U_i^+ := \{x \in \mathbb{R}^n \mid x_i > 0\} \qquad \text{and} \qquad U_i^- := \{x \in \mathbb{R}^n \mid x_i < 0\}.$$

The intersections $U_i^+ \cap \mathbb{S}^{n-1}$ and $U_i^- \cap \mathbb{S}^{n-1}$ are open hemispheres which cover all of $\mathbb{S}^{n-1}$. Each such hemisphere is the graph of an analytic function; more precisely, letting $\Omega := \{\xi \in \mathbb{R}^{n-1} \mid \|\xi\| < 1\}$, the hemisphere $H_i^\pm$ is the graph of the function $f : \Omega \to \mathbb{R}^n$ given by

$$f(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \;=\; \pm\sqrt{1 - \sum\nolimits_{j \neq i} x_j^2}\,.$$

A diffeomorphism $F : \Omega \times \mathbb{R} \to \Omega \times \mathbb{R}$ which deforms Ω into H_i^+ is given by

$$F(x_1, \ldots, x_n) \; := \; \Big(x_1, \, \ldots, \, x_{i-1}, \, x_i \mp \sqrt{1 - \sum_{j \neq i} x_j^2}, \, x_{i+1}, \, \ldots, \, x_n\Big)$$

where we identify Ω with $\{(x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_n) \mid \sum_{j\neq i} x_j^2 < 1\}$. Finally, to parametrize $H_i^\pm$ we can use the mapping $\varphi : \Omega \to \mathbb{R}^n$ given by

$$\varphi(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \; := \; \Big(x_1, \, \ldots, \, x_{i-1}, \, \pm\sqrt{1 - \sum_{j \neq i} x_j^2}, \, x_{i+1}, \, \ldots, \, x_n\Big).$$

There are altogether $2n$ open hemispheres $H_i^\pm$, and since $(0, \ldots, \pm 1, \ldots 0)$ lies in $H_i^\pm$, but not in any of the hemispheres $H_j^\pm$ for $j \neq i$, all $2n$ are needed to cover all of $\mathbb{S}^{n-1}$. There is a different way to parametrize $\mathbb{S}^{n-1}$ which requires only two different parameter domains, namely, by **stereographic projection**.

Here we identify $\mathbb{R}^n$ with $\mathbb{R}^{n-1} \times \mathbb{R}$; we call $N = (0, 1)$ the **north pole** and $\mathbb{S}^{n-1} \cap (\mathbb{R}^{n-1} \times \{0\})$ the **equator** of $\mathbb{S}^{n-1}$. Then the stereographic projection $\varphi : \mathbb{R}^{n-1} \to \mathbb{S}^{n-1} \backslash \{N\}$ is defined as follows: Given $\xi \in \mathbb{R}^{n-1}$, let $\varphi(\xi)$ be the unique point of intersection of the line segment between $(\xi, 0)$ and $(0, 1) = N$ with $\mathbb{S}^{n-1}$.

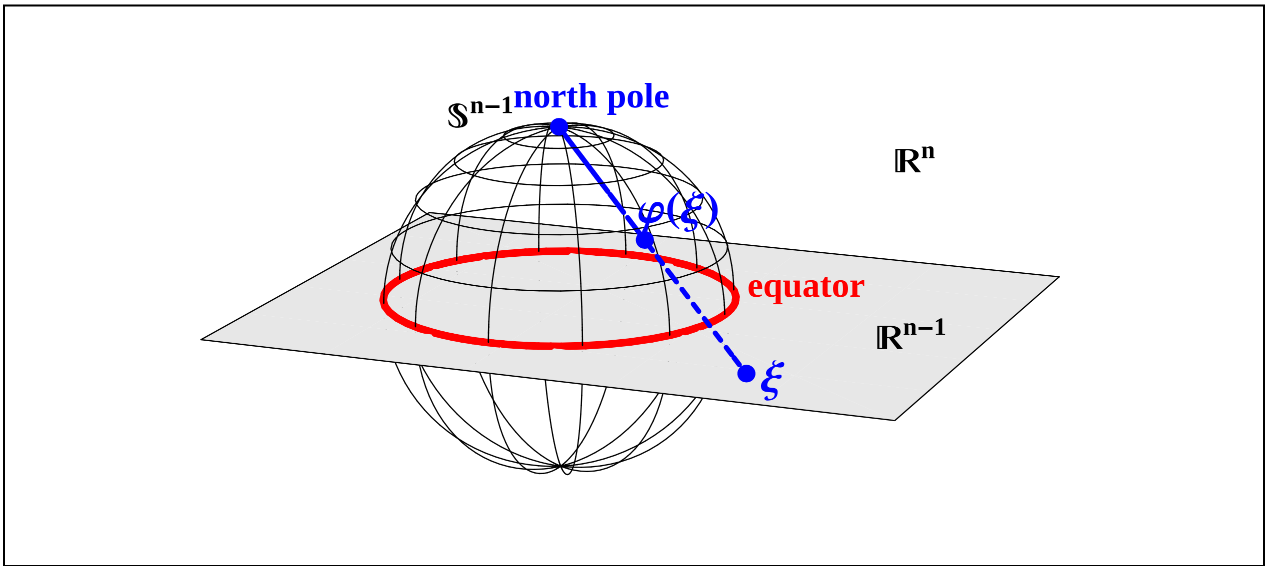

Fig. 2.10: Stereographic projection $\varphi : \mathbb{R}^{n-1} \to \mathbb{S}^{n-1} \setminus \{\text{north pole}\}$.

Then φ parametrizes all of $\mathbb{S}^{n-1}$ with the single exception of the north pole. Defining a second parametrization which covers all of $\mathbb{S}^{n-1}$ with the exception of one other point (for example the south pole), we see that $\mathbb{S}^{n-1}$ can be covered by two local parametrizations. To find an explicit formula for φ, we note that

$$\varphi(\xi) \; = \; \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \lambda \begin{bmatrix} \xi - 0 \\ 0 - 1 \end{bmatrix} \; = \; \begin{bmatrix} \lambda\xi \\ 1 - \lambda \end{bmatrix}$$

where λ is the unique number in $(0,1)$ for which $1 = \|\varphi(\xi)\|^2 = \lambda^2\|\xi\|^2 + (1-\lambda)^2 = (1+\|\xi\|^2)\lambda^2 - 2\lambda + 1$ and hence $0 = (1+\|\xi\|^2)\lambda^2 - 2\lambda = \lambda \cdot \big((1+\|\xi\|^2) - 2\big)$, which yields $\lambda = 2/(1+\|\xi\|^2)$. (The solution $\lambda = 0$ belongs to the north pole N.) Thus we have

$$\varphi(\xi) \;=\; \frac{1}{\|\xi\|^2+1}\begin{bmatrix} 2\xi \\ \|\xi\|^2 - 1\end{bmatrix}.$$

A different parametrization of $\mathbb{S}^{n-1}$ is by **polar coordinates**, which we now discuss. Generalizing polar coordinates in the plane and spherical coordinates in space, we define n-dimensional polar coordinates by the mapping

$$\Phi(r,\theta_1,\ldots,\theta_{n-1}) \;:=\; r\, x_{n-1}(\theta_1,\ldots,\theta_{n-1})$$

where $x_n : \mathbb{R}^n \to \mathbb{S}^n$ is defined by

$$(\star) \qquad x_n(\lambda,\theta_1,\ldots,\theta_{n-2}) \;:=\; \begin{bmatrix} \cos\lambda\cos\theta_1\cos\theta_2\cdots\cos\theta_{n-3}\cos\theta_{n-2} \\ \sin\lambda\cos\theta_1\cos\theta_2\cdots\cos\theta_{n-3}\cos\theta_{n-2} \\ \sin\theta_1\cos\theta_2\cdots\cos\theta_{n-3}\cos\theta_{n-2} \\ \sin\theta_2\cdots\cos\theta_{n-3}\cos\theta_{n-2} \\ \vdots \\ \sin\theta_{n-3}\cos\theta_{n-2} \\ \sin\theta_{n-2}\end{bmatrix}$$

We want to calculate the Jacobian determinant $\partial\Phi(r,\theta_1,\ldots,\theta_{n-1})/\partial(r,\theta_1,\ldots,\theta_{n-1})$. Writing for short x_{n-1} instead of $x_{n-1}(\theta_1,\ldots,\theta_{n-1})$, we have the recursive formula $x_n = \cos(\theta_n)\begin{bmatrix} x_{n-1} \\ 0\end{bmatrix} + \sin(\theta_n)\begin{bmatrix} 0 \\ 1\end{bmatrix}$ which, written more explicitly, means

$$(\star\star) \qquad x_n(\theta_1,\ldots,\theta_n) \;=\; \begin{bmatrix} \cos(\theta_n)\, x_{n-1}(\theta_1,\ldots,\theta_{n-1}) \\ \sin(\theta_n)\end{bmatrix}.$$

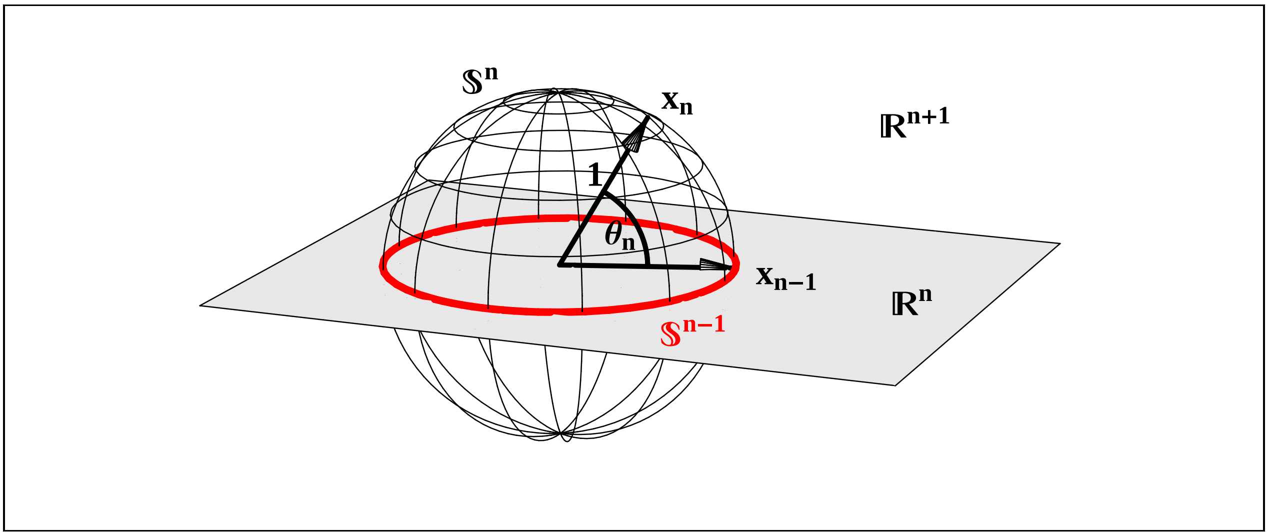

Fig. 2.11: Parametrization of $\mathbb{S}^{n-1}$ (with a hemisphere in dimension $n-2$ removed) by polar coordinates.

The mapping x_1 is one-to-one on $[0, 2\pi)$ (or any other semi-open interval of length 2π). Since the domain of a parametrization should be an open set, we restrict x_1 to an open interval of length 2π; thus x_1 parametrizes the unit circle with one point removed. The angles θ_k with $k > 1$ are defined on the interval $[-\pi/2, \pi/2]$; however, we have to remove the boundary points $\pm\pi/2$ (which correspond to the north and south pole of $\mathbb{S}^n$) to ensure injectivitity. Thus we restrict x_n to an injective mapping

$$x_n : (0, 2\pi) \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \times \cdots \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \to \mathbb{R}^{n+1}$$

whose image is the sphere $\mathbb{S}^n$ with half of a sphere of dimension $n - 1$ removed. (For example, spherical coordinates in $\mathbb{R}^3$ parametrize $\mathbb{S}^2$ with half a great circle joining the north and the south poles removed.) Letting

$$A_{n-1} := \frac{\partial x_{n-1}(\theta_1, \ldots, \theta_n)}{\partial(\theta_1, \ldots, \theta_{n-1})} \in \mathbb{R}^{n\times(n-1)} \qquad \text{and} \qquad J_n := \frac{\partial \Phi(r, \theta_1, \ldots, \theta_{n-1})}{\partial(r, \theta_1, \ldots, \theta_{n-1})} \in \mathbb{R}^{n\times n},$$

our goal is to determine ther Jacobian determinant $\det(J_n)$. Since $\Phi(r, \theta_1, \ldots, \theta_{n-1}) = r \cdot x_{n-1}(\theta_1, \ldots, \theta_{n-1})$, we have the block representation

$$J_n = [x_{n-1} \mid rA_{n-1}] .$$

Moreover, because of ($\star\star$) above, we have the recursive formula

$$A_n = \begin{bmatrix} \cos(\theta_n)A_{n-1} & -\sin(\theta_n)x_{n-1} \\ \mathbf{0} & \cos(\theta_n) \end{bmatrix} .$$

For $J_{n+1} = [x_n \mid rA_n]$ this yields

$$J_{n+1} = \begin{bmatrix} \cos(\theta_n)x_{n-1} & r\cos(\theta_n)A_{n-1} & -r\sin(\theta_n)x_{n-1} \\ \sin(\theta_n) & \mathbf{0} & r\cos(\theta_n) \end{bmatrix}$$

and, after expansion with respect to the last row, also

$$\det(J_{n+1}) = (-1)^{n-1}\sin(\theta_n) \cdot D_1 + r\cos(\theta_n) \cdot D_2$$

where

$$\begin{aligned} D_1 &:= \det[r\cos(\theta_n)A_{n-1} \mid -r\sin(\theta_n)x_{n-1}] \\ &= -r\sin(\theta_n)\cos(\theta_n)^{n-1}\det[rA_{n-1} \mid x_{n-1}] \\ &= (-1)^{n+1}r\sin(\theta_n)\cos(\theta_n)^{n-1}\det[x_{n-1} \mid rA_{n-1}] \\ &= (-1)^{n+1}r\sin(\theta_n)\cos(\theta_n)^{n-1}\det(J_n) \end{aligned}$$

and

$$\begin{aligned} D_2 &:= \det[\cos(\theta_n)x_{n-1} \mid r\cos(\theta_n)A_{n-1}] \\ &= \cos(\theta_n)^n \det[x_{n-1} \mid rA_{n-1}] = \cos(\theta_n)^n \det(J_n). \end{aligned}$$

Thus we obtain

$$\begin{aligned}\det(J_{n+1}) &= (-1)^{n-1}\sin(\theta_n)\cdot D_1 + r\cos(\theta_n)\cdot D_2\\ &= r\sin(\theta_n)^2\cos(\theta_n)^{n-1}\det(J_n) + r\cos(\theta_n)^{n+1}\det(J_n)\\ &= r\cos(\theta_n)^{n-1}(\sin(\theta_n)^2+\cos(\theta_n)^2)\det(J_n)\\ &= r\cos(\theta_n)^{n-1}\det(J_n).\end{aligned}$$

Starting with the well-known formula

$$\det(J_1) = \det\begin{bmatrix}\cos(\theta_1) & -r\sin(\theta_1)\\ \sin(\theta_1) & r\cos(\theta_1)\end{bmatrix} = r$$

for two-dimensional polar coordinates we find

$$\begin{aligned}&\det(J_1) = r,\\ &\det(J_2) = r\cos(\theta_2)\cdot\det(J_1) = r^2\cos(\theta_2),\\ &\det(J_3) = r\cos(\theta_3)^2\cdot\det(J_2) = r^3\cos(\theta_2)\cos(\theta_3)^2,\\ &\det(J_4) = r\cos(\theta_4)^3\cdot\det(J_3) = r^4\cos(\theta_2)\cos(\theta_3)^2\cos(\theta_4)^3\end{aligned}$$

and, in general,

$$\det(J_n) = r^n\cos(\theta_2)\cos(\theta_3)^2\cdots\cos(\theta_n)^{n-1}.$$

Letting $r := 1$, this results in a parametrization of the sphere $\mathbb{S}^{n-1}$ in terms of $n-1$ angles. This parametrization is regular on all of $\mathbb{S}^{n-1}$ with half a meridian (which is one half of an $(n-2)$-dimensional sphere) removed.

(2.7) Example. The orthogonal group

$$\mathrm{O}(3) = \{A\in\mathbb{R}^{3\times 3} \mid A^TA = \mathbf{1}\}$$

is a 3-dimensionale analytical submanifold of $\mathbb{R}^{3\times 3}$. This can be seen as follows: Writing a real (3×3)-matrix as $(a|b|c)$ with the column vectors $a,b,c\in\mathbb{R}^3$, we see that $\mathrm{O}(3)$ is the zero set of the analytical functions

$$\begin{aligned}g_1(a,b,c) &:= \|a\|^2-1,\\ g_2(a,b,c) &:= \|b\|^2-1,\\ g_3(a,b,c) &:= \|c\|^2-1,\\ g_4(a,b,c) &:= \langle a,b\rangle,\\ g_5(a,b,c) &:= \langle a,c\rangle,\\ g_6(a,b,c) &:= \langle b,c\rangle.\end{aligned}$$

The gradients $(\nabla g_i)(A)$ of these functions are linearly independent at any element $A\in\mathrm{O}(3)$, as can be seen by verifying that, given $(a|b|c)\in\mathrm{O}(3)$, the (9×6)-matrix

$$\begin{aligned}P &:= (\nabla g_1|\nabla g_2|\nabla g_3|\nabla g_4|\nabla g_5|\nabla g_6)\\ &= \begin{bmatrix}2a & 0 & 0 & b & c & 0\\ 0 & 2b & 0 & a & 0 & c\\ 0 & 0 & 2c & 0 & a & b\end{bmatrix}\end{aligned}$$

satisfies $P^T P = 2 \cdot \mathbf{1}_6$ and hence is of rank 6. More generally (but with the same proof) we see that the orthogonal group

$$\mathrm{O}(n) \;=\; \{A \in \mathbb{R}^{n\times n} \mid A^T A = \mathbf{1}\}$$

is an analytical submanifold of $\mathbb{R}^{n\times n} = \mathbb{R}^n \times \cdots \times \mathbb{R}^n$ of dimension $n(n-1)/2$. The special orthogonal group

$$\mathrm{SO}(n) \;=\; \{A \in \mathrm{O}(n) \mid \det(A) = 1\} \;=\; \{A \in \mathrm{O}(n) \mid \det(A) > 0\}$$

is an open submanifold of $\mathrm{O}(n)$; letting $\theta := \mathrm{diag}(-1, 1, \ldots, 1)$ we have the disjoint union $\mathrm{O}(n) = \mathrm{SO}(n) \cup (\theta \mathrm{SO}(n))$. For $n = 3$ a parametrization of $\mathrm{SO}(n)$ is obtained by observing that each rotation in space can be obtained by concatenating first a rotation about the third coordinate axis, then a rotation about the first coordinate axis and then again a rotation about the third coordinate axis, as follows. Choose an arbitrary Cartesian coordinate system (a_1, a_2, a_3). Given a rotation g, denote by b_1, b_2, b_3 the vectors which are obtained by applying g to a_1, a_2, a_3. Then (b_1, b_2, b_3) is again a Cartesian coordinate system, and the rotation g is uniquely determined by this system. The transition from the frame (a_1, a_2, a_3) to the frame (b_1, b_2, b_3) can be accomplished as follows.

Step 1: Rotate (a_1, a_2, a_3) by an angle γ about a_3 in such a way that the vector a_1 is moved into the plane spanned by b_1 and b_2. Denote the resulting frame by (a_1', a_2', a_3') (where $a_3' = a_3$).

Step 2: Rotate the frame (a_1', a_2', a_3') by an angle β about a_1' in such a way that the vector a_3' is transferred to the vector b_3. (This is possible because a_3' and b_3 are both perpendiculat to a_1'.) Denote the resulting frame by (a_1'', a_2'', a_3'') (where $a_1'' = a_1'$ and $a_3'' = b_3$).

Step 3: Rotate the frame (a_1'', a_2'', a_3'') by an angle α about $a_3'' = b_3$ in such a way that the vector a_1'' is transformed into the vector b_1. Then automatically also a_2'' is transformed to b_2, so that the resulting new frame (a_1''', a_2''', a_3''') coincides with (b_1, b_2, b_3).

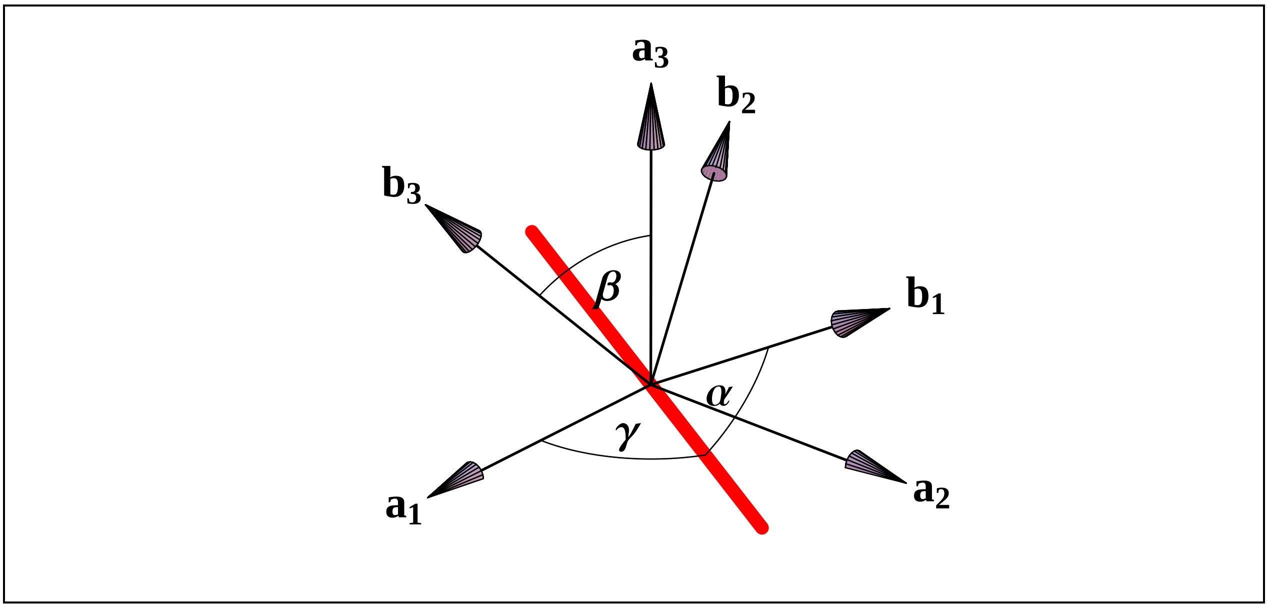

Fig. 2.12: Definition of Euler angles.

The angles α, β, γ occurring in this construction are called the **Euler angles** of g. Since the first rotation is about the third axis, the second one about the first axis and the third one again about the third axis, this yields

$$\begin{aligned} g &= \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\beta & -\sin\beta \\ 0 & \sin\beta & \cos\beta \end{bmatrix} \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \cos\alpha\cos\gamma - \sin\alpha\cos\beta\sin\gamma & -\cos\alpha\sin\gamma - \sin\alpha\cos\beta\cos\gamma & \sin\alpha\sin\beta \\ \sin\alpha\cos\gamma + \cos\alpha\cos\beta\sin\gamma & -\sin\alpha\sin\gamma + \cos\alpha\cos\beta\cos\gamma & -\cos\alpha\sin\beta \\ \sin\beta\sin\gamma & \sin\beta\cos\gamma & \cos\beta \end{bmatrix}. \end{aligned}$$

The manifold $\mathrm{SO}(3)$ can be interpreted physically as the set of all possible spatial orientations of a rigid body.

(2.8) Remark. Let $M \subseteq \mathbb{R}^n$ be a d-dimensional manifold of class C^k. It is a direct consequence of the definition of a manifold that, given an open set $U \subseteq \mathbb{R}^n$, the intersection $M \cap U$ is again a d-dimensional manifold of class C^k. Every manifold obtained in this way is called an **open submanifold** of M. (Since the sets of the form $M \cap U$ with $U \subseteq \mathbb{R}^n$ open are exactly the open subsets of M if equipped with the subspace topology inherited from $\mathbb{R}^n$, this can be restated by saying that open subsets of manifolds are themselves manifolds.)

(2.9) Remark. For $i = 1, 2$ let $M_i \subseteq \mathbb{R}^{n_i}$ be a d_i-dimensional manifold of class C^k. It is a direct consequence of the definition of a manifold that in this case the cartesian product $M_1 \times M_2 \subseteq \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} = \mathbb{R}^{n_1+n_2}$ is a $(d_1 + d_2)$-dimensional manifold of class C^k. This may be stated by saying that the direct product of two manifolds is again a manifold.