## Chapter 3

## Submanifolds

## Topics :

1. Euclidean m-Space
2. Linear Submanifolds
3. The Inverse Mapping Theorem
4. Smooth Submanifolds

Copyright © Claudiu C. Remsing, 2006.
All rights reserved.

### 3.1 Euclidean m-Space

Let $\mathbb{R}$ be the set of real numbers and let $\mathbb{R}^{m}(m \geq 1)$ denote the Cartesian product of $m$ copies of $\mathbb{R}$. Clearly, $\mathbb{R}^{1}=\mathbb{R}$. The elements of $\mathbb{R}^{m}$ are ordered $m$-tuples of real numbers. Under the usual addition and scalar multiplication, $\mathbb{R}^{m}$ is a vector space over $\mathbb{R}$.

Note : The set $\mathbb{R}^{m}$ may be equipped with various "natural" structures (e.g., group structure, vector space structure, topological structure, etc.) thus yielding various spaces (having the same underlying set) $\mathbb{R}^{m}$. We must usually decide from the context which structure is intended. We shall find it convenient to refer to the vector space $\mathbb{R}^{m}$ equipped with its canonical topology as the Cartesian $m$-space.

For $0<\ell<m$ the canonical inclusion $\mathbb{R}^{\ell} \hookrightarrow \mathbb{R}^{m}$ is defined as the map $\left(x_{1}, \ldots, x_{\ell}\right) \mapsto\left(x_{1}, \ldots, x_{\ell}, 0, \ldots, 0\right)$. Similarly, the map $\left(x_{1}, \ldots, x_{\ell}, \ldots x_{m}\right) \mapsto$ $\left(\left(x_{1}, \ldots, x_{\ell}\right),\left(x_{\ell+1}, \ldots x_{m}\right)\right)$ defines a canonical isomorphism between (vector spaces) $\mathbb{R}^{m}$ and $\mathbb{R}^{\ell} \times \mathbb{R}^{m-\ell}$. We write $\mathbb{R}^{m}=\mathbb{R}^{\ell} \times \mathbb{R}^{m-\ell}$.

The concept of Euclidean (2- or 3-dimensional) space extends straightforwardly to higher dimensions. We make the following definition.
3.1.1 Definition. The (standard) Euclidean $m$-space is the set $\mathbb{R}^{m}$ together with the Euclidean distance between points $x=\left(x_{1}, \ldots, x_{m}\right)$ and $y=\left(y_{1}, \ldots, y_{m}\right)$ given by

$$
d(x, y)=\sqrt{\left(y_{1}-x_{1}\right)^{2}+\cdots+\left(y_{m}-x_{m}\right)^{2}}
$$

The distance function $d: \mathbb{R}^{m} \times \mathbb{R}^{m} \rightarrow \mathbb{R}, \quad(x, y) \mapsto d(x, y)$ is a metric (see Exercise 7) and hence Euclidean m-space $\mathbb{R}^{m}$ is a metric space.

Note : Any metric space is a topological space and so any (standard) Euclidean space is, by definition, a Cartesian space. It is important to realize that these two structures are distinct : a Euclidean space has "more structure" than a Cartesian space; this distinction will subsequently play an important role.

We denote the open ball of center $p$ and radius $\rho>0$ by

$$
\mathcal{B}(x, \rho):=\left\{x \in \mathbb{R}^{m} \mid d(x, p)<\rho\right\} .
$$

It turns out that the open sets are exactly the (arbitrary) unions of such open balls. In the usual sense one can introduce concepts like closed sets, connected sets, convergence (of sequences), completeness, compact sets, etc. Also, one can speak of continuous mappings.

Under the usual addition and scalar multiplication, Euclidean m-space $\mathbb{R}^{m}$ is a vector space. This vector space is rather special in the sense that it has a built-in positive definite inner product (i.e., a positive definite symmetric bilinear form), the so-called dot product,

$$
x \bullet y:=x_{1} y_{1}+x_{2} y_{2}+\cdots+x_{m} y_{m}
$$

and an orthonormal basis

$$
\left\{e_{1}, e_{2}, \ldots, e_{m}\right\} \quad \text { with } \quad e_{i} \bullet e_{j}=\delta_{i j} .
$$

Note : (1) The Euclidean metric $d$ can be defined using the standard inner product on $\mathbb{R}^{m}$. We define $\|x\|$, the norm of the element (vector) $x$, by $\|x\|=\sqrt{x \bullet x}$. Then we have

$$
d(x, y)=\|x-y\| .
$$

This notation is frequently useful even when we are dealing with the Euclidean $m$ space $\mathbb{R}^{m}$ as a metric space and not using its vector space structure. In particular, $\|x\|=d(x, 0)$.
(2) An abstract concept of Euclidean space (i.e., a space satisfying the axioms of Euclidean geometry) can be introduced. It is defined as a structure $(\mathcal{E}, \vec{E}, \varphi)$, consisting of a (nonempty) set $\mathcal{E}$, an associated standard vector space (which is a real vector space equipped with an arbitrary positive definite inner product $\langle\cdot, \cdot\rangle$ ), and a structure map

$$
\varphi: \mathcal{E} \times \mathcal{E} \rightarrow \vec{E}, \quad(p, q) \mapsto \vec{p} \vec{q}
$$

such that
(AS1) $\overrightarrow{p q}+\overrightarrow{q r}=\overrightarrow{p r}$ for every $p, q, r \in \mathcal{E}$;
(AS2) For every $o \in \mathcal{E}$ and every $v \in \vec{E}$, there is a unique $p \in \mathcal{E}$ such that $\overrightarrow{o p}=v$.
Elements of $\mathcal{E}$ are called points, whereas elements of $\vec{E}$ are called vectors. ( $\overrightarrow{o p}$ is the position vector of $p$ with the initial point o.) The dimension of $\mathcal{E}$ is the dimension of (the vector space) $\vec{E}$. It turns out that
(i) if we fix an arbitrary point $o \in \mathcal{E}$, there is a one-to-one correspondence between $\mathcal{E}$ and $\vec{E}$ (the mapping $p \mapsto \overrightarrow{o p}$ is a bijection);
(ii) in addition, if we fix an arbitrary orthonormal basis $e_{1}, e_{2}, \ldots, e_{m}$ of $\vec{E}$, the (inner product) spaces $\vec{E}$ and $\mathbb{R}^{m}$ are isomorphic. (In other words, the inner product on $\vec{E}$ "is" a dot product: for $v, w \in \vec{E}$,

$$
\begin{aligned}
\langle v, w\rangle & =\left\langle v_{1} e_{1}+\cdots+v_{m} e_{m}, w_{1} e_{1}+\cdots+w_{m} e_{m}\right\rangle \\
& \left.=v_{1} w_{1}+\cdots+v_{m} w_{m} .\right)
\end{aligned}
$$

In this sense, we identify the (abstract) $m$-dimensional Euclidean space $\mathcal{E}=\mathcal{E}^{m}$ with the (concrete) standard Euclidean $m$-space $\mathbb{R}^{m}$.

Elements of Euclidean $m$-space $\mathbb{R}^{m}$, when thought of as points, will be written as $m$-tuples. When thought of as vectors, they will be written as column $m$-matrices. Euclidean 1 -space $\mathbb{R}^{1}=\mathbb{R}$ will be referred to as the real line.

Let $U \subseteq \mathbb{R}^{m}$. Let $x=\left(x_{1}, \ldots, x_{m}\right)$ denote the general (variable) point of $U$ and let $p=\left(p_{1}, \ldots, p_{m}\right)$ be a fixed but arbitrary point of $U . U$ is an open set if (and only if) for each point $x \in U$ there is an open ball $\mathcal{B}(x, \rho) \subset U$; intuitively, this means that points in $U$ are entirely surrounded by points of $U$ (or that points sufficiently close to points of $U$ still belong to $U$ ). Let $\emptyset \neq A \subseteq \mathbb{R}^{m}$. An open neighborhood of $A$ is an open set containing $A$, and a neighborhood of $A$ is any set containing an open neighborhood of $A$. In particular, a neighborhood of a set $\{p\}$ is also called a neighborhood of the point $p$.

Henceforth, throughout this chapter, $U$ will denote an open set.

## Continuity

A mapping $F: U \subseteq \mathbb{R}^{m} \rightarrow \mathbb{R}^{n}$ is continuous at $p \in U$ if (and only if) given $\varepsilon>0$, there exists a $\delta>0$ such that

$$
F(\mathcal{B}(p, \delta)) \subseteq \mathcal{B}(F(p), \varepsilon)
$$

In other words, $F$ is continuous at $p$ if (and only if) points arbitrarily close to $F(p)$ are images of points sufficiently close to $p$. We say that $F$ is continuous provided it is continuous at each $p \in U$.

Given a mapping $F: U \subseteq \mathbb{R}^{m} \rightarrow \mathbb{R}^{n}$, we can determine $n$ functions (of $m$ variables) as follows. Let $x=\left(x_{1}, \ldots, x_{m}\right) \in U$ and $F(x)=\left(y_{1}, \ldots, y_{n}\right)$. Then we can write

$$
y_{1}=f_{1}\left(x_{1}, \ldots, x_{m}\right), \quad y_{2}=f_{2}\left(x_{1}, \ldots, x_{m}\right), \quad \ldots, \quad y_{n}=f_{n}\left(x_{1}, \ldots, x_{m}\right) .
$$

The functions $f_{i}: U \rightarrow \mathbb{R}, i=1,2, \ldots, n$ are the component functions of $F$. The continuity of the mapping $F$ is equivalent to the continuity of its component functions.
$\diamond$ Exercise 123 Prove that a mapping $F: U \subseteq \mathbb{R}^{m} \rightarrow \mathbb{R}^{n}$ is continuous if and only if each component function $f_{i}: U \subseteq \mathbb{R}^{m} \rightarrow \mathbb{R}, i=1,2, \ldots, n$ is continuous.

The following results are standard (and easy to prove).
3.1.2 Proposition. Let $F, G: U \subseteq \mathbb{R}^{m} \rightarrow \mathbb{R}^{n}$ be continuous mappings and let $\lambda \in \mathbb{R}$. Then $F+G, \lambda F$, and $F \bullet G$ are each continuous. If $n=1$ and $G(x) \neq 0$ for all $x \in U$, then the quotient $\frac{F}{G}$ is also continuous.
3.1.3 Proposition. Let $F: U \subseteq \mathbb{R}^{\ell} \rightarrow \mathbb{R}^{m}$ and $G: V \subseteq \mathbb{R}^{m} \rightarrow \mathbb{R}^{n}$ be continuous mappings, where $U$ and $V$ are open sets such that $F(U) \subseteq V$. Then $G \circ F$ is a continuous mapping.
$\diamond$ Exercise 124 Show that the following mappings (or functions) are continuous.
(a) The identity mapping id: $\mathbb{R}^{m} \rightarrow \mathbb{R}^{m}, \quad x \mapsto x$.
(b) The norm function $\nu: \mathbb{R}^{m} \rightarrow \mathbb{R}, \quad x \mapsto\|x\|$.
(c) The $i^{\text {th }}$ natural projection $\mathrm{pr}_{i}: \mathbb{R}^{m} \rightarrow \mathbb{R}, \quad x \mapsto x_{i}$.

Hence derive that every polynomial function (in several variables)

$$
p_{k}: \mathbb{R}^{m} \rightarrow \mathbb{R}, \quad x=\left(x_{1}, \ldots, x_{m}\right) \mapsto \sum_{\substack{i_{1}, \ldots, i_{m}=0 \\ i_{1}+\ldots+i_{m} \leq k}}^{k} a_{i_{1} \ldots i_{m}} x_{1}^{i_{1}} \ldots x_{m}^{i_{m}}
$$

is continuous.

Note : More generally, every rational function (i.e., a quotient of two polynomial functions) is continuous. In can be shown that elementary functions like exp, log, sin, and cos are also continuous.

Mappings $L: \mathbb{R}^{m} \rightarrow \mathbb{R}^{n}$ that preserve the linear structure of the Euclidean space (i.e., linear mappings) play an important role in differentiation. Such mappings are continuous (see also Exercise 128).
$\diamond$ Exercise 125 Show that every linear mapping $L: \mathbb{R}^{m} \rightarrow \mathbb{R}^{n}$ is continuous.

In most applications it is convenient to express continuity in terms of neighborhoods instead of open balls.
$\diamond$ Exercise 126 Prove that a mapping $F: U \subseteq \mathbb{R}^{m} \rightarrow \mathbb{R}^{n}$ is continuous at $p \in$ $U$ if and only if given a neighborhood $\mathcal{N}$ of $F(p)$ in $\mathbb{R}^{n}$ there exists a neighborhood $\mathcal{M}$ of $p$ in $\mathbb{R}^{m}$ such that $F(\mathcal{M}) \subseteq \mathcal{N}$.

It is often necessary to deal with mappings (or functions) defined on arbitrary (i.e., not necessarily open) sets. To extend the previous ideas to this situation, we shall proceed as follows.

Let $F: A \subseteq \mathbb{R}^{m} \rightarrow \mathbb{R}^{n}$ be a mapping, where $A$ is an arbitrary set. We say that $F$ is continuous on $A$ provided there exists an open set $U \subseteq \mathbb{R}^{m}, A \subseteq$ $U$, and a continuous mapping $\bar{F}: U \rightarrow \mathbb{R}^{n}$ such that the restriction $\left.\bar{F}\right|_{A}=F$. In other words, $F$ is continuous (on $A$ ) if it is the restriction of a continuous mapping defined on an open set containing $A$.

NOTE : It is clear that if $F: A \subseteq \mathbb{R}^{m} \rightarrow \mathbb{R}^{n}$ is continuous, then given a neigh$\operatorname{borhood} \mathbb{N}$ of $F(p)$ in $\mathbb{R}^{n}, p \in A$, there exists a neighborhood $\mathcal{M}$ of $p$ in $\mathbb{R}^{m}$ such that $F(\mathcal{M} \cap A) \subseteq \mathcal{N}$. For this reason, it is convenient to call the set $W \cap A$ a neighborhood of $p$ in $A$.

We say that a continuous mapping $F: A \subseteq \mathbb{R}^{m} \rightarrow \mathbb{R}^{m}$ is a homeomorphism onto $F(A)$ if $F$ is one-to-one and the inverse $F^{-1}: F(A) \subseteq \mathbb{R}^{m} \rightarrow$ $\mathbb{R}^{m}$ is continuous. In this case $A$ and $F(A)$ are homeomorphic sets.
3.1.4 Example. Let $F: \mathbb{R}^{3} \rightarrow \mathbb{R}^{3}$ be given by

$$
F\left(x_{1}, x_{2}, x_{3}\right)=\left(a x_{1}, b x_{2}, c x_{3}\right)
$$

$F$ is clearly continuous, and the restriction of $F$ to the (unit) sphere

$$
\mathbb{S}^{2}=\left\{x=\left(x_{1}, x_{2}, x_{3}\right) \in \mathbb{R}^{3} \mid x_{1}^{2}+x_{2}^{2}+x_{3}^{2}=1\right\}
$$

is a continuous mapping $\widetilde{F}: \mathbb{S}^{2} \rightarrow \mathbb{R}^{3}$. Observe that $\widetilde{F}\left(\mathbb{S}^{2}\right)=E$, where $E$ is the ellipsoid

$$
E=\left\{x=\left(x_{1}, x_{2}, x_{3}\right) \in \mathbb{R}^{3} \left\lvert\, \frac{x_{1}^{2}}{a^{2}}+\frac{x_{2}^{2}}{b^{2}}+\frac{x_{3}^{2}}{c^{2}}=1\right.\right\}
$$

It is also clear that $F$ is one-to-one and that

$$
F^{-1}\left(x_{1}, x_{2}, x_{3}\right)=\left(\frac{x_{1}}{a}, \frac{x_{2}}{b}, \frac{x_{3}}{c}\right) .
$$

Thus $\widetilde{F}^{-1}=\left.F^{-1}\right|_{E}$ is continuous. Therefore, $\widetilde{F}$ is a homeomorphism of the sphere $\mathbb{S}^{2}$ onto the ellipsoid $E$.

## Differentiability

A function $f: U \subseteq \mathbb{R}^{m} \rightarrow \mathbb{R}$ is differentiable at $p \in U$ if there exists a linear functional $L_{p}: \mathbb{R}^{m} \rightarrow \mathbb{R}$ such that

$$
\lim _{x \rightarrow p} \frac{f(x)-f(p)-L_{p}(x-p)}{\|x-p\|}=0
$$

or, equivalently, if there exist a linear functional $L_{p}: \mathbb{R}^{m} \rightarrow \mathbb{R}$ and a function $R(\cdot, p)$, defined on an open neighborhood $V$ of $p \in U$, such that

$$
f(x)=f(p)+L_{p}(x-p)+\|x-p\| \cdot R(x, p), \quad x \in V
$$

and

$$
\lim _{x \rightarrow p} R(x, p)=0
$$

Then $L_{p}$ is called a derivative (or differential) of $f$ at $p$. We say that $f$ is differentiable provided it is differentiable at each $p \in U$.

NOTE : We think of a derivative $L_{p}$ as a linear approximation of $f$ near $p$. By the definition, the error involved in replacing $f(x)$ by $L_{p}(x-p)$ is negligible compared to the distance from $x$ to $p$, provided that this distance is sufficiently small.

If $L_{p}(x)=b_{1} x_{1}+\cdots+b_{m} x_{m}$ is a derivative of $f$ at $p$, then

$$
b_{i}=\frac{\partial f}{\partial x_{i}}(p):=\lim _{t \rightarrow 0} \frac{1}{t}\left(f\left(p+t e_{i}\right)-f(p)\right), \quad i=1,2, \ldots, m
$$

In particular, if $f$ is differentiable at $p$, these partial derivatives exist and the derivative $L_{p}$ is unique. We denote by $D f(p)$ (or sometimes $f^{\prime}(p)$ ) the derivative of $f$ at $p$, and write (by a slight abuse of notation)

$$
D f(p)=\frac{\partial f}{\partial x_{1}}(p)\left(x_{1}-p_{1}\right)+\frac{\partial f}{\partial x_{2}}(p)\left(x_{2}-p_{2}\right)+\cdots+\frac{\partial f}{\partial x_{n}}(p)\left(x_{m}-p_{m}\right) .
$$

$\diamond$ Exercise 127 Show that any linear functional $f: \mathbb{R}^{m} \rightarrow \mathbb{R}$ is differentiable and $D f(p)=f$ for all $p \in \mathbb{R}^{m}$.
$\diamond$ Exercise 128 Prove that any differentiable function $f: U \subseteq \mathbb{R}^{m} \rightarrow \mathbb{R}$ is continuous.

Note : Mere existence of partial derivatives is not sufficient for differentiability (of the function $f$ ). For example, the function $f: \mathbb{R}^{2} \rightarrow \mathbb{R}$ defined by

$$
f\left(x_{1}, x_{2}\right)=\frac{x_{1} x_{2}}{x_{1}^{2}+x_{2}^{2}} \quad \text { and } \quad f(0,0)=0
$$

is not continuous at $(0,0)$, yet both partial derivatives are defined there. However, if all partial derivatives $\frac{\partial f}{\partial x_{i}}, i=1,2, \ldots, m$ are defined and continuous in a neighborhood of $p \in U$, then $f$ is differentiable at $p$.

If the function $f: U \subseteq \mathbb{R}^{m} \rightarrow \mathbb{R}$ has all partial derivatives continuous (on $U$ ) we say that $f$ is continuously differentiable (or of class $C^{1}$ ) on $U$. We denote this class of functions by $C^{1}(U)$. (The class of continuous functions on $U$ is denoted by $C^{0}(U)$.)

Note: We have seen that
$f \in C^{1}(U) \Rightarrow f$ is differentiable (on $U$ ) $\Rightarrow$ all partial derivatives $\frac{\partial f}{\partial x_{i}}$ exist (on $U$ ) but the converse implications may fail. Many results actually need $f$ to be of class $C^{1}$ rather than differentiable.

If $r \geq 1$, the class $C^{r}(U)$ of functions $f: U \subseteq \mathbb{R}^{m} \rightarrow \mathbb{R}$ that are $r$ fold continuously differentiable (or $C^{r}$ functions) is specified inductively by requiring that the partial derivatives of $f$ exist and belong to $C^{r-1}(U)$. If $f$ is of class $C^{r}$ for all $r$, then we say that $f$ is of class $C^{\infty}$ or simply smooth. The class of smooth functions on $U$ is denoted by $C^{\infty}(U)$.

Note : If $f \in C^{r}(U)$, then (at any point of $U$ ) the value of the partial derivatives of order $k, 1<k \leq r$ is independent of the order of differentiation; that is, if $\left(j_{1}, \ldots, j_{k}\right)$ is a permutation of $\left(i_{1}, \ldots, i_{k}\right)$, then

$$
\frac{\partial^{k} f}{\partial x_{i_{1}} \ldots \partial x_{i_{k}}}=\frac{\partial^{k} f}{\partial x_{j_{1}} \ldots \partial x_{j_{k}}}
$$

We are now interested in extending the notion of differentiability to mappings $F: U \subseteq \mathbb{R}^{m} \rightarrow \mathbb{R}^{n}$. We say that $F$ is differentiable at $p \in U$ if (and only if) its component functions are differentiable at $p$; that is, by writing

$$
F\left(x_{1}, \ldots, x_{m}\right)=\left(f_{1}\left(x_{1}, \ldots, x_{m}\right), \ldots, f_{n}\left(x_{1}, \ldots, x_{m}\right)\right)
$$

the functions $f_{i}: U \rightarrow \mathbb{R}, i=1,2, \ldots, n$ have partial derivatives at $p \in U$. $F$ is differentiable provided it is differentiable at each $p \in U$.

The class $C^{r}\left(U, \mathbb{E}^{n}\right), 1 \leq r \leq \infty$ of $C^{r}$-mappings $F: U \subseteq \mathbb{R}^{m} \rightarrow \mathbb{R}^{n}$ is defined in the obvious way. We will be concerned primarily with smooth (i.e., of class $C^{\infty}$ ) mappings. So if $F$ is a smooth mapping, then its component functions $f_{i}, i=1,2, \ldots, n$ have continuous partial derivatives of all orders and each such derivative is independent of the order of differentiation.

Note : For the case $m=1$, we obtain the notion of (parametrized) smooth curve in Euclidean $n$-space $\mathbb{R}^{n}$. In Chapter 2, we have already seen such an object in $\mathbb{E}^{3}$. (Most of the concepts introduced in Chapter 2 can be extended to higher dimensions; in particular, the concept of tangent vector.)

Let $T_{p} \mathbb{R}^{m}$ be the tangent space to $\mathbb{R}^{m}$ at $p$; this vector space can be identified with $\mathbb{R}^{m}$ via

$$
\left.v_{1} \frac{\partial}{\partial x_{1}}\right|_{p}+\cdots+\left.v_{m} \frac{\partial}{\partial x_{m}}\right|_{p} \mapsto\left(v_{1}, \cdots, v_{m}\right)
$$

Let $\alpha: U \subseteq \mathbb{R} \rightarrow \mathbb{R}^{m}$ be a smooth (parametrized) curve with component functions $\alpha_{1}, \ldots, \alpha_{m}$. The velocity vector (or tangent vector) to $\alpha$ at $t \in U$ is the element

$$
\dot{\alpha}(t):=\left(\frac{d \alpha_{1}}{d t}(t), \cdots, \frac{d \alpha_{m}}{d t}(t)\right) \in T_{\alpha(t)} \mathbb{R}^{m}
$$

3.1.5 Example. Given a point $p \in U \subseteq \mathbb{R}^{m}$ and a (tangent) vector $v \in$ $T_{p} \mathbb{R}^{m}$, we can always find a smooth curve $\alpha:(-\varepsilon, \varepsilon) \rightarrow U$ with $\alpha(0)=$ $p$ and $\dot{\alpha}(0)=v$. Simply define $\alpha(t)=p+t v, t \in(-\varepsilon, \varepsilon)$. By writing $p=\left(p_{1}, \ldots, p_{m}\right)$ and $v=\left(v_{1}, \ldots, v_{m}\right)$, the component functions of $\alpha$ are $\alpha_{i}(t)=p_{i}+t v_{i}, i=1,2, \ldots, m$. Thus $\alpha$ is smooth, $\alpha(0)=p$ and

$$
\dot{\alpha}(0)=\left(\frac{d \alpha_{1}}{d t}(0), \cdots, \frac{d \alpha_{m}}{d t}(0)\right)=\left(v_{1}, \ldots, v_{m}\right)=v .
$$

We shall now introduce the concept of derivative (or differential) of a differentiable mapping. Let $F: U \subseteq \mathbb{R}^{m} \rightarrow \mathbb{R}^{n}$ be a differentiable mapping. To each $p \in U$ we associate a linear mapping

$$
D F(p): \mathbb{R}^{m}=T_{p} \mathbb{R}^{m} \rightarrow \mathbb{R}^{n}=T_{F(p)} \mathbb{R}^{n}
$$

which is called the derivative (or differential) of $F$ at $p$ and is defined as follows. Let $v \in T_{p} \mathbb{E}^{m}$ and let $\alpha:(-\varepsilon, \varepsilon) \rightarrow U$ be a differentiable curve such that $\alpha(0)=p$ and $\dot{\alpha}(0)=v$. By the chain rule (for functions), the curve $\beta=F \circ \alpha:(-\varepsilon, \varepsilon) \rightarrow \mathbb{E}^{n}$ is also differentiable. Then

$$
D F(p) \cdot v:=\dot{\beta}(0) .
$$

Note: The above definition of $D F(p)$ does not depend on the choice of the curve which passes through $p$ with tangent vector $v$, and $D F(p)$ is, in fact, linear. So

$$
D F(p) \cdot v=\left.\frac{d}{d t} F(\alpha(t))\right|_{t=0} \in T_{F(p)} \mathbb{R}^{n}=\mathbb{R}^{n} .
$$

The derivative $D F(p)$ is also denoted by $F_{*, p}$ and called the tangent mapping of $F$ at $p$ (see Section 2.1 for the special case when $F$ is an isometry on Euclidean 3-space $\left.\mathbb{R}^{3}\right)$.

The matrix of the linear mapping $D F(p)$ (relative to bases $\left(\left.\frac{\partial}{\partial x_{1}}\right|_{p}, \ldots,\left.\frac{\partial}{\partial x_{m}}\right|_{p}\right)$ of $T_{p} \mathbb{R}^{m}$ and $\left(\left.\frac{\partial}{\partial y_{1}}\right|_{F(p)}, \ldots,\left.\frac{\partial}{\partial y_{n}}\right|_{F(p)}\right)$ of $\left.T_{F(p)} \mathbb{R}^{n}\right)$ is the Jacobian matrix

$$
\frac{\partial F}{\partial x}(p)=\frac{\partial\left(f_{1}, \ldots, f_{n}\right)}{\partial\left(x_{1}, \ldots, x_{m}\right)}(p):=\left[\begin{array}{ccc}
\frac{\partial f_{1}}{\partial x_{1}}(p) & \cdots & \frac{\partial f_{1}}{\partial x_{m}}(p) \\
\vdots & & \vdots \\
\frac{\partial f_{n}}{\partial x_{1}}(p) & \cdots & \frac{\partial f_{n}}{\partial x_{m}}(p)
\end{array}\right] \in \mathbb{R}^{n \times m}
$$

of $F$ at $p$. When $m=n$ this is a square matrix and its determinant is then defined. This determinant is called the Jacobian of $F$ at $p$ and is denoted by $J_{F}(p)$. Thus

$$
J_{F}(p)=\left|\frac{\partial F}{\partial x}(p)\right|:=\operatorname{det} \frac{\partial F}{\partial x}(p) .
$$

$\diamond$ Exercise 129 Let $f: I \rightarrow \mathbb{R}$ and $g: J \rightarrow \mathbb{R}$ be differentiable functions, where $I$ and $J$ are open intervals such that $f(I) \subseteq J$. Show that the function $g \circ f$ is differentiable and (for $t \in I$ )

$$
(g \circ f)^{\prime}(t)=g^{\prime}(f(t)) \cdot f^{\prime}(t)
$$

The standard chain rule (for functions) extends to mappings.
3.1.6 Proposition. (The General Chain rule) Let $F: U \subseteq \mathbb{R}^{\ell} \rightarrow$ $\mathbb{R}^{m}$ and $G: V \subseteq \mathbb{R}^{m} \rightarrow \mathbb{R}^{n}$ be differentiable mappings, where $U$ and $V$ are open sets such that $F(U) \subseteq V$. Then $G \circ F$ is a differentiable mapping and (for $p \in U$ )

$$
D(G \circ F)(p)=D G(F(p)) \circ D F(p) .
$$

Proof: The fact that $G \circ F$ is differentiable is a consequence of the chain rule for functions. Now, let $v \in T_{p} \mathbb{E}^{\ell}$ be given and let us consider a (differentiable) curve $\alpha:(-\varepsilon, \varepsilon) \rightarrow U$ with $\alpha(0)=p$ and $\dot{\alpha}(0)=v$. Set $D F(p) \cdot v=w$ and observe that

$$
D G(F(p)) \cdot w=\left.\frac{d}{d t}(G \circ F \circ \alpha)\right|_{t=0}
$$

Then

$$
\begin{aligned}
D(G \circ F)(p) \cdot v & =\left.\frac{d}{d t}(G \circ F \circ \alpha)\right|_{t=0} \\
& =D G(F(p)) \cdot w \\
& =D G(F(p)) \circ D F(p) \cdot v .
\end{aligned}
$$

Note: In terms of Jacobian matrices, the general chain rule can be written

$$
\frac{\partial(G \circ F)}{\partial x}(p)=\frac{\partial G}{\partial y}(F(p)) \cdot \frac{\partial F}{\partial x}(p) .
$$

Thus if $H=G \circ F$ and $y=F(x)$, then

$$
\frac{\partial H}{\partial x}=\left[\begin{array}{ccc}
\frac{\partial g_{1}}{\partial y_{1}} & \cdots & \frac{\partial g_{1}}{\partial y_{m}} \\
\vdots & & \vdots \\
\frac{\partial g_{n}}{\partial y_{1}} & \cdots & \frac{\partial g_{n}}{\partial y_{m}}
\end{array}\right]\left[\begin{array}{ccc}
\frac{\partial f_{1}}{\partial x_{1}} & \cdots & \frac{\partial f_{1}}{\partial x_{\ell}} \\
\vdots & & \vdots \\
\frac{\partial f_{m}}{\partial x_{1}} & \cdots & \frac{\partial f_{m}}{\partial x_{\ell}}
\end{array}\right]
$$

where $\frac{\partial g_{1}}{\partial y_{1}}, \ldots, \frac{\partial g_{n}}{\partial y_{m}}$ are evaluated at $y=F(x)$ and $\frac{\partial f_{1}}{\partial x_{1}}, \cdots, \frac{\partial F_{m}}{\partial x_{\ell}}$ at $x$. Writing this out, we obtain

$$
\frac{\partial h_{i}}{\partial x_{j}}=\frac{\partial g_{i}}{\partial y_{1}} \frac{\partial y_{1}}{\partial x_{j}}+\cdots+\frac{\partial g_{i}}{\partial y_{m}} \frac{\partial y_{m}}{\partial x_{j}} \quad(i=1,2, \ldots, n ; j=1,2, \ldots, \ell)
$$

## $\diamond$ Exercise 130 Let

$$
F\left(x_{1}, x_{2}\right)=\left(x_{1}^{2}-x_{2}^{2}+x_{1} x_{2}, x_{2}^{2}-1\right) \quad \text { and } \quad G\left(y_{1}, y_{2}\right)=\left(y_{1}+y_{2}, 2 y_{1}, y_{2}^{2}\right) .
$$

(a) Show that $F$ and $G$ are differentiable, and that $G \circ F$ exists.
(b) Compute $D(G \circ F)(1,1)$
i. directly
ii. using the chain rule.
$\diamond$ Exercise 131 Show that
(a) if $\sigma: \mathbb{R}^{2} \rightarrow \mathbb{R}$ is defined by $\sigma(x, y)=x+y$, then $D \sigma(a, b)=\sigma$.
(b) if $\pi: \mathbb{R}^{2} \rightarrow \mathbb{R}$ is defined by $\pi(x, y)=x \cdot y$, then $D \pi(a, b) \cdot(x, y)=b x+a y$.

Hence deduce that if the functions $f, g: U \subseteq \mathbb{R}^{m} \rightarrow \mathbb{R}$ are differentiable at $p \in U$, then

$$
\begin{aligned}
D(f+g)(p) & =D F(p)+D g(p) \\
D(f \cdot g)(p) & =g(p) D F(p)+f(p) D G(p)
\end{aligned}
$$

If moreover $g(p) \neq 0$, then

$$
D\left(\frac{f}{g}\right)=\frac{g(p) D F(p)-f(p) D G(p)}{(g(p))^{2}}
$$

Note : The precise sense in which the derivative $D F(p)$ of the (differentiable) mapping $F$ at $p$ is a linear approximation of $F$ near $p$ is given by the following
result (in which $D F(p)$ is interpreted as a linear mapping from $\mathbb{R}^{m}$ to $\mathbb{R}^{n}$ ): If the mapping $F: U \subseteq \mathbb{R}^{m} \rightarrow \mathbb{R}^{n}$ is differentiable, then for each $p \in U$,

$$
\lim _{x \rightarrow p} \frac{F(x)-F(p)-D F(p) \cdot(x-p)}{\|x-p\|}=0 .
$$

If $A \subseteq \mathbb{R}^{m}$ is an arbitrary set, then $C^{\infty}(A)$ denotes the set of all functions $f: A \rightarrow \mathbb{R}$ such that $f=\left.\bar{f}\right|_{A}$, where $\bar{f}: U \rightarrow \mathbb{R}$ is a smooth function on some open neighborhood $U$ of $A$.

### 3.2 Linear Submanifolds

Smooth curves in Euclidean 3 -space $\mathbb{R}^{3}$ represent an important class of "geometrically interesting" subsets that are one-dimensional and can be thoroughly studied with the methods of calculus (and linear algebra). The simplest type of such geometric curve is the line, which is "straight". A two-dimensional analogue of the line is the plane, which is "flat". We shall briefly discuss these two simple cases before considering their natural higher-dimensional analogues, the linear submanifolds.

## Lines and planes in $\mathbb{R}^{3}$

Let $p \in \mathbb{R}^{3}$ and $0 \neq v \in T_{p} \mathbb{R}^{3}=\mathbb{R}^{3}$. The line through the point $p$ with direction vector $v$ is the subset

$$
L:=p+\operatorname{span}\{v\} \subset \mathbb{R}^{3}
$$

We can write

$$
L=\{p+\lambda v \mid \lambda \in \mathbb{R}\}
$$

and refer to the equation

$$
x=p+\lambda v, \quad \lambda \in \mathbb{R}
$$

as the vector equation of the line.

Note : In the vector equation of line $L$, the elements $x, p$, and $v$ are all viewed as geometric vectors, hence written as column matrices :

$$
\left[\begin{array}{l}
x_{1} \\
x_{2} \\
x_{3}
\end{array}\right]=\left[\begin{array}{l}
p_{1} \\
p_{2} \\
p_{3}
\end{array}\right]+\lambda\left[\begin{array}{l}
v_{1} \\
v_{2} \\
v_{3}
\end{array}\right], \quad \lambda \in \mathbb{R} .
$$

The vector equation is equivalent to the following set of three scalar equations:

$$
\begin{aligned}
x_{1} & =p_{1}+\lambda v_{1} \\
x_{2} & =p_{2}+\lambda v_{2} \\
x_{3} & =p_{3}+\lambda v_{3}, \quad \lambda \in \mathbb{R}
\end{aligned}
$$

called parametric equations for the line $L$. Alternatively, the line $L$ can be viewed as the image set of the linear mapping

$$
G: \mathbb{R} \rightarrow \mathbb{R}^{3}, \quad t \mapsto\left(p_{1}+t v_{1}, p_{2}+t v_{2}, p_{3}+t v_{3}\right)
$$

Now let $p \in \mathbb{R}^{3}$ and consider two linearly independent vectors $v, w \in$ $T_{p} \mathbb{R}^{3}=\mathbb{R}^{3}$. The plane through the point $p$ with direction subspace $\vec{P}=$ $\operatorname{span}\{v, w\}$ is the subset

$$
P:=p+\operatorname{span}\{v, w\} \subset \mathbb{R}^{3}
$$

Likewise, we can write

$$
P=\{p+\lambda v+\mu w \mid \lambda, \mu \in \mathbb{R}\}
$$

and refer to the equation

$$
x=p+\lambda v+\mu w, \quad \lambda, \mu \in \mathbb{R}
$$

as the vector equation of the plane. The vector equation is equivalent to the following set of three scalar equations :

$$
\begin{aligned}
& x_{1}=p_{1}+\lambda v_{1}+\mu w_{1} \\
& x_{2}=p_{2}+\lambda v_{2}+\mu w_{2} \\
& x_{3}=p_{3}+\lambda v_{3}+\mu w_{3}, \quad \lambda, \mu \in \mathbb{R}
\end{aligned}
$$

called parametric equations for the plane $P$.
Note: The fact that the vectors $v$ and $w$ are linearly independent is equivalent to the following rank condition :

$$
\operatorname{rank}\left[\begin{array}{cc}
v & w
\end{array}\right]=\operatorname{rank}\left[\begin{array}{ll}
v_{1} & w_{1} \\
v_{2} & w_{2} \\
v_{3} & w_{3}
\end{array}\right]=2 .
$$

Alternatively, the plane $P$ can be viewed as the image set of the linear mapping

$$
G^{\prime}: \mathbb{R}^{2} \rightarrow \mathbb{R}^{3}, \quad(s, t) \mapsto\left(p_{1}+s v_{1}+t w_{1}, p_{2}+s v_{2}+t w_{2}, p_{3}+s v_{3}+t w_{3}\right) .
$$

$\diamond$ Exercise 132 Show that the system of linear equations (in unknowns $\lambda$ and $\mu$ )

$$
\begin{aligned}
& \lambda v_{1}+\mu w_{1}=x_{1}-p_{1} \\
& \lambda v_{2}+\mu w_{2}=x_{2}-p_{2} \\
& \lambda v_{3}+\mu w_{3}=x_{3}-p_{3}
\end{aligned}
$$

(where $\operatorname{rank}\left[\begin{array}{ll}v & w\end{array}\right]=2$ ) is consistent if and only if

$$
\left|\begin{array}{ccc}
x_{1}-p_{1} & x_{2}-p_{2} & x_{3}-p_{3} \\
v_{1} & v_{2} & v_{3} \\
w_{1} & w_{2} & w_{3}
\end{array}\right|=0 .
$$

(Hint : A system of linear equations $A x=b$ is consistent if and only if $\operatorname{rank}\left[\begin{array}{ll}A & b\end{array}\right]=$ rank (A).)
$\diamond$ Exercise 133 Show that the condition $x-p=\lambda v+\mu w\left(\right.$ where $\operatorname{rank}\left[\begin{array}{ll}v & w\end{array}\right]=$ 2 ) is equivalent to

$$
(x-p) \bullet v \times w=0 .
$$

The plane

$$
P=p+\vec{P}=p+\operatorname{span}\{v, w\}, \quad \operatorname{rank}\left[\begin{array}{ll}
v & w
\end{array}\right]=2
$$

can be described by the scalar equation

$$
a_{1}\left(x_{1}-p_{1}\right)+a_{2}\left(x_{2}-p_{2}\right)+a_{3}\left(x_{3}-p_{3}\right)=0
$$

or by the so-called (general) Cartesian equation

$$
a_{1} x_{1}+a_{2} x_{1}+a_{3} x_{3}+c=0, \quad a_{1}^{2}+a_{2}^{2}+a_{3}^{2} \neq 0
$$

(Here

$$
\left.a_{1}=\left|\begin{array}{cc}
v_{2} & v_{3} \\
w_{2} & w_{3}
\end{array}\right|, \quad a_{2}=\left|\begin{array}{cc}
v_{3} & v_{1} \\
w_{3} & w_{1}
\end{array}\right|, \quad a_{3}=\left|\begin{array}{cc}
v_{1} & v_{2} \\
w_{1} & w_{2}
\end{array}\right| .\right)
$$

Exercise 134 Show that any equation of the form

$$
a_{1} x_{1}+a_{2} x_{2}+a_{3} x_{3}+c=0, \quad a_{1}^{2}+a_{2}^{2}+a_{3}^{2} \neq 0
$$

represents a plane $P$ in $\mathbb{R}^{3}$.

Note: The Cartesian equation for the plane $P$ can be put into the form

$$
u \bullet x+c=0
$$

where $u=v \times w$ and $c=-p \bullet v \times w$. The (nonzero) vector $u$ defines the normal direction of $P$. We can see that the line with vector direction $u=v \times w$ is orthogonal to the plane with vector subspace $\operatorname{span}\{v, w\}$.

Let $P_{1}$ and $P_{2}$ be two planes (not necessarily distinct) in $\mathbb{R}^{3}$. So

$$
P_{i}=p_{i}+\vec{P}_{i}, \quad i=1,2
$$

and it is easy to see that

$$
P_{1}=P_{2} \Longleftrightarrow p_{2}-p_{1} \in \vec{P}_{1}=\vec{P}_{2}
$$

Hence

$$
P_{1} \neq P_{2} \Longleftrightarrow\left(\vec{P}_{1} \neq \vec{P}_{2} \quad \text { or } \quad p_{2}-p_{1} \notin \vec{P}_{1}=\vec{P}_{2}\right) .
$$

It turns out that condition $p_{2}-p_{1} \notin \vec{P}_{1}=\vec{P}_{2}$ is equivalent to $P_{1} \cap P_{2}=\emptyset$; in this case, we say that the planes $P_{1}$ and $P_{2}$ are strictly parallel: $P_{1} \| P_{2}$ but $P_{1} \neq P_{2}$. Otherwise, $P_{1}$ and $P_{2}$ are two intersecting planes.

On intuitive grounds we "know" that the intersection of two distinct planes is either the empty set (when the planes are strictly parallel) or a line.
3.2.1 Proposition. The intersection of two distinct, intersectiong planes is a line.

Proof : Let $P_{1}$ and $P_{2}$ be two distinct, intersecting planes. We can describe each of these planes by a Cartesian equation of the form

$$
a_{i 1} x_{1}+a_{i 2} x_{2}+a_{i 3} x_{3}+c_{i}=0
$$

where each set of coefficients is such that $a_{i 1}^{2}+a_{i 2}^{2}+a_{i 3}^{2} \neq 0, \quad i=1,2$. The facts that the planes are distinct and are not parallel translate into the following rank condition :

$$
\operatorname{rank}\left[\begin{array}{ccc}
a_{11} & a_{12} & a_{13} \\
a_{21} & a_{22} & a_{23}
\end{array}\right]=2
$$

But this means that the system of two linear equations in three unknowns $x_{1}, x_{2}$, and $x_{3}$

$$
\begin{aligned}
& a_{11} x_{1}+a_{12} x_{2}+a_{13} x_{3}=-c_{1} \\
& a_{21} x_{1}+a_{22} x_{2}+a_{23} x_{3}=-c_{2}
\end{aligned}
$$

is consistent and, moreover, there is one basic variable $(v \neq 0)$ and one free variable $(\lambda)$. As a result, the general solution has the form

$$
x=p+\lambda v, \quad \lambda \in \mathbb{R}
$$

which represents a line.
$\diamond$ Exercise 135 Show that any line can be represented as an intersection of two (distinct) planes. (Hint : Write the parametric equations of your line in "symmetric form" :

$$
\left.\frac{x_{1}-p_{1}}{v_{1}}=\frac{x_{2}-p_{2}}{v_{2}}=\frac{x_{3}-p_{3}}{v_{3}} .\right)
$$

Note : Any line can be represented as the intersection of an arbitrary family of planes. Indeed, given a line $L$ described by the (Cartesian) equations

$$
\begin{array}{r}
(P) \quad a_{1} x_{1}+a_{2} x_{2}+a_{3} x_{3}+c=0 \\
\left(P^{\prime}\right) \quad a_{1}^{\prime} x_{1}+a_{2}^{\prime} x_{2}+a_{3}^{\prime} x_{3}+c^{\prime}=0
\end{array}
$$

where the coefficients satisfy the rank condition

$$
\operatorname{rank}\left[\begin{array}{lll}
a_{1} & a_{2} & a_{3} \\
a_{1}^{\prime} & a_{2}^{\prime} & a_{3}^{\prime}
\end{array}\right]=2
$$

(i.e., the line $L$ is represented as an intersection of two planes : $L=P \cap P^{\prime}$ ), then the family of planes

$$
\nu_{1}\left(a_{1} x_{1}+a_{2} x_{2}+a_{3} x_{3}+c\right)+\nu_{2}\left(a_{1}^{\prime} x_{1}+a_{2}^{\prime} x_{2}+a_{3}^{\prime} x_{3}+c^{\prime}\right)=0, \quad \nu_{1}, \nu_{2} \in \mathbb{R}
$$

contains all planes through the line $L$. (For $\nu_{1}=0$ we get the plane $P$. If $\nu_{1} \neq 0$, put $\nu:=\frac{\nu_{2}}{\nu_{1}}$ and we may write our family of planes - excluding the plane $P^{\prime}$ - as follows

$$
a_{1} x_{1}+a_{2} x_{2}+a_{3} x_{3}+c+\nu\left(a_{1}^{\prime} x_{1}+a_{2}^{\prime} x_{2}+a_{3}^{\prime} x_{3}+c^{\prime}\right)=0, \quad \nu \in \mathbb{R}
$$

So

$$
L=P \cap P^{\prime}=\bigcap_{\nu \in \mathbb{R}} P_{\nu}
$$

Clearly, $P=P_{0} \in\left(P_{\nu}\right)_{\nu \in \mathbb{R}}$ but $P^{\prime} \notin\left(P_{\nu}\right)_{\nu \in \mathbb{R}}$. The "exclusion" of the plane $P^{\prime}$ can be easily fixed by simply putting $P^{\prime}=P_{\infty}:=\lim _{\nu \rightarrow \infty} P_{\nu}$. Hence any subfamily, finite or infinite, of $\left(P_{\nu}\right)_{\nu \in \overline{\mathbb{R}}}, \overline{\mathbb{R}}:=\mathbb{R} \cup\{\infty\}$ has the desired property.
$\diamond$ Exercise 136 Show that the Cartesian equation of the plane through three noncolinear points $p, q, r$ can be put into the form

$$
\left|\begin{array}{llll}
x_{1} & x_{2} & x_{3} & 1 \\
p_{1} & p_{2} & p_{3} & 1 \\
q_{1} & q_{2} & q_{3} & 1 \\
r_{1} & r_{2} & r_{3} & 1
\end{array}\right|=0
$$

What do we get when the points are collinear ?
$\diamond$ Exercise 137 Prove that the lines

$$
\text { (L) } \quad x=p+\lambda v \quad \text { and } \quad\left(L^{\prime}\right) \quad x=p^{\prime}+\mu v^{\prime}
$$

lie in the same plane if and only if

$$
\left|\begin{array}{ccc}
p_{1}-p_{1}^{\prime} & p_{2}-p_{2}^{\prime} & p_{3}-p_{3}^{\prime} \\
v_{1} & v_{2} & v_{3} \\
v_{1}^{\prime} & v_{2}^{\prime} & v_{3}^{\prime}
\end{array}\right|=0
$$

## $\ell$-Planes in $\mathbb{R}^{m}$

Higher-dimensional analogues of lines and planes can be now defined without difficulty.
3.2.2 Definition. A (nonempty) subset $L \subseteq \mathbb{R}^{m}$ of the form

$$
L=p+\vec{L},
$$

where $p \in \mathbb{R}^{m}$ and $\vec{L}$ is a vector subspace of $T_{o} \mathbb{R}^{m}=\mathbb{R}^{m}$, is said to be a linear submanifold of Euclidean m-space $\mathbb{R}^{m}$.

The vector subspace $\vec{L}$ is called the direction subspace of the linear submanifold $L$. If the dimension of $\vec{L}$ (as a vector subspace) is $\ell$, then we say that $L$ is a linear submanifold of dimension $\ell$ (or, simply, a linear $\ell$ submanifold); in this case, $m-\ell$ is referred to as the codimension of $L$.
Note : A linear submanifold $L=p+\vec{L}$ is the result of "shifting" a vector subspace $\vec{L}$ by (a vector) $p$. In this vein, linear $\ell$-submanifolds are also called $\ell$-planes (or even $\ell$-flats).
3.2.3 Example. Vector subspaces of $\mathbb{R}^{m}$ are linear submanifolds. Indeed, if $p \in \vec{L}$ (in particular, if $p=o$ ), then $L=\vec{L}$.
3.2.4 Example. A linear 0 -submanifold is simply a point (in fact, a singleton). In this case, $L=p+\overrightarrow{0}=p$, hence $L=\{p\}$.
3.2.5 Example. A linear 1 -submanifold is a line (in $\mathbb{R}^{m}$ ).

A linear submanifold of dimension $m-1$ is called a hyperplane. A hyperplane has codimension 1 . What about linear submanifolds of codimension zero ? There is only one such linear submanifold, the space itself. Indeed, in this case,

$$
L=p+\operatorname{span}\left\{v_{1}, v_{2}, \ldots, v_{m}\right\}=p+\mathbb{R}^{m}=\mathbb{R}^{m} .
$$

Exercise 138 Let $L=p+\vec{L}$ be a linear submanifold and let $q \in L$. Show that

$$
L=q+\vec{L} .
$$

Hence deduce that a linear submanifold $L$ is a vector subspace if and only if $o \in L$.
$\diamond$ Exercise 139 Prove that if $p+\vec{L}=p^{\prime}+\vec{L}^{\prime}$, then $\vec{L}=\vec{L}^{\prime}$.
$\diamond$ Exercise 140 Let $\left(L_{\alpha}\right)_{\alpha \in \mathfrak{A}}$ be a family of linear submanifolds such that $\bigcap_{\substack{\alpha \in \mathscr{A} \\ \text { deduce }}} L_{\alpha} \neq \emptyset$. Show that the subset $L=\bigcap_{\alpha \in \mathscr{A}} L_{\alpha}$ is a linear submanifolds. Hence deduce that

$$
\operatorname{dim}(L)=\operatorname{dim} \bigcap_{\alpha \in \mathscr{A}} \vec{L}_{\alpha} .
$$

3.2.6 Proposition. Given two distinct points $p, q \in \mathbb{R}^{m}$, there exists a unique line $\overleftrightarrow{p q}$ containing $p$ and $q$.
Proof: (Existence) The line $p+\operatorname{span}\{q-p\}$ contains both points $p, q$. (Uniqueness) Let $L$ be a line such that $p, q \in L$. We must show that

$$
L=p+\operatorname{span}\{q-p\} .
$$

We have

$$
L=p+\vec{L}
$$

and so

$$
q \in p+\vec{L}
$$

Thus the 1 -dimensional vector subspace $\vec{L}$ contains the nonzero vector $q-p$. Hence

$$
\vec{L}=\operatorname{span}\{q-p\} .
$$

Note: The line $\overleftrightarrow{p q}$ throught the points $p$ and $q$ can be expressed as follows

$$
\overleftrightarrow{p q}=\{(1-\lambda) p+\lambda q \mid \lambda \in \mathbb{R}\} .
$$

We can now characterize linear submanifolds in terms of lines.
3.2.7 Theorem. A subset $\emptyset \neq L \subseteq \mathbb{R}^{m}$ is a linear submanifold if and only if for every two distinct points $x, y \in \mathbb{R}^{m}$, the line $\overleftrightarrow{x y}$ is contained in $L$.

Proof : Observe that this condition is equivalent to

$$
(x, y \in L, \lambda \in \mathbb{R}) \Rightarrow(1-\lambda) x+\lambda y \in L
$$

$(\Rightarrow) \quad$ Let $x, y \in L$. Then $L=x+\vec{L}$, so $y-x \in \vec{L}$ and hence

$$
\lambda(y-x) \in \vec{L}
$$

We have

$$
(1-\lambda) x+\lambda y=x+\lambda(y-x) \in x+\vec{L}=L
$$

$(\Leftarrow)$ Let $p \in L$ and denote $\vec{L}:=L-p$. Let

$$
y_{1}=x_{1}-p \in \vec{L} \quad \text { and } \quad y_{2}=x_{2}-p \in \vec{L}
$$

Then

$$
\begin{aligned}
(1-\lambda) y_{1}+\lambda y_{2} & =(1-\lambda)\left(x_{1}-p\right)+\lambda\left(x_{2}-p\right) \\
& =(1-\lambda) x_{1}+\lambda x_{2}-p \in L-p
\end{aligned}
$$

Hence

$$
\left(y_{1}, y_{2} \in \vec{L}, \lambda \in \mathbb{R}\right) \Rightarrow(1-\lambda) y_{1}+\lambda y_{2} \in \vec{L}
$$

In particular, for $y_{1}=0$, we get

$$
(y \in \vec{L}, \lambda \in \mathbb{R}) \Rightarrow \lambda y \in \vec{L}
$$

Now let $\mu \in \mathbb{R} \backslash\{0,1\}$ and let $y, y^{\prime} \in \vec{L}$. Then $y_{1}=\frac{1}{1-\mu} y, y_{2}=\frac{1}{\mu} y^{\prime} \in \vec{L}$ and thus

$$
\begin{aligned}
y+y^{\prime} & =(1-\mu) \frac{1}{1-\mu} y+\mu \frac{1}{\mu} y^{\prime} \\
& =(1-\mu) y_{1}+\mu y_{2} \in \vec{L} .
\end{aligned}
$$

Hence

$$
y, y^{\prime} \in \vec{L} \Rightarrow y+y^{\prime} \in \vec{L}
$$

It follows that $\vec{Y}$ is a vector subspace of $\mathbb{R}^{m}$. But $L=p+\vec{L}$, which proves the result.

This result can be easily generalized.

Exercise 141 Prove that a subset $\emptyset \neq L \subseteq \mathbb{R}^{m}$ is a linear submanifold if and only if

$$
\left(x_{1}, \ldots, x_{m} \in L, \lambda_{1}, \ldots, \lambda_{m} \in \mathbb{R}, \sum_{i=1}^{m} \lambda_{i}=1\right) \Rightarrow \sum_{i=1}^{m} \lambda_{i} x_{i} \in L
$$

Note : A linear combination $\sum \lambda_{i} x_{i}$ where the coefficients $\lambda_{i}$ satisfy the condition $\sum \lambda_{i}=1$ is called an affine combination. A linear submanifold can be characterized by the condition that it contains all the affine combinations of any (finite collection) of its elements; such special subsets (of some "affine space") are called affine subspaces. So linear submanifolds are just affine subspaces of $\mathbb{R}^{m}$.

In general, the union of two linear submanifolds is not a linear submanifold. Let $L_{1}$ and $L_{2}$ be two linear submanifolds of Euclidean m-space $\mathbb{R}^{m}$. Then the set $L_{1} \cup L_{2}$ does generate a linear submanifold, denoted by $L_{1} \vee L_{2}$, by taking the intersection of all linear submanifolds of $\mathbb{R}^{m}$ that contain $L_{1} \cup L_{2}$. Thus

$$
L_{1} \vee L_{2}:=\bigcap_{L_{1} \cup L_{2} \subseteq L} L \subseteq \mathbb{R}^{m}
$$

Note : $\quad L_{1} \vee L_{2}$ is the smallest linear submanifold that contains (as subsets) $L_{1}$ and $L_{2}$. It is sometimes referred to as the affine span of $L_{1} \cup L_{2}$. It turns out that for $L_{i}=p_{i}+\vec{L}_{i}, i=1,2$ one has

$$
L_{1} \vee L_{2}=p_{1}+\vec{L}_{1}+\vec{L}_{2}+\operatorname{span}\left\{p_{2}-p_{1}\right\}
$$

(Here $L_{1}+L_{2}$ denotes the sum of the vector subspaces $L_{1}$ and $L_{2}$.)
$\diamond$ Exercise 142 Given linear submanifolds $L_{i}=p_{i}+\vec{L}_{i}, i=1,2$, show that

$$
L_{1} \cap L_{2} \neq \emptyset \Longleftrightarrow \operatorname{span}\left\{p_{2}-p_{1}\right\} \subseteq \vec{L}_{1}+\vec{L}_{2}
$$

Hence deduce that if $p \in L_{1} \cap L_{2}$, then

$$
\begin{aligned}
& L_{1} \cap L_{2}=p+\vec{L}_{1} \cap \vec{L}_{2} \\
& L_{1} \vee L_{2}=p+\vec{L}_{1}+\vec{L}_{2}
\end{aligned}
$$

3.2.8 Theorem. (Dimension Theorem) Let $L_{i}=p_{i}+\vec{L}_{i}, i=1,2$ be linear submanifolds.
(a) If $L_{1} \cap L_{2} \neq \emptyset$, then

$$
\operatorname{dim}\left(L_{1} \vee L_{2}\right)=\operatorname{dim} L_{1}+\operatorname{dim} L_{2}-\operatorname{dim}\left(L_{1} \cap L_{2}\right)
$$

(b) If $L_{1} \cap L_{2}=\emptyset$, then

$$
\operatorname{dim}\left(L_{1} \vee L_{2}\right)=\operatorname{dim}\left(\vec{L}_{1}+\vec{L}_{2}\right)+1
$$

Proof: (a) We have (see Exercise 142)

$$
\begin{aligned}
\operatorname{dim}\left(L_{1} \vee L_{2}\right) & =\operatorname{dim}\left(\vec{L}_{1}+\vec{L}_{2}\right) \\
\operatorname{dim}\left(L_{1} \cap L_{2}\right) & =\operatorname{dim}\left(\vec{L}_{1} \cap \vec{L}_{2}\right) .
\end{aligned}
$$

But

$$
\operatorname{dim}\left(\vec{L}_{1}+\vec{L}_{2}\right)=\operatorname{dim} \vec{L}_{1}+\operatorname{dim} \vec{L}_{2}-\operatorname{dim}\left(\vec{L}_{1} \cap \vec{L}_{2}\right)
$$

and the first result follows.
(b) We have

$$
\begin{aligned}
\operatorname{dim}\left(L_{1} \vee L_{2}\right) & =\operatorname{dim}\left(\vec{L}_{1}+\vec{L}_{2}+\operatorname{span}\left\{p_{2}-p_{1}\right\}\right) \\
& =\operatorname{dim}\left(\vec{L}_{1}+\vec{L}_{2}\right)+1
\end{aligned}
$$

3.2.9 Example. The linear submanifold $L_{1} \vee L_{2}$ generated by the lines $L_{1}$ and $L_{2}$

- is a plane if $L_{1} \cap L_{2}=\{p\}$.
- is a plane if $L_{1} \cap L_{2}=\emptyset$ and $\vec{L}_{1}=\vec{L}_{2}$.
- has dimension 3 (i.e., is a 3-flat) if $L_{1} \cap L_{2}=\emptyset$ and $\vec{L}_{1} \neq \vec{L}_{2}$.
$\diamond$ Exercise 143 In Euclidean 4 -space $\mathbb{R}^{4}$, write (parametric) equations for the linear submanifold generated by te lines

$$
\frac{x_{1}}{2}=\frac{x_{2}-1}{1}=\frac{x_{3}+1}{-1}=\frac{x_{4}}{3}
$$

and

$$
\frac{x_{1}-1}{3}=\frac{x_{2}}{2}=\frac{x_{3}}{1}=\frac{x_{4}-2}{-1} .
$$

Consider an affine map

$$
F: \mathbb{R}^{m} \rightarrow \mathbb{R}^{n}, \quad x \mapsto A x+c .
$$

(Here $A$ is an $n \times m$ matrix and $c$ a column $n$-matrix, both with real entries.) We can see that such a map preserves affine combinations of points.
3.2.10 Proposition. Let $L=p+\vec{L}$ be a linear submanifold of $\mathbb{R}^{m}$. Then the image of $L$ under the affine map $F, x \mapsto A x+c$ is also a linear submanifold (of $\mathbb{R}^{n}$ ).

Proof: We shall show that

$$
F(L)=F(p)+A(\vec{L}) .
$$

Let $y=F(x), x \in L$; then $x-p \in \vec{L}$ and hence

$$
\begin{aligned}
y-F(p) & =F(x)-F(p) \\
& =A(x-p) \in A(\vec{L}) .
\end{aligned}
$$

Thus $F(L) \subseteq F(p)+A(\vec{L})$.
Conversely, let $y-F(p) \in A(\vec{L})$. Then

$$
y-F(p)=A(x-p)
$$

for some $x \in L$. This implies $y=F(x)$ and thus $F(L) \supseteq F(p)+A(\vec{L})$. The result now follows.
$\diamond$ Exercise 144 Given a linear submanifold $L=p+\vec{L}$ of $\mathbb{R}^{m}$ and an affine map $F: \mathbb{R}^{m} \rightarrow \mathbb{R}^{n}, x \mapsto A x+c$, show that the inverse image of any $y \in F(L)$ under $F$ is a linear submanifold. (The direction subspace of $F^{-1}(y)$ is $\operatorname{ker}(A) \subseteq \vec{L}$.)

Note : The linear submanifold $F^{-1}(y), y \in F\left(\mathbb{R}^{m}\right)=\operatorname{im}(F)$ may be referred to as the fibre of (the affine map) $F$ over (the point) $y$. All the fibres of $F$ have the same direction subspace. So the space $\mathbb{R}^{m}$ decomposes into a family of parallel submanifolds of the same dimension :

$$
\mathbb{R}^{m}=\bigcup_{y \in \operatorname{im}(F)} F^{-1}(y), \quad \operatorname{dim} F^{-1}(y)=\operatorname{dim} \operatorname{ker}(A) .
$$

Recall that, for an $n \times m$ matrix $A$, the following basic relation holds :

$$
\operatorname{dim} \operatorname{ker}(A)+\operatorname{dimim}(A)=m
$$

(the rank-nullity formula). Geometrically, this means that, for the linear map $x \mapsto$ $A x$, the nullity of $A(=\operatorname{dim} \operatorname{ker}(A))$ counts for the number of dimensions that collapse as we perform $A$ and the rank of $A(=\operatorname{dimim}(A))$ counts for the number of dimensions that survive after we perform $A$.

It follows that the dimension of any of the fibres of the affine map $F, x \mapsto A x+c$ is $m-\operatorname{rank}(A)$.

A function $f: \mathbb{R}^{m} \rightarrow \mathbb{R}$ of the form

$$
x=\left(x_{1}, x_{2}, \ldots, x_{m}\right) \mapsto a_{1} x_{1}+a_{2} x_{2}+\cdots+a_{m} x_{m}+c
$$

is called an affine functional on $\mathbb{R}^{m}$. We shall find it convenient to assume that not all the coefficients $a_{1}, \ldots, a_{m}$ are zero; so, in other words, we rule out the constant function $x \mapsto c$.

NOTE: A nonconstant affine functional is an affine map (function)

$$
f: \mathbb{R}^{m} \rightarrow \mathbb{R}, \quad x \mapsto A x+c
$$

with

$$
\operatorname{rank}(A)=\operatorname{rank}\left[\begin{array}{llll}
a_{1} & a_{2} & \cdots & a_{m}
\end{array}\right]=1
$$

Hence the fibres of $f$ are linear submanifolds (of $\mathbb{R}^{m}$ ) of dimension $m-1$ (i.e., hyperplanes).

The Cartesian equation

$$
a_{1} x_{1}+a_{2} x_{2}+\cdots+a_{m} x_{m}+c=0 \quad \text { with } \operatorname{rank}\left[\begin{array}{lll}
a_{1} & \cdots & a_{m}
\end{array}\right]=1
$$

represents the hyperplane $f^{-1}(0) \subseteq \mathbb{R}^{m}$.
$\diamond$ Exercise 145 Show that any nonconstant affine functional $f: \mathbb{R}^{m} \rightarrow \mathbb{R}$ is surjective.

Note: A system of linear equations (in unknowns $x_{1}, x_{2}, \ldots, x_{m}$ )

$$
\begin{aligned}
& a_{11} x_{1}+a_{12} x_{2}+\cdots+a_{1 m} x_{m}=b_{1} \\
& a_{21} x_{1}+a_{22} x_{2}+\cdots+a_{2 m} x_{m}=b_{2} \\
& a_{m-\ell, 1} x_{1}+a_{m-\ell, 2} x_{2}+\cdots+a_{m-\ell, m} x_{m}=b_{m-\ell}
\end{aligned}
$$

with

$$
\operatorname{rank}(A)=\operatorname{rank}\left[\begin{array}{ccc}
a_{11} & \ldots & a_{1 m} \\
\vdots & & \vdots \\
a_{m-\ell, 1} & \ldots & a_{m-\ell, m}
\end{array}\right]=m-\ell
$$

represents (geometrically) the intersection of $m-\ell$ hyperplanes in $\mathbb{R}^{m}$.
Let

$$
L=p+\vec{L}=p+\operatorname{span}\left\{v_{1}, v_{2}, \ldots, v_{\ell}\right\}
$$

be a linear submanifold of dimension $\ell$. (It is assumed, of course, that the vectors $v_{1}, v_{2}, \ldots, v_{\ell}$ are linearly independent.) Then we can write

$$
L=\left\{p+\lambda_{1} v_{1}+\cdots+\lambda_{\ell} v_{\ell} \mid \lambda_{1}, \ldots, \lambda_{\ell} \in \mathbb{R}\right\}
$$

and refer to the equation

$$
x=p+\lambda_{1} v_{1}+\cdots+\lambda_{\ell} v_{\ell}, \quad \lambda_{1}, \ldots, \lambda_{\ell} \in \mathbb{R}
$$

as the vector equation of the linear submanifold.
Equivalently, we can express (in coordinates) the linear submanifold $L$ by the following set of $m$ scalar equations

$$
\begin{aligned}
x_{1} & =p_{1}+\lambda_{1} v_{11}+\lambda_{2} v_{12}+\cdots+\lambda_{\ell} v_{1 \ell} \\
x_{2} & =p_{2}+\lambda_{1} v_{21}+\lambda_{2} v_{22}+\cdots+\lambda_{\ell} v_{2 \ell} \\
& \vdots \\
x_{m} & =p_{m}+\lambda_{1} v_{m 1}+\lambda_{2} v_{m 2}+\cdots+\lambda_{\ell} v_{m \ell}, \quad \lambda_{1}, \ldots, \lambda_{\ell} \in \mathbb{R}
\end{aligned}
$$

called parametric equations for $L$. (Here $v_{i}=\left[\begin{array}{c}v_{1 i} \\ \vdots \\ v_{m i}\end{array}\right], \quad i=1,2, \ldots, \ell$.) Alternatively, the linear submanifold $L$ can be viewed as the image set of the following affine mapping

$$
\left(t_{1}, \ldots, t_{\ell}\right) \mapsto\left(p_{1}+t_{1} v_{11}+\cdots+t_{\ell} v_{1 l}, \ldots, p_{m}+t_{1} v_{m 1}+\cdots+t_{\ell} v_{m \ell}\right)
$$

Note : Linear submanifolds are in fact solution sets for (consistent) systems of linear equations. More precisely, let $A x=b$ (where $A \in \mathbb{R}^{n \times m}$ and $b \in \mathbb{R}^{n \times 1}$ ) be a
system of $n$ linear equations in $m$ unknows $x_{1}, x_{2}, \ldots, x_{n}$. Suppose that $\operatorname{rank}(A)=$ $k$ with $0<k \leq \min \{m, n\}$. The system is consistent (i.e., it has at least one solution) if and only if the rank of the augmented matrix of the system equals the rank of the coefficient matrix (Kronecker-Capelli) :

$$
\operatorname{rank}\left[\begin{array}{ll}
A & b
\end{array}\right]=\operatorname{rank}(A)
$$

(When $b=0$, the system is said to be homogeneous and, clearly, it is consistent. A homogeneous system possesses a unique solution - the trivial solution - if and only if $\operatorname{rank}(A)=m$.) Reducing the matrix $\left[\begin{array}{ll}A & b\end{array}\right]$ to a row echelon form using Gaussian elimination and then solving for the basic variables in terms of the free variables leads to the general solution

$$
x=p+\lambda_{1} v_{1}+\lambda_{2} v_{2}+\cdots+\lambda_{m-k} v_{m-k} .
$$

As the free variables $\lambda_{i}$ range over all possible values, this general solution generates all possible solutions of the system. ( $p$ is a particular solution of the nonhomogeneous system, whereas the expression $\lambda_{1} v_{1}+\cdots+\lambda_{m-k} v_{m-k}$ is the general solution of the associated homogeneous system.) We see that the solution set $S$ of the system (assumed to be consistent) is a linear submanifold of dimension $m-k$ :

$$
S=p+\operatorname{span}\left\{v_{1}, \ldots, v_{m-k}\right\} \subset \mathbb{R}^{m}
$$

(The basic vectors form a basis of the direction subspace of $S$.) This algebraic viewpoint makes it clear that linear submanifolds can be studied, at least in principle, only by (linear) algebraic means. On the other hand, the alternative geometric viewpoint offers a broader perspective : linear submanifolds are simple, special cases of nonlinear objects/subspaces, the so-called smooth submanifolds; these are the natural higher-dimensional analogues of regular curves.

We can interpret the parametric equations for (the linear $\ell$-submanifold) $L$ as the general solution of a system of linear equations (in unknowns $x_{1}, x_{2}, \ldots, x_{m}$ ). If we write down one such system (i.e., if we eliminate the parameters $\lambda_{1}, \ldots, \lambda_{\ell}$ ) we get Cartesian equations for $L$ :

$$
\begin{aligned}
& a_{11} x_{1}+a_{12} x_{2}+\cdots+a_{1 m} x_{m}+c_{1}=0 \\
& a_{21} x_{1}+a_{22} x_{2}+\cdots+a_{2 m} x_{m}+c_{2}=0
\end{aligned}
$$

$$
a_{m-\ell, 1} x_{1}+a_{m-\ell, 2} x_{2}+\cdots+a_{m-\ell, m} x_{m}+c_{m-\ell}=0
$$

with

$$
\operatorname{rank}\left[\begin{array}{ccc}
a_{11} & \ldots & a_{1 m} \\
\vdots & & \vdots \\
a_{m-\ell, 1} & \ldots & a_{m-\ell, m}
\end{array}\right]=m-\ell .
$$

(The linear $\ell$-submanifold $L$ is represented as an intersection of $m-\ell$ distinct hyperplanes.)

We can summarize all these characterizations of a linear submanifold in the following
3.2.11 ThEOREM. Let $\emptyset \neq L$ be a subset of $\mathbb{R}^{m}$ and assume $0 \leq \ell \leq m$. The following statements are equivalent.
(i) $L$ is a linear $\ell$-submanifold of $\mathbb{R}^{m}$.
(ii) There exist linearly independent affine functions
$f_{i}: \mathbb{R}^{m} \rightarrow \mathbb{R}, \quad\left(x_{1}, \ldots, x_{m}\right) \mapsto a_{i 1} x_{1}+\cdots+a_{i m} x_{m}+c_{i} \quad(i=1,2, \ldots, m-\ell)$ (i.e., the row matrices $a_{i}=\left[\begin{array}{lll}a_{i 1} & \cdots & a_{i m}\end{array}\right], i=1,2, \ldots, m-\ell$ are linearly independent) such that

$$
L=\bigcap_{i=1}^{m-\ell} f_{i}^{-1}(0) .
$$

(iii) There exists an affine mapping

$$
F: \mathbb{R}^{m} \rightarrow \mathbb{R}^{m-\ell}, \quad x \mapsto A x+c
$$

with $\operatorname{rank}(A)=m-\ell$ such that

$$
L=F^{-1}(0) .
$$

(iv) There exist affine functions

$$
h_{i}: \mathbb{R}^{m-\ell} \rightarrow \mathbb{R}, \quad i=1,2, \ldots, m-\ell
$$

such that (possibly after a permutation of coordinates) $L$ is the graph of the mapping

$$
H=\left(h_{1}, \ldots, h_{m-\ell}\right): \mathbb{R}^{m-\ell} \rightarrow \mathbb{R}^{m-\ell} \subseteq \mathbb{R}^{m}
$$

(under the canonical isomorphism).
(v) There exists an affine mapping

$$
G: \mathbb{R}^{m-\ell} \rightarrow \mathbb{R}^{m}, \quad t=\left(t_{1}, \ldots, t_{m-\ell}\right) \mapsto B t+d
$$

with $\operatorname{rank}(B)=m-\ell$ such that $L$ is the image set of $G$.
NOTE : In (ii) we think of a linear submanifold as an intersection of hyperplanes, in (iii) as the zero-set of a certain affine mapping, in (iv) as a graph, and in (v) as the image set of a certain affine mapping (i.e., a parametrized set).

## Parallelism and orthogonality

Let $L_{i}=p_{i}+\vec{L}_{i}, i=1,2$ be linear submanifolds of $\mathbb{R}^{m}$.
3.2.12 Definition. We say that $L_{1}$ and $L_{2}$ are parallel, denoted $L_{1} \|$ $L_{2}$, provided $\vec{L}_{1} \subseteq \vec{L}_{2}$ or $\vec{L}_{2} \subseteq \vec{L}_{1}$.
$\diamond$ Exercise 146 Show that if $L_{1} \| L_{2}$, then either $L_{1} \subseteq L_{2}$ or $L_{2} \subseteq L_{1}$ or $L_{1} \cap L_{2}=\emptyset$.
$\diamond$ Exercise 147 Given two planes

$$
\begin{array}{ll}
(P) & a_{1} x_{1}+a_{2} x_{2}+a_{3} x_{3}+c=0 \\
\left(P^{\prime}\right) & a_{1}^{\prime} x_{1}+a_{2}^{\prime} x_{2}+a_{3}^{\prime} x_{3}+c^{\prime}=0
\end{array}
$$

in Euclidean 3 -space $\mathbb{R}^{3}$, show that a necessary and sufficient condition for them to be parallel is

$$
\frac{a_{1}}{a_{1}^{\prime}}=\frac{a_{2}}{a_{2}^{\prime}}=\frac{a_{3}}{a_{3}^{\prime}} .
$$

(The convention is made that if a denominator is zero, the corresponding numerator is also zero.)
$\diamond$ Exercise 148 Show that a necessary and sufficient condition for the plane

$$
a_{1} x_{1}+a_{2} x_{2}+a_{3} x_{3}+c=0
$$

and the line

$$
\begin{aligned}
& x_{1}=p_{1}+t u_{1} \\
& x_{2}=p_{2}+t u_{2} \\
& x_{3}=p_{3}+t u_{3}, \quad t \in \mathbb{R}
\end{aligned}
$$

to be parallel is

$$
a_{1} u_{1}+a_{2} u_{2}+a_{3} u_{3}=0 .
$$

3.2.13 Proposition. Let $L$ and $H$ be an arbitrary linear submanifold an a hyperplane (i.e., a linear submanifold of codimension 1), respectively. If $L \cap H=\emptyset$, then $L \| H$.

Proof : Let $L=p+\vec{L}$ and $H=q+\vec{H}$. It is clear that

$$
\operatorname{dim}(L \vee H)=m
$$

Since

$$
\operatorname{dim}(L \vee H)=\operatorname{dim}(\vec{L}+\vec{H})+1
$$

it follows that

$$
\operatorname{dim}(\vec{L}+\vec{H})=m-1=\operatorname{dim} \vec{H}
$$

We have $\vec{H} \subseteq \vec{L}+\vec{H}$ and thus

$$
\vec{H}=\vec{L}+\vec{H}
$$

Hence $\vec{L} \subseteq \vec{H}$. This shows that $L \| H$.

### 3.3 The Inverse Mapping Theorem

One of the most important results of differential calculus is the so-called inverse mapping theorem. (Another fundamental result is the existence theorem for ordinary differential equations.) In order to simplify the terminology of this and later sections we introduce first the notion of diffeomorphism (or differentiable homeomorphism) between two spaces.

Note : This concept can have no meaning unless the spaces are such that differentiability is defined, which - at the present moment - means that they must be subsets of Euclidean spaces.

Let $U \subseteq \mathbb{R}^{m}$ and $V \subseteq \mathbb{R}^{n}$ be open sets. We say that a mapping $F: U \rightarrow$ $V$ is a $C^{r}$ diffeomorphism $(1 \leq r \leq \infty)$ if $F$ is a homeomorphism and both $F$ and $F^{-1}$ are of class $C^{r}$. (When $r=1$ we simply say diffeomorphism.)

Note : A diffeomorphism is thus necessarily bijective, but a differentiable bijective mapping may not be a diffeomorphism. For example, the function $f: \mathbb{R} \rightarrow \mathbb{R}, t \mapsto t^{3}$ is a homeomorphism and $f$ is differentiable (in fact, smooth), but $f^{-1}: \mathbb{R} \rightarrow \mathbb{R}, s \mapsto$ $\sqrt[3]{s}$ is not differentiable (since it has no derivative at $s=0$ ).
$\diamond$ Exercise 149 Let $A$ be an $n \times m$ matrix and $B$ an $m \times n$ matrix. Prove that if $B A=I_{m}$ and $A B=I_{n}$, then $m=n$ and $A$ is invertible with inverse $B$. (Hint : Show that if $B A=I_{m}$, then $\operatorname{rank}(A)=\operatorname{rank}(B)=m$.)
3.3.1 Proposition. If $F: U \rightarrow V$ is a diffeomorphism (of an open subset of $\mathbb{R}^{m}$ onto an open subset of $\left.\mathbb{R}^{n}\right)$ and $p \in U$, then the derivative $D F(p)$ : $\mathbb{R}^{m}=T_{p} \mathbb{R}^{m} \rightarrow \mathbb{R}^{n}=T_{F(p)} \mathbb{R}^{n}$ is a linear isomorphism. In particular, $m=n$.

Proof : Since

$$
F^{-1} \circ F=i d_{U}=\left.i d_{\mathbb{R}^{m}}\right|_{U}
$$

( $i d_{\mathbb{R}^{m}}$ is a linear mapping), we have

$$
D\left(F^{-1} \circ F\right)(p)=i d_{\mathbb{R}^{m}}
$$

or, by the general chain rule,

$$
D F^{-1}(F(p)) \circ D F(p)=i d_{\mathbb{R}^{m}}
$$

Likewise,

$$
D F(p) \circ D F^{-1}(F(p))=i d_{\mathbb{R}^{n}}
$$

(It is safe to identify

$$
\left.T_{p} U=T_{p} \mathbb{R}^{m}=\mathbb{R}^{m} \quad \text { and } \quad T_{F(p)} V=T_{F(p)} \mathbb{R}^{n}=\mathbb{R}^{n} .\right)
$$

It follows that the linear mapping $D F(p)$ is invertible with inverse $D\left(F^{-1}\right)(F(p))$.

Note : It would not be possible to have a diffeomorphism between open subsets of Euclidean spaces of different dimensions; indeed, a famous (and deep) result of algebraic topology - Brouwer's theorem on invariance of domain - asserts that even homeomorphisms between open subsets of Euclidean spaces of different dimensions is impossible. (In fact, the result says that if $U \subseteq \mathbb{R}^{m}$ is open and $f: U \rightarrow \mathbb{R}^{n}$ is continuous and one-to-one, then $f(U)$ is open. It is then easy to derive the fact that if $U \subseteq \mathbb{R}^{m}$ and $V \subseteq \mathbb{R}^{n}$ are open subsets such that $U$ is homeomorphic to $V$, then $m=n$.)

We have seen that if the mapping $F: U \rightarrow V$ is a diffeomorphism between open subsets of $\mathbb{R}^{m}$, then the Jacobian matrix $\frac{\partial F}{\partial x}(p)$ is nonsingular (or, equivalently, the Jacobian $J_{F}(p) \neq 0$ ) for every $p \in U$. While the converse is not exactly true, it is true locally. The following fundamental result holds.
3.3.2 Theorem. (Inverse Mapping Theorem) Let $U \subseteq \mathbb{R}^{m}$ be an open set and let $F: U \rightarrow \mathbb{R}^{m}$ be of class $C^{r}(1 \leq r \leq \infty)$. Let $p \in U$ and suppose that $D F(p)$ is a linear isomorphism (i.e., the Jacobian matrix $\frac{\partial F}{\partial x}(p)$ is nonsingular). Then there exists an open neighborhood $W$ of $p$ in $U$ such that $\left.F\right|_{W}: W \rightarrow F(W)$ is a $C^{r}$ diffeomorphism. Moreover, for $y \in F(W)$ we have the following formula for the derivatives of $F^{-1}$ at $y$ :

$$
D F^{-1}(y)=(D F(x))^{-1}, \quad \text { where } y=F(x)
$$

This is a remarkable result. From a single piece of linear information at one point, it concludes to information in a whole neighborhood of that point. The proof is quite involved and will be omitted.

Note : The following two results are consequences of the inverse mapping theorem

- If $D F$ is invertible at every point of $U$, then $F$ is an open mapping (i.e., it carries $U$ and open subsets of $\mathbb{R}^{m}$ contained in $U$ into open subsets of $\left.\mathbb{R}^{m}\right)$.
- A necessary and sufficient condition for the $C^{1}$ mapping $F$ to be a diffeomorphism (from $U$ to $F(U)$ ) is that it be one-to-one and $D F$ be invertible at every point of $U$.
$\diamond$ Exercise 150 Let $F: \mathbb{R}^{2} \rightarrow \mathbb{R}^{2}$ be given by

$$
F\left(x_{1}, x_{2}\right)=\left(e^{x_{1}} \cos x_{2}, e^{x_{1}} \sin x_{2}\right) .
$$

Show that the (smooth) mapping $F$ is locally invertible, but not invertible.
Exercise 151 Show that the system

$$
\begin{aligned}
& y_{1}=x_{1}^{3} x_{2}+x_{2}^{2} \\
& y_{2}=\ln \left(x_{1}+x_{2}\right)
\end{aligned}
$$

has a unique solution $x_{1}=f\left(y_{1}, y_{2}\right), x_{2}=g\left(y_{1}, y_{2}\right)$ in a neighborhood of $(6, \ln 3)$ with $f(6, \ln 3)=1$ and $g(6, \ln 3)=2$. Find

$$
\frac{\partial f}{\partial y_{1}}, \frac{\partial f}{\partial y_{2}}, \frac{\partial g}{\partial y_{1}}, \quad \text { and } \quad \frac{\partial g}{\partial y_{2}} .
$$

There is a generalization of Theorem 3.3.2, called the constant rank theorem, which is actually equivalent to the inverse function theorem.

A $C^{1}$ mapping $F: U \subseteq \mathbb{R}^{m} \rightarrow \mathbb{R}^{n}$ has constant rank $k$ if the rank of the linear mapping $D F(x): \mathbb{R}^{m}=T_{x} \mathbb{R}^{m} \rightarrow \mathbb{R}^{n}=T_{F(x)} \mathbb{R}^{n}$ is $k$ at every point $x \in U$. Equivalently, the Jacobian matrix $\frac{\partial F}{\partial x}$ has constant rank $k$ on $U$.

Note : In linear algebra, the rank of a matrix $A \in \mathbb{R}^{n \times m}$ is defined in three equivalent ways : ( $i$ ) the dimension of the subspace of $\mathbb{R}^{m}$ spanned by the rows, (ii) the dimension of the subspace of $\mathbb{R}^{n}$ spanned by the columns, or (iii) the maximum order of any nonvanishing minor determinant. We see at once from (i) and (ii) that $\operatorname{rank}(A) \leq m, n$.

The rank of a linear mapping is defined to be the dimension of the image, and one proves that this is the rank of any matrix which represents the mapping. From this it follows that, if $P$ and $Q$ are nonsingular matrices, then $\operatorname{rank}(P A Q)=\operatorname{rank}(A)$. When $F: U \subseteq \mathbb{R}^{m} \rightarrow \mathbb{R}^{n}$ is a $C^{1}$ mapping, then the linear mapping $D F(x)$ has a rank at each $x \in U$. Because the value of the determinant is a continuous function of its entries, we see from (iii) that if $\operatorname{rank}(D F(p))=k$, then for some neighborhood $V$ of $p, \operatorname{rank}(D F(x)) \geq k$; and, if $k=\min \{m, n\}$, then $\operatorname{rank}(D F(x))=k$ on $V$. We shall refer to the rank of $D F(x)$ as the $\operatorname{rank}$ of $F$ at $x$.

If we compose $F$ with diffeomorphisms, then the facts cited and the general chain rule imply that the rank of the composition is the rank of $F$, since diffeomorphisms have nonsingular Jacobian matrices.
3.3.3 Example. Consider the composition

$$
\mathbb{R}^{k} \times \mathbb{R}^{m-k} \xrightarrow{\pi} \mathbb{R}^{k} \xrightarrow{i} \mathbb{R}^{n} \quad(1 \leq k<m, n)
$$

where

$$
\begin{aligned}
\pi\left(x_{1}, \ldots, x_{k}, y_{1}, \ldots, y_{m-k}\right) & =\left(x_{1}, \ldots, x_{k}\right) \\
i\left(x_{1}, \ldots, x_{k}\right) & =\left(x_{1}, \ldots, x_{k}, 0, \ldots, 0\right)
\end{aligned}
$$

The Jacobian matrix of $i \circ \pi$ is constantly the matrix

$$
\left[\begin{array}{cc}
I_{k} & 0 \\
0 & 0
\end{array}\right] \in \mathbb{R}^{n \times m}
$$

The rank is constantly $k$.

The constant rank theorem asserts that, in a certain precise sense, mappings of constant rank $k$ locally "look like" the above example.
3.3.4 Theorem. (Constant Rank Theorem) Let $U \subseteq \mathbb{R}^{m}$ and $V \subseteq$ $\mathbb{R}^{n}$ be open sets and let $F: U \rightarrow V$ be of class $C^{r}(1 \leq r \leq \infty)$. Let $p \in U$ and suppose that, in some neighborhood of $p, F$ has constant rank $k$. Then there are open neighborhoods $W$ of $p$ in $U$ and $Z \supseteq F(W)$ of $F(p)$ in $V$, respectively, together with $C^{r}$ diffeomorphisms

$$
G: W \rightarrow \widetilde{W} \subseteq \mathbb{R}^{m} \quad \text { and } \quad H: Z \rightarrow \widetilde{Z} \subseteq \mathbb{R}^{n}
$$

such that (on $\widetilde{W}$ )

$$
H \circ F \circ G^{-1}\left(z_{1}, \ldots, z_{m}\right)=\left(z_{1}, \ldots, z_{k}, 0, \ldots, 0\right)
$$

Note : The diffeomorphisms $G: W \rightarrow \widetilde{W}$ and $H: Z \rightarrow \widetilde{Z}$ should be thought of as changes of coordinates in these open sets. For instance, one could write

$$
\begin{aligned}
z_{1} & =g_{1}\left(x_{1}, \ldots, x_{m}\right) \\
z_{2} & =g_{2}\left(x_{1}, \ldots, x_{m}\right) \\
& \vdots \\
z_{m} & =g_{m}\left(x_{1}, \ldots, x_{m}\right)
\end{aligned}
$$

viewing $\left(z_{1}, \ldots, z_{m}\right)$ as new coordinates of the point $\left(x_{1}, \ldots, x_{m}\right)$. The new coordinates depend differentiably on the original ones and, $G$ being a diffeomorphism,
the original coordinates depend differentiably on the new ones. Thus, all of calculus, formulated in the coordinates $x_{i}$ has a completely equivalent formulation in the coordinates $z_{i}$. (The specific formulas change, but the "realities" they express do not.) According to this philosophy, the point of the constant rank theorem is that the most general mapping of constant rank can be expressed locally using the same formula as the simple Example 3.3.3, provided the coordinates in the domain and the range are suitably changed.

## The immersion and submersion theorems

There are two important special cases of Theorem 3.3.4, the immersion theorem and the submersion theorem. A $C^{r}$ mapping $F: U \subseteq \mathbb{R}^{m} \rightarrow V \subseteq \mathbb{R}^{n}$ is

- an immersion if it has constant rank $m$
- a submersion if it has constant rank $n$
on $U$.
Note : If $F$ is an immersion, then $m \leq n$. If it is a submersion, then $m \geq n$. If it is both an immersion and a submersion, then $n=m$ and $F$ is locally a diffeomorphism (such a mapping is also said to be regular).
$\diamond$ Exercise 152 Let $F: U \subseteq \mathbb{R}^{m} \rightarrow V \subseteq \mathbb{R}^{n}$ be a $C^{r}$ mapping (between open sets), and $m \leq n$. Show that $F$ is an immersion if and only if the derivative $D F(x)$ is one-to-one at every point $x \in U$.

When $m=1$, let $U$ be an open interval $J \subseteq \mathbb{R}$. In this case, the mapping $F: J \rightarrow \mathbb{R}^{n}$ is a parametrized curve in the Euclidean space $\mathbb{R}^{n}$. To verify that $F$ is an immersion it is necessary to check that the Jacobian matrix of $F$ has rank 1 (i.e., one of the derivatives, with respect to $t$, of the components of $F$ differs from zero for every $t \in J)$.
$\diamond$ Exercise 153 Verify that the following mappings are immersions.
(a) $F_{1}: \mathbb{R} \rightarrow \mathbb{R}^{3}, \quad t \mapsto(\cos t, \sin t, t)$.
(The image of $F_{1}$ is a circular helix.)
(b) $F_{2}: \mathbb{R} \rightarrow \mathbb{R}^{2}, \quad t \mapsto(\cos t, \sin t)$.
(The image of $F_{2}$ is the unit circle $\mathbb{S}^{1}$.)
(c) $F_{3}:(1, \infty) \rightarrow \mathbb{R}^{2}, \quad t \mapsto\left(\frac{1}{t} \cos (2 \pi t), \frac{1}{t} \sin (2 \pi t)\right)$.
(The image of $F_{3}$ is a curve spiraling to the origin as $t \rightarrow \infty$ and tending to the point $(1,0)$ as $t \rightarrow 1$.)
3.3.5 Corollary. (Immersion Theorem) Let $F: U \rightarrow V$ be a $C^{r}$ immersion. Then there are open neighborhoods $W$ of $p$ in $U$ and $Z \supseteq F(W)$ of $F(p)$ in $V$, respectively, together with $C^{r}$ diffeomorphisms

$$
G: W \rightarrow \widetilde{W} \subseteq \mathbb{R}^{m} \quad \text { and } \quad H: Z \rightarrow \widetilde{Z} \subseteq \mathbb{R}^{n}
$$

such that (on $\widetilde{W}$ )

$$
H \circ F \circ G^{-1}\left(y_{1}, \ldots, y_{m}\right)=\left(y_{1}, \ldots, y_{m}, 0, \ldots, 0\right)
$$

An immersion is locally - but not necessarily globally - one-to-one. For instance, the standard parametrization of the unit circle is an immersion which is clearly not one-to-one. Two more instructive examples are given below.
3.3.6 Example. Consider the mapping

$$
F: \mathbb{R} \rightarrow \mathbb{R}^{2}, \quad t \mapsto\left(2 \cos \left(t-\frac{\pi}{2}\right), \sin 2\left(t-\frac{\pi}{2}\right)\right)
$$

It is easy to check that $F$ is an immersion which is not one-to-one. The image of $F$ is a "figure eight" (a self-intersecting geometric curve) with the image point making a complete circuit starting at the origin as $t$ goes from 0 to $2 \pi$.

### 3.3.7 Example. The mapping

$$
G: \mathbb{R} \rightarrow \mathbb{R}^{2}, \quad t \mapsto F(g(t))=\left(2 \cos \left(g(t)-\frac{\pi}{2}\right), \sin 2\left(g(t)-\frac{\pi}{2}\right)\right)
$$

where $g(t)=\pi+2 \arctan t$, is again an immersion. The image is the "eight figure" as in the previous example, but with an important difference : the image point passes through the origin only once, when $t=0$; for $t \rightarrow-\infty$ and $t \rightarrow \infty$ it only approaches the origin as limit. Hence $G$ is an one-to-one immersion.

Exercise 154 Is the mapping

$$
F: \mathbb{R} \rightarrow \mathbb{R}^{2}, \quad t \mapsto\left(t^{2}, t^{3}\right)
$$

an immersion ? What about the restriction $\left.F\right|_{U}$ of $F$ to $U=\mathbb{R} \backslash\{0\}$ ? Investigate for injectivity this restriction.

Note : An immersion $F: U \subseteq \mathbb{R}^{\ell} \rightarrow \mathbb{R}^{m}$ is said to be an embedding if, in addition,

- $F$ is injective. (Observe that the induced mapping $F: U \rightarrow \mathbb{F}(U)$ is bijective.)
- $F^{-1}: F(U) \rightarrow U$ is continuous.

In particular, the mapping $F: U \rightarrow F(U)$ is bijective, continuous, and possesses a continuous inverse; hence, is is a homeomorphism. Accordingly, an embedding is an immersion which is also a homeomorphism onto its image.

### 3.3.8 Example. The mapping

$$
F: \mathbb{R} \rightarrow \mathbb{R}^{2}, \quad t \mapsto(\cos t, \sin t)
$$

is a smooth immersion (see Exercise 153). Its image set is the unit circle

$$
\mathbb{S}^{1}=\left\{x \in \mathbb{R}^{2} \mid\|x\|=1\right\}
$$

We can see that $F$ is not one-to-one. However, we can make it so by restricting $F$ to the open interval $J_{0}=(0,2 \pi)$ (or, more generally, to an interval of the form $J_{a}=(a, a+2 \pi)$ with $\left.a \in \mathbb{R}\right)$. The image of this interval under $F$ is a circle with one point left out (a punctured circle) :

$$
F\left(J_{0}\right)=\mathbb{S}^{1} \backslash\{(1,0)\}
$$

The maping

$$
F^{-1}: F\left(J_{0}\right) \rightarrow J_{0}
$$

is continuous. Consequently, $F: J_{0} \rightarrow \mathbb{R}^{2}$ is a smooth embedding.
3.3.9 Example. The mapping

$$
\widetilde{F}: \mathbb{R} \rightarrow \mathbb{R}^{2}, \quad t \mapsto\left(t^{2}-1, t^{3}-t\right)
$$

is a smooth immersion. One has

$$
\widetilde{F}(s)=\widetilde{F}(t) \Longleftrightarrow t=s \text { or } s, t \in\{-1,1\} .
$$

This makes the restriction $F:=\left.\widetilde{F}\right|_{(-\infty, 1)}$ one-to-one. But it does not make $F$ an embedding.
$\diamond$ Exercise 155 Show that the mapping

$$
F^{-1}: F((-\infty, 1)) \rightarrow(-\infty, 1)
$$

is not continuous at the point $(0,0)$.
$\diamond$ Exercise 156 Let $F: U \subseteq \mathbb{R}^{m} \rightarrow V \subseteq \mathbb{R}^{n}$ be a $C^{r}$ mapping (between open sets), and $m \geq n$. Show that $F$ is a submersion if and only if the derivative $D F(x)$ is onto at every point $x \in U$.

When $n=1$, the mapping $F=f: U \subseteq \mathbb{R}^{m} \rightarrow \mathbb{R}$ is a (differentiable) function defined on the open set $U$. To verify that $f$ is a submersion it is necessary to check that the Jacobian matrix of $f$ has rank 1 (i.e., one of the partial derivatives of $f$ differs from zero for every $t \in U$ ).
$\diamond$ Exercise 157 Verify that the following functions are submersions.
(a) $f_{1}: \mathbb{R}^{m} \rightarrow \mathbb{R}, \quad x \mapsto a_{1} x_{1}+\cdots+a_{m} x_{m}+c \quad\left(a_{1}^{2}+\cdots+a_{m}^{2} \neq 0\right)$.
(The inverse image of the origin under $f_{1}$ is a hyperplane.)
(b) $f_{2}: \mathbb{R}^{m} \backslash\{0\} \rightarrow \mathbb{R}, \quad x \mapsto x_{1}^{2}+\cdots+x_{m}^{2}-1$.
(The inverse image of the origin under $f_{2}$ is the unit sphere $\mathbb{S}^{m-1}$.)
3.3.10 Corollary. (Submersion Theorem) Let $F: U \rightarrow V$ be a $C^{r}$ submersion. Then there are open neighborhoods $W$ of $p$ in $U$ and $Z \supseteq$ $F(W)$ of $F(p)$ in $V$, respectively, together with $C^{r}$ diffeomorphisms

$$
G: W \rightarrow \widetilde{W} \subseteq \mathbb{R}^{m} \quad \text { and } \quad H: Z \rightarrow \widetilde{Z} \subseteq \mathbb{R}^{n}
$$

such that (on $\widetilde{W}$ )

$$
H \circ F \circ G^{-1}\left(y_{1}, \ldots, y_{m}\right)=\left(y_{1}, \ldots, y_{n}\right) .
$$

3.3.11 Example. Let $G L(n, R)$ denote the set (group) of all invertible (i.e., nonsingular) $n \times n$ matrices with real entries. (It can be shown that $\mathrm{GL}(n, \mathbb{R})$ may be viewed as an open subset of Euclidean space $\mathbb{R}^{n^{2}}$.) The map

$$
\operatorname{det}: \mathrm{GL}(n, \mathbb{R}) \rightarrow \mathbb{R}^{\times}, \quad A \mapsto \operatorname{det}(A)
$$

is differentiable (in fact, smooth) and its derivative is given by

$$
D \operatorname{det}(A) \cdot B=(\operatorname{det} A) \operatorname{tr}\left(A^{-1} B\right)
$$

The differentiability of det is clear from its formula in terms of matrix elements. Now

$$
\operatorname{det}\left(I_{n}+\lambda C\right)=1+\lambda \operatorname{tr} C+\cdots+\lambda^{n} \operatorname{det} C
$$

implies

$$
\left.\frac{d}{d \lambda} \operatorname{det}\left(I_{n}+\lambda C\right)\right|_{\lambda=0}=\operatorname{tr} C
$$

and hence

$$
\begin{aligned}
D \operatorname{det}(A) \cdot B & =\left.\frac{d}{d \lambda} \operatorname{det}(A+\lambda B)\right|_{\lambda=0} \\
& =\frac{d}{d \lambda}\left[(\operatorname{det} A) \operatorname{det}\left(I_{n}+\lambda A^{-1} B\right)\right]_{\lambda=0} \\
& =(\operatorname{det} A)\left(\operatorname{tr}\left(A^{-1} B\right)\right)
\end{aligned}
$$

In particular (for $A=I_{n}$ ),

$$
D \operatorname{det}\left(I_{n}\right) \cdot B=\operatorname{tr} B
$$

The map $\operatorname{tr}$ is onto, and so the function det is a (smooth) submersion.
$\diamond$ Exercise 158 Let $\operatorname{Sym}(n)$ denote the set (vector space) of all symmetric $n \times n$ matrices with real entries, and consider the mapping

$$
\Psi: \mathrm{GL}(n, \mathbb{R}) \rightarrow \operatorname{Sym}(n), \quad A \mapsto A A^{T} .
$$

Show that $\Psi$ is differentiable (in fact, smooth) and its derivative is given by

$$
D \Psi(A) \cdot B=A B^{T}+B A^{T}
$$

Hence derive that $\Psi$ is a (smooth) submersion.

## The Implicit Mapping Theorem

The following result follows easily from the Inverse Mapping Theorem.
3.3.12 Proposition. Let $U \subseteq \mathbb{R}^{k} \times \mathbb{R}^{m-k}$ be an open set and let $F$ : $U \rightarrow \mathbb{R}^{m-k}$ be of class $C^{r}(1 \leq r \leq \infty)$. Let $(p, q) \in U$ and suppose that $F(p, q)=0$ and the matrix $\frac{\partial F}{\partial y}(p, q) \in \mathbb{R}^{(m-k) \times(m-k)}$ is nonsingular. Then there exist an open neighborhood $W \subseteq \mathbb{R}^{k}$ of $p$, an open neighborhood $W^{\prime} \subseteq \mathbb{R}^{m-k}$ of $q$ and a unique $C^{r}$ mapping $\Phi: W \rightarrow W^{\prime}$ such that $\Phi(p)=q$, and for all $x \in W, \quad(x, \Phi(x)) \in U$ and

$$
F(x, \Phi(x))=0
$$

Note : This result is the so-called Implicit Mapping Theorem. It gives sufficient conditions for local solvability of a system of equations of the form

$$
\begin{aligned}
f_{1}\left(x_{1}, \ldots, x_{k}, y_{1}, \ldots, y_{m-k}\right) & =0 \\
f_{2}\left(x_{1}, \ldots, x_{k}, y_{1}, \ldots, y_{m-k}\right) & =0 \\
& \vdots \\
f_{m-k}\left(x_{1}, \ldots, x_{k}, y_{1}, \ldots, y_{m-k}\right) & =0
\end{aligned}
$$

where the functions $f_{i}$ are differentiable. (We want to solve for these $m-k$ unknown $y_{1}, \ldots, y_{m-k}$ in the $m-k$ equations in terms of $x_{1}, \ldots, x_{k}$.)

Proof : Define the mapping $\widetilde{F}: U \rightarrow \mathbb{R}^{k} \times \mathbb{R}^{m-k}=\mathbb{R}^{m}$ by

$$
\widetilde{F}(x, y):=(x, F(x, y))
$$

and observe that $\widetilde{F}$ satisfies the hypotheses of the Inverse Mapping Theorem: $\widetilde{F} \in$ $C^{r}\left(U, \mathbb{R}^{m}\right)$ and $J_{\widetilde{F}}(p, q)=\left|\frac{\partial F}{\partial y}(p, q)\right| \neq 0$. Thus there is an open neighborhood $\widetilde{W}=W_{0}^{\prime} \times W^{\prime}$ of $(p, q)$ and an open neighborhood $W \times W_{0}$ of $\widetilde{F}(p, q)=(p, 0)$ such that $\widetilde{F}: W_{0}^{\prime} \times W^{\prime} \rightarrow W \times W_{0}$ has a $C^{r}$ inverse $\widetilde{F}^{-1}: W \times W_{0} \rightarrow W_{0}^{\prime} \times W^{\prime}$; clearly, $\widetilde{F}^{-1}$ is of the form $\widetilde{F}^{-1}(x, y)=(x, H(x, y))$. Now define

$$
\Phi: W \rightarrow W^{\prime}, \quad \Phi(x):=H(x, 0)
$$

Then $\Phi \in C^{r}\left(W, \mathbb{R}^{m-k}\right)$ and

$$
(p, \Phi(p))=\left(p, H(p, 0)=\widetilde{F}^{-1}(p, 0)=(p, q)\right.
$$

which implies $\Phi(p)=q$. For $x \in W,(x, \Phi(x)) \in U$ and

$$
F(x, \Phi(x))=\left(F \circ \widetilde{F}^{-1}\right)(x, 0)=\left(\operatorname{pr}_{2} \circ \widetilde{F} \circ \widetilde{F}^{-1}\right)(x, 0)=\operatorname{pr}_{2}(x, 0)=0
$$

$\diamond$ Exercise 159 Show that (in Proposition 3.3.10) when $m-k=1$ we get

$$
\frac{\partial \Phi}{\partial x_{j}}=-\frac{\frac{\partial F}{\partial x_{j}}}{\frac{\partial F}{\partial y}} \quad(j=1,2, \ldots, k) .
$$

Note : More generally, the partial derivatives $\frac{\partial \Phi_{i}}{\partial x_{j}}$ are given by

$$
\left[\begin{array}{ccc}
\frac{\partial \Phi_{1}}{\partial x_{1}} & \cdots & \frac{\partial \Phi_{1}}{\partial x_{k}} \\
\vdots & & \vdots \\
\frac{\partial \Phi_{m-k}}{\partial x_{1}} & \cdots & \frac{\partial \Phi_{m-k}}{\partial x_{k}}
\end{array}\right]=-\left[\begin{array}{ccc}
\frac{\partial f_{1}}{\partial y_{1}} & \cdots & \frac{\partial f_{1}}{\partial y_{m-k}} \\
\vdots & & \vdots \\
\frac{\partial f_{m-k}}{\partial y_{1}} & \cdots & \frac{\partial f_{m-k}}{\partial y_{m-k}}
\end{array}\right]^{-1}\left[\begin{array}{ccc}
\frac{\partial f_{1}}{\partial x_{1}} & \cdots & \frac{\partial f_{1}}{\partial x_{k}} \\
\vdots & & \vdots \\
\frac{\partial f_{m-k}}{\partial x_{1}} & \cdots & \frac{\partial f_{m-k}}{\partial x_{k}}
\end{array}\right] .
$$

$\diamond$ Exercise 160 Show that the equations

$$
\begin{aligned}
x+y+t & =0 \\
x y t+\sin (x y t) & =0
\end{aligned}
$$

define $x$ and $y$ implicitly as functions of $t$ in an open neighborhood of the point $(t, x, y)=(-1,0,1)$. Calculate the derivatives $x^{\prime}(-1)$ and $y^{\prime}(-1)$.

### 3.4 Smooth Submanifolds

Linear submanifolds (of some Euclidean space $\mathbb{R}^{m}$ ) are a generalization of the notion of line; they are higher-dimensional geometrical objects (subsets) which can be studied rather easily because of their simple algebraic structure : linear submanifolds are "linear"! The natural "non-linear" analogues are the smooth submanifolds; smooth submanifolds are a significant generalization of the notion of smooth curve.

Note : All the results proven so far are valid for $C^{r}$ mappings (or functions). However, the class $C^{r}$ is not strong enough for some purposes. For this reason, and
since it is very convenient to know that we do not lose differentiability as a result of taking derivatives (the derivatives of a smooth mapping are also smooth). $C^{\infty}$ is the preferred differentiability class in much of (differentiable) manifold theory. Henceforth we will be concerned almost exclusively with smooth mappings (or functions).

We make the following definition.
3.4.1 Definition. A (nonempty) subset $S$ of $\mathbb{R}^{m}$ is said to be a smooth submanifold if, for every $x \in S$, there exist an open neighborhood $U$ of $x$ in $\mathbb{R}^{m}$ and a smooth diffeomorphism $\phi: U \rightarrow \widetilde{U} \subseteq \mathbb{R}^{m}$ such that

$$
\phi(S \cap U)=\widetilde{U} \cap \mathbb{R}^{\ell}
$$

where $0 \leq \ell \leq m$.
We say that $S$ is a smooth submanifold of dimension $\ell$ (or, simply, an $\ell$-submanifold). The codimension of $S$ is $m-\ell$.
Note : Roughly speaking, the condition

$$
S \cap U=\phi^{-1}\left(\widetilde{U} \cap \mathbb{R}^{\ell}\right)
$$

says that the set $S$ looks like $\mathbb{R}^{\ell}$ and is "flat" in $\mathbb{R}^{m}$. We may assume, without any loss of generality, that $\phi(x)=o$ (the origin).
3.4.2 Theorem. Let $\emptyset \neq S$ be a subset of $\mathbb{R}^{m}$ and suppose $0 \leq \ell \leq m$. The following statements are equivalent.
(i) $S$ is an $\ell$-submanifold of $\mathbb{R}^{m}$.
(ii) For every $x \in S$ there exist an open neighborhood $U$ of $x$ in $\mathbb{R}^{m}$ and smooth functions $f_{i}: U \rightarrow \mathbb{R}, \quad i=1,2, \ldots, m-\ell$ such that the linear functionals $D f_{i}(x)$ are linearly independent and

$$
S \cap U=\bigcap_{i=1}^{m-\ell} f_{i}^{-1}(0)
$$

(iii) For every $x \in S$ there exist an open neighborhood $U$ of $x$ in $\mathbb{R}^{m}$ and a smooth submersion $F: U \rightarrow \mathbb{R}^{m-\ell}$ such that

$$
S \cap U=F^{-1}(0)
$$

(iv) For every $x \in S$ there exist an open neighborhood $U$ of $x=\left(x_{1}, \ldots, x_{m}\right)$ in $\mathbb{R}^{m}$, an open neighborhood $U^{\prime}$ of $x^{\prime}=\left(x_{1}, \ldots, x_{\ell}\right)$ in $\mathbb{R}^{\ell}$ and smooth functions $h_{i}: U^{\prime} \rightarrow \mathbb{R}, \quad i=1,2, \ldots, m-\ell$ such that, possibly after a permutation of coordinates, the intersection $S \cap U$ is the graph of the mapping $H:=\left(h_{1}, \ldots, h_{m-\ell}\right): U^{\prime} \rightarrow \mathbb{R}^{m-\ell}$ (under the canonical isomorphism):

$$
S \cap U=\operatorname{graph}(H)
$$

(v) For every $x \in S$ there exist an open neighborhood $U$ of $x$ in $\mathbb{R}^{m}$, an open neighborhood $V$ of 0 in $\mathbb{R}^{\ell}$ and a smooth embedding $\Phi: V \rightarrow \mathbb{R}^{m}$ such that $\Phi(0)=x$ and

$$
S \cap U=\operatorname{im} \Phi:=\{\Phi(y) \mid y \in V\} .
$$

Note : In (ii) we think of a smooth submanifold as an intersection of hypersurfaces (i.e., codimension-1 smooth submanifolds) defined by local equations, in (iii) as the zero-set of a smooth submersion, in (iv) as a graph, and in $(v)$ as the image set of a smooth embedding (i.e., a parametrized set). All these are local descriptions. (In $(v)$ it is sufficient to assume that the smooth mapping $\Phi$ is an embedding only at the origin because if $D G(0)$ is injective, so is $D G(x)$ for $x$ close enough to o.)

Proof : We shall show that

$$
(i i i) \Rightarrow(i) \Rightarrow(v) \Rightarrow(i v) \Rightarrow(i i) \Rightarrow(i i i)
$$

$(i i i) \Rightarrow(i)$. This is just the Submersion Theorem.
$(i) \Rightarrow(v)$. We may assume (by using a translation, if necessary) that $\phi(x)=0$. Take $V=\phi(S \cap U)$ and $\Phi=\phi^{-1} \circ i$, where $i: \mathbb{R}^{\ell} \rightarrow \mathbb{R}^{m}$ is the canonical inclusion.
$(v) \Rightarrow(i v)$. After permuting indices, if necessary, we may assume that $D \Phi(0)\left(\mathbb{R}^{\ell}\right) \cap \mathbb{R}^{m-\ell}=0$. Let $\mathrm{pr}_{1}: \mathbb{R}^{m}=\mathbb{R}^{\ell} \times \mathbb{R}^{m-\ell} \rightarrow \mathbb{R}^{\ell}$ be the projection on the first factor. From $D \Phi(0)\left(\mathbb{R}^{\ell}\right) \cap \mathbb{R}^{m-\ell}=0$ we deduce that

$$
D\left(\mathrm{pr}_{1} \circ \Phi\right)(0)\left(\mathbb{R}^{\ell}\right)=\mathbb{R}^{\ell} .
$$

In other words, the mapping $\operatorname{pr}_{1} \circ \Phi$ is regular at 0 . By the Inverse Mapping Theorem, there exists an open neighborhood $V^{\prime}$ of 0 such that $\mathrm{pr}_{1} \circ \Phi$ is a (smooth) diffeomorphism between $V^{\prime}$ and $U^{\prime}=\operatorname{pr}_{1}\left(\Phi\left(V^{\prime}\right)\right) \subseteq \mathbb{R}^{\ell}$. Thus (iv) is satisfied if we take this $U^{\prime}$ and $h_{1}, \ldots, h_{m-\ell}$ equal to the $m-\ell$ last component functions of the mapping $H=\Phi \circ\left(\operatorname{pr}_{1} \circ \Phi\right)^{-1} \in C^{\infty}\left(U^{\prime}, \mathbb{R}^{m}\right)$. In fact, $H\left(U^{\prime}\right)=\Phi\left(V^{\prime}\right)$ by assumption, and so there exists an open set $U^{\prime \prime} \subseteq \mathbb{R}^{m}$ (containing $U$ ) such that

$$
\Phi\left(V^{\prime}\right)=H\left(U^{\prime}\right)=U^{\prime \prime} \cap V .
$$

Thus $U^{\prime \prime} \cap V$ is the graph of $\left(h_{1}, \ldots, h_{m-\ell}\right)=H$.
$(i v) \Rightarrow(i i)$. Just set

$$
f_{i}\left(x_{1}, \ldots, x_{m}\right)=h_{i}\left(x_{1}, \ldots, x_{\ell}\right)-x_{i+\ell}
$$

for $i=1,2, \ldots, m-\ell$.
(ii) $\Rightarrow$ (iii). The mapping $F: U \rightarrow \mathbb{R}^{m-\ell}$ with component functions $f_{1}, \ldots, f_{m-\ell}$ is a smooth submersion at $x$, and remains a submersion on an open neighborhood of $x$, since the determinant is a continuous function.

The following result follows easily from the Constant Rank Theorem.
3.4.3 Proposition. Let $U \subseteq \mathbb{R}^{m}$ and $V \subseteq \mathbb{R}^{n}$ be open sets and let $F$ : $U \rightarrow V$ be a smooth mapping of constant rank $k$. Let $q \in F(U) \subseteq V$. Then $F^{-1}(q)$ is a smooth submanifold of $U$ of dimension $m-k$.

Proof : Let $x \in F^{-1}(q)$. Choose a neighborhood of $x$ as in the Constant Rank Theorem. Without loss of generality, we can replace $W$ with $\widetilde{W}$ and $\left.F\right|_{W}$ with $H \circ F \circ G^{-1}$ on $\widetilde{W}$, all as in that theorem. That is, on $W$, we assume that

$$
F\left(x_{1}, \ldots, x_{m}\right)=\left(x_{1}, \ldots, x_{m}, 0, \ldots, 0\right) .
$$

Thus $q=\left(a_{1}, \ldots, a_{k}, 0, \ldots, 0\right)$ and $W \cap F^{-1}(q)$ is the set of all points in $W$ of the form

$$
\left(a_{1}, \ldots, a_{k}, x_{k+1}, \ldots, x_{m}\right) .
$$

The desired diffeomorphism $\phi: W \rightarrow \phi(W) \subseteq \mathbb{R}^{m}$ will be

$$
\phi\left(x_{1}, \ldots, x_{m}\right)=\left(x_{k+1}, \ldots, x_{m}, x_{1}-a_{1}, \ldots, x_{k}-a_{k}\right)
$$

## Examples of smooth submanifolds

3.4.4 EXAMPLE. 0-submanifolds of $\mathbb{R}^{m}$ are exactly sets of isolated points.
$\diamond$ Exercise 161 Show that linear submanifolds are smooth submanifolds.
3.4.5 EXAMPLE. A parametrized curve in $\mathbb{R}^{m}$ is a smooth mapping

$$
\alpha: J \rightarrow \mathbb{R}^{m}
$$

where $J \subseteq \mathbb{R}$ is an open interval. If the mapping $\alpha$ is an immersion (i.e., $\dot{\alpha}(t) \neq 0$ for all $t \in J)$, we say that the curve is regular. In this case, one can show that every $t \in J$ has a neighborhood $W$ such that $\alpha(W) \subseteq \mathbb{R}^{m}$ is a 1 -submanifold of $\mathbb{R}^{m}$.

Note : In general, the trace $\alpha(J)$ of a regular curve is not a submanifold, even if the mapping $\alpha$ is one-to-one. For instance, neither the "figure eight" (see Example 3.3.6) nor its variation, without self-intersection (see Example 3.3.7) are submanifolds of $\mathbb{R}^{2}$. Both these geometric curves are images of a smooth submanifold - the open interval $J$ - under some smooth immersion.

We have just seen that, in general, the image of a submanifold under an immersion (even a one-to-one immersion) is not a submanifold. However, the inverse image of a point (i.e., a connected 0-dimensional submanifold) under a submersion is either the empty set or a submanifold. (This is a special case of Proposition 3.4.3.)

### 3.4.6 ExAMPLE. The sphere

$$
\mathbb{S}^{m-1}:=\left\{x \in \mathbb{R}^{m} \mid\|x\|=1\right\}
$$

is a compact, $(m-1)$-submanifold of $\mathbb{R}^{m}$. ( $\mathbb{S}^{1}$ is the unit circle; $\mathbb{S}^{0}$ is equal to two points.)

To see this, write

$$
\mathbb{S}^{m-1}=\left\{x=\left(x_{1}, \ldots, x_{m}\right) \mid x_{1}^{2}+\cdots+x_{m}^{2}=1\right\}
$$

Thus the sphere $\mathbb{S}^{m-1}$ is the zero-set of the smooth function

$$
f: \mathbb{R}^{m} \rightarrow \mathbb{R}, \quad\left(x_{1}, \ldots, x_{m}\right) \mapsto x_{1}^{2}+\cdots+x_{m}^{2}-1
$$

That is, $\mathbb{S}^{m-1}=f^{-1}(0)$. Since the function $f$ is a smooth submersion, the result follows.
3.4.7 EXAMPLE. A smooth submanifold of codimension one is usually referred to as a (smooth) hypersurface. Hyperplanes and spheres are simple examples of hypersurfaces. More generally, (nonempty) subsets of the form

$$
S=\left\{x=\left(x_{1}, \ldots, x_{m}\right) \in \mathbb{R}^{m} \mid f\left(x_{1}, \ldots, x_{m}\right)=0\right\}
$$

where $f: \mathbb{R}^{m} \rightarrow \mathbb{R}$ is a smooth submersion, are hypersurfaces (of $\mathbb{R}^{m}$ ).

Another simple way of constructing smooth submanifolds is given now.
3.4.8 Proposition. Let $S_{1}$ be an $\ell_{1}$-submanifold of $\mathbb{R}^{m}$ and $S_{2}$ an $\ell_{2}$ submanifold of $\mathbb{R}^{n}$. Then $S_{1} \times S_{2}$ is an $\left(\ell_{1}+\ell_{2}\right)$-submanifold of $\mathbb{R}^{m+n}$.

Proof : Theorem 3.4.2, applied to $x \in S_{1}$, and $y \in S_{2}$, gives $n+$ $m-\left(\ell_{1}+\ell_{2}\right)$ (smooth) functions $f_{i}$ defined on an open neighborhood $U=$ $U_{1} \times U_{2} \subseteq \mathbb{R}^{m+n}$ of $(x, y)$ and satisfying condition (ii) for $S_{1} \times S_{2}$.
3.4.9 Example. The $k$-torus

$$
\mathbb{T}^{k}:=\mathbb{S}^{1} \times \cdots \times \mathbb{S}^{1} \subseteq \mathbb{R}^{2} \times \cdots \times \mathbb{R}^{2}=\mathbb{R}^{2 k}
$$

is a compact, $k$-submanifold of $\mathbb{R}^{2 k}$.
3.4.10 EXAMPLE. $\quad m$-submanifolds of $\mathbb{R}^{m}$ are exactly open subsets of $\mathbb{R}^{m}$. We shall see that the set (group) $\mathrm{GL}(n, \mathbb{R})$ of all invertible $n \times n$ matrices with real entries - the so-called (real) general linear group - is an open subset of Euclidean space $\mathbb{R}^{n^{2}}$. Hence the general linear group $G L(n, \mathbb{R})$ is a smooth submanifold (of $\mathbb{R}^{n^{2}}$ ).

Note : Any closed subgroup of $G L(n, \mathbb{R})$ turns out to be a smooth submanifold (of $\mathbb{R}^{n^{2}}$ ). This result (by no means obvious) will be proved in the chapter devoted to (abstract) Lie groups.
$\diamond$ Exercise 162 Prove that
(a) each of the following sets is a smooth submanifold of $\mathbb{R}^{2}$ (of dimension $1)$ :
i. $\left\{x \in \mathbb{R}^{2} \mid x_{2}=x_{1}^{3}\right\} ;$
ii. $\left\{x \in \mathbb{R}^{2} \mid x_{1}=x_{2}^{3}\right\}$;
iii. $\left\{x \in \mathbb{R}^{2} \mid x_{1} x_{2}=1\right\}$.
(b) none of the following sets is a smooth submanifold of $\mathbb{R}^{2}$ :
i. $\left\{x \in \mathbb{R}^{2}\left|x_{2}=\left|x_{1}\right|\right\}\right.$;
ii. $\left\{x \in \mathbb{R}^{2} \mid\left(x_{1} x_{2}-1\right)\left(x_{1}^{2}+x_{2}^{2}-2\right)=0\right\}$;
iii. $\left\{x \in \mathbb{R}^{2} \mid x_{2}=-x_{1}^{2}\right.$ for $x_{1} \leq 0 ; x_{2}=x_{1}^{2}$ for $\left.x_{1} \geq 0\right\}$.
$\diamond$ Exercise 163 Why is that

$$
\left\{x \in \mathbb{R}^{2} \mid\|x\|<1\right\} \quad \text { and } \quad\left\{x \in \mathbb{R}^{2}| | x_{1}\left|<1,\left|x_{2}\right|<1\right\}\right.
$$

are submanifolds of $\mathbb{R}^{2}$, but not

$$
\left\{x \in \mathbb{R}^{2} \mid\|x\| \leq 1\right\} ?
$$

$\diamond$ Exercise 164 Which of the following sets are smooth submanifolds (of some appropriate Euclidean space $\mathbb{R}^{m}$ ) ?
(a) $\left\{\left(t^{2}, t^{3}\right) \mid t \in \mathbb{R}\right\}$;
(b) $\left\{\left(x_{1}, x_{2}\right) \in \mathbb{R}^{2} \mid x_{1}=0\right.$ or $\left.x_{2}=0\right\}$;
(c) $\left\{\left(t, t^{2}\right) \mid t<0\right\} \cup\left\{\left(t,-t^{2}\right) \mid t>0\right\}$;
(d) $\{(\cos t, \sin t, t) \mid t \in \mathbb{R}\}$;
(e) $\left\{\left(x_{1}, x_{2}, x_{3}\right) \in \mathbb{R}^{3} \mid x_{1}^{3}+x_{2}^{3}+x_{3}^{3}-3 x_{1} x_{2} x_{3}=1\right\}$;
(f) $\left\{\left(x_{1}, x_{2}, x_{3}\right) \in \mathbb{R}^{3} \mid x_{1}^{2}+x_{2}^{2}+x_{3}^{2}=1\right.$ and $\left.x_{1}+x_{2}-x_{3}=0\right\}$.
$\diamond$ Exercise 165 Define

$$
f: \mathbb{R}^{2} \rightarrow \mathbb{R}, \quad x \mapsto x_{1}^{3}-x_{2}^{3}
$$

(a) Prove that $f$ is a surjective smooth function.
(b) Prove that $f$ is a smooth submersion at every point $x \in \mathbb{R}^{2} \backslash\{0\}$.
(c) Prove that for all $c \in \mathbb{R}$ the set

$$
\left\{x \in \mathbb{R}^{2} \mid f(x)=c\right\}
$$

is a submanifold of $\mathbb{R}^{2}$ of dimension 1 .
$\diamond$ Exercise 166 Define

$$
g: \mathbb{R}^{3} \rightarrow \mathbb{R}, \quad x \mapsto x_{1}^{2}+x_{2}^{2}-x_{2}^{2}
$$

(a) Prove that $g$ is a surjective smooth function.
(b) Prove that $g$ is a smooth submersion at every point $x \in \mathbb{R}^{3} \backslash\{0\}$.
(c) Prove that the two sheets of the cone

$$
g^{-1}(0) \backslash\{0\}=\left\{x \in \mathbb{R}^{3} \backslash\{0\} \mid x_{1}^{2}+x_{2}^{2}=x_{3}^{2}\right\}
$$

form a submanifold of $\mathbb{R}^{3}$ of dimension 2.
$\diamond$ Exercise 167 A nondegenerate quadric in $\mathbb{R}^{m}$ is a set of the form

$$
\begin{aligned}
Q & :=\left\{x \in \mathbb{R}^{m} \mid(A x) \bullet x+b \bullet x+c=0\right\} \\
& =\left\{x \in \mathbb{R}^{m} \mid x^{\top} A x+b^{\top} x+c=0\right\},
\end{aligned}
$$

where $A$ is a symmetric (i.e., $A^{\top}=A$ ) invertible $m \times m$ matrix with real entries, $b$ is a column $m$-matrix with real entries, and $c \in \mathbb{R}$. Introduce the discriminant $\Delta:=b^{\top} A^{-1} b-4 c \in \mathbb{R}$.
(a) Show that $Q$ is a smooth hypersurface (i.e., a smooth submanifold of dimension $m-1$ ) of $\mathbb{R}^{m}$.
(b) Suppose $\Delta=0$. Verify that $p:=-\frac{1}{2} A^{-1} b \in Q$ and then show that $S=Q \backslash\{p\}$ is also a smooth hypersurface of $\mathbb{R}^{m}$.

## Tangent spaces

Let $S$ be an $\ell$-submanifold of Euclidean space $\mathbb{R}^{m}$ and let $p \in S$. We want to define the (geometric) tangent space to $S$ at the point $p$; this is, locally at $p$, the "best" approximation of $S$ by a linear $\ell$-submanifold.

We shall base our definition of tangent space on the concept of (geometric) tangent vector to a curve in $\mathbb{R}^{m}$.

Let $\gamma: J \rightarrow \mathbb{R}^{m}$ be a parametrized curve in $\mathbb{R}^{m}$. (This means that $J$ is an open interval of $\mathbb{R}$ and $\gamma$ is a smooth mapping. Also, recall that the image set $\gamma(J) \subseteq \mathbb{R}^{m}$, the so-called trace of $\gamma$, is generally not a submanifold of $\mathbb{R}^{m}$.) The (geometric) tangent vector to $\gamma$ at (the point) $\gamma(t)$ is the element

$$
\dot{\gamma}(t)=\left(\frac{d \gamma_{1}}{d t}(t), \cdots, \frac{d \gamma_{m}}{d t}(t)\right) \in \mathbb{R}^{m}=T_{\gamma(t)} \mathbb{R}^{m}
$$

where $\gamma_{i}: J \rightarrow \mathbb{R}, \quad i=1,2, \ldots, m$ are the component functions of $\gamma$.
3.4.11 Definition. Let $S$ be an $\ell$-submanifold of $\mathbb{R}^{m}$ and let $p \in S$. A tangent vector $v \in \mathbb{R}^{m}=T_{p} \mathbb{R}^{m}$ is said to be a geometric tangent vector of $S$ at $p$ if there exist a parametrized curve $\gamma: J \rightarrow \mathbb{R}^{m}$ and $t_{0} \in J$ such that

$$
\begin{array}{ll}
\text { (GTV1) } & \gamma(t) \in S \text { for all } t \in J ; \\
\text { (GTV2) } & \gamma\left(t_{0}\right)=p ; \\
\text { (GTV3) } & \dot{\gamma}\left(t_{0}\right)=v .
\end{array}
$$

NOTE : We are dealing with two kinds of tangent vectors : those that are "tangent" to the whole space (i.e., the Euclidean space $\mathbb{R}^{m}$ ) and those that are tangent to a specific submanifold; the latter will be referred to as geometric tangent vectors in order to avoid ambiguity.

The set of all geometric tangent vectors of $S$ at $p$ is denoted by $T_{p} S$ and is called the tangent space to $S$ at $p$.

Note: By definition, $T_{p} S$ is a subset of (the vector space) $T_{p} \mathbb{R}^{m}=\mathbb{R}^{m}$. It turns out that it is, in fact, a vector subspace of the tangent space $T_{p} \mathbb{R}^{m}$. When regarded as a subset of (Euclidean space) $\mathbb{R}^{m}$, the tangent space $T_{p} S$ is better viewed as a linear submanifold which is tangent to (i.e., has a contact of order one with) the smooth submanifold $S$ at the point $p$. It is common to refer to $p+T_{p} S$, as the geometric tangent space of $S$ at $p$. (Obviously, the point $p$ plays an important role in this linear submanifold. By choosing the point $p$ as the origin, one obtains the identification of the geometric tangent space with the vector space $T_{p} S$.) Experience
shows that is convenient to regard the tangent space to $S$ at $p$ as a vector space (i.e., to identify the linear submanifold $p+T_{p} S$ with its direction space $T_{p} S$ ).

Let $v \in T_{p} S$ and $\lambda \in \mathbb{R}$. Then $\lambda v \in T_{p} S$. Indeed, we may assume that $\gamma(t) \in S$ for al $t \in J$ with $\gamma(0)=p$ and $\dot{\gamma}(0)=v$. Consider the parametrized curve $\gamma_{\lambda}: t \mapsto \gamma(\lambda t)$. One has, for sufficiently small $t, \gamma_{\lambda}(t) \in S$. Also

$$
\gamma_{\lambda}(0)=p \quad \text { and } \quad \dot{\gamma}_{\lambda}(0)=\lambda v
$$

Hence $\lambda v \in T_{p} S$. It is less obvious that if $v, w \in T_{p} S$, then $v+w \in T_{p} S$.
3.4.12 Theorem. Let $S$ be an $\ell$-submanifold of $\mathbb{R}^{m}$ and let $p \in S$. Assume that, locally at $p, S$ is described as in THEOREM 3.4.2. Then

$$
\begin{aligned}
T_{p} S & =\operatorname{ker}(D F(p)) \\
& =\operatorname{graph}(D H(z)) \\
& =\operatorname{im}(D \Phi(0))
\end{aligned}
$$

In particular, $T_{p} S$ is an $\ell$-dimensional vector subspace of $\mathbb{R}^{m}=T_{p} \mathbb{R}^{m}$.
Proof : Since $S$ is an $\ell$-submanifold of $\mathbb{R}^{m}$ and $p \in S$, there exists an open neighborhood $U$ of $p$ in $\mathbb{R}^{m}$ such that we can write

- $S \cap U=F^{-1}(0)$, where $F: U \rightarrow \mathbb{R}^{m-\ell}$ is a smooth submersion;
- $S \cap U=\operatorname{graph}(H)$, where $H: W \subseteq \mathbb{R}^{\ell} \rightarrow \mathbb{R}^{m-\ell}$ is a smooth mapping;
- $S \cap U=\operatorname{im}(\Phi)$, where $\Phi: V \subseteq \mathbb{R}^{\ell} \rightarrow \mathbb{R}^{m}$ is a smooth embedding.

In particular, we assume that

$$
\begin{aligned}
p & =(z, H(z)), \quad z \in W \subseteq \mathbb{R}^{\ell} \\
& =\Phi(y), \quad y \in V \subseteq \mathbb{R}^{\ell}
\end{aligned}
$$

and

$$
F(p)=0 \in \mathbb{R}^{m-\ell}
$$

Let $h \in \mathbb{R}^{\ell}$. Then there exists an $\varepsilon>0$ such that $z+t h \in W$ for all $|t|<\varepsilon$. Consequently,

$$
\gamma: t \mapsto(z+t h, H(z+t h)), \quad|t|<\varepsilon
$$

is a smooth curve in $\mathbb{R}^{m}$ such that

$$
\begin{array}{ll}
- & \gamma(t) \in V \\
- & \gamma(0)=(z, H(z))=p \\
- & \dot{\gamma}(0)=(h, D H(z) \cdot h)
\end{array}
$$

This implies

$$
\operatorname{graph}(H) \subset T_{p} S
$$

It is equally true that

$$
\operatorname{im}(D \Phi(y)) \subset T_{p} S
$$

Hence

$$
\begin{equation*}
\operatorname{graph}(H) \cup \operatorname{im}(D \Phi(y)) \subset T_{p} S \tag{*}
\end{equation*}
$$

Now let $v \in T_{p} S$ and assume $v=\dot{\gamma}\left(t_{0}\right)$. Then we have $(F \circ \gamma)(t)=0, t \in J$ and hence (by differentiation)

$$
0=D(F \circ \gamma)\left(t_{0}\right)=D F(p) \circ \dot{\gamma}\left(t_{0}\right)=D F(p) \cdot v
$$

Therefore

$$
T_{p} S \subset \operatorname{ker}(D F(p)) . \quad(* *)
$$

Since the linear mappings $h \mapsto(h, D H(z) \cdot h)$ and $D \Phi(p)$ are injective and surjective, respectively, from $(*)$ and $(*)$ it follows that

$$
\operatorname{dim} \operatorname{graph}(D H(z))=\operatorname{dimim}(D \Phi(y))=\operatorname{dim} \operatorname{ker}(D F(p))=\ell
$$

This proves the result.
$\diamond$ Exercise 168 Let $S$ be an $\ell$-submanifold of $\mathbb{R}^{m}$ and let $p \in S$. Assume that, locally at $p, S$ is described as in Theorem 3.4.2 (i). Prove that

$$
T_{p} S=\left(D \phi^{-1}(0)\right)\left(\mathbb{R}^{\ell}\right)
$$

3.4.13 EXAMPLE. Let $H=\left(h_{1}, h_{2}\right): \operatorname{dom}(H) \subseteq \mathbb{R} \rightarrow \mathbb{R}^{2}$ be a smooth mapping. The submanifold

$$
S=\left\{\left(t, h_{1}(t), h_{2}(t)\right) \in \mathbb{R}^{3} \mid t \in \operatorname{dom}(H)\right\}
$$

is the (geometric) curve in $\mathbb{R}^{3}$ given as the graph of $H$.

Then

$$
\operatorname{graph}(D H(t))=\mathbb{R}\left(1, \dot{h}_{1}(t), \dot{h}_{2}(t)\right)
$$

A parametric representation of the geometric tangent line of $S$ at $(t, H(t))$ (with $t \in R$ fixed) is

$$
x=\left(t, h_{1}(t), h_{2}(t)\right)+\lambda\left(1, \dot{h}_{1}(t), \dot{h}_{2}(t)\right), \quad \lambda \in \mathbb{R}
$$

3.4.14 EXAMPLE. Let $S \subset \mathbb{R}^{3}$ be the (geometric) helix such that

$$
\begin{gathered}
S \subset\left\{x \in \mathbb{R}^{3} \mid x_{1}^{2}+x_{2}^{2}=1\right\} \quad \text { and } \\
S \cap\left\{x \in \mathbb{R}^{3} \mid x_{3}=2 k \pi\right\}=\{(1,0,2 k \pi)\}, \quad k \in \mathbb{Z} .
\end{gathered}
$$

Then $S$ is the graph of the smooth mapping

$$
H: \mathbb{R} \rightarrow \mathbb{R}^{2}, \quad t \mapsto(\cos t, \sin t)
$$

That is,

$$
S=\{(\cos t, \sin t, t) \mid t \in \mathbb{R}\}
$$

It follows that $S$ is a smooth submanifold of $\mathbb{R}^{3}$ of dimension 1. Moreover, $S$ is a zero-set. Indeed, we have

$$
x \in S \Longleftrightarrow F(x)=\left(x_{1}-\cos x_{3}, x_{2}-\sin x_{3}\right)=0
$$

For $x=(H(t), t)$ we obtain

$$
\begin{gathered}
T_{x} S=\operatorname{graph}(D H(t)) \\
=\mathbb{R}(-\sin t, \cos t, 1) \\
\\
=\mathbb{R}\left(-x_{2}, x_{1}, 1\right) \\
D F(x)=\left[\begin{array}{ccc}
1 & 0 & \sin x_{3} \\
0 & 1 & -\cos x_{3}
\end{array}\right]
\end{gathered}
$$

The parametric representation of the geometric tangent line of $S$ at $x=$ $(H(t), t)$ is

$$
(\cos t, \sin t, t)+\lambda(-\sin t, \cos t, 1), \quad \lambda \in \mathbb{R}
$$

3.4.15 Example. The submanifold $S \subset \mathbb{R}^{3}$ is given by a smooth embed$\operatorname{ding} \Phi: \operatorname{dom}(\Phi) \subseteq \mathbb{R}^{2} \rightarrow \mathbb{R}^{3}$. That is,

$$
S=\left\{\Phi(y) \in \mathbb{R}^{3} \mid y \in \operatorname{dom}(\Phi)\right\}
$$

Then

$$
\begin{aligned}
D \Phi(y) & =\left[\begin{array}{ll}
\frac{\partial \Phi}{\partial y_{1}}(y) & \frac{\partial \Phi}{\partial y_{2}}(y)
\end{array}\right] \\
& =\left[\begin{array}{ll}
\frac{\partial \Phi_{1}}{\partial y_{1}}(y) & \frac{\partial \Phi_{1}}{\partial y_{2}}(y) \\
\frac{\partial \Phi_{2}}{\partial y_{1}}(y) & \frac{\partial \Phi_{2}}{\partial y_{2}}(y) \\
\frac{\partial \Phi_{3}}{\partial y_{1}}(y) & \frac{\partial \Phi_{3}}{\partial y_{2}}(y)
\end{array}\right] .
\end{aligned}
$$

The tangent space $T_{x} S$, with $x=\Phi(y)$, is spanned by the (tangent) vectors

$$
\frac{\partial \Phi}{\partial y_{1}}(y) \quad \text { and } \quad \frac{\partial \Phi}{\partial y_{2}}(y)
$$

Therefore, a parametric representation of the geometric tangent plane of $S$ at $\Phi(y)$ is

$$
\begin{aligned}
u_{1} & =\Phi_{1}(y)+\lambda_{1} \frac{\partial \Phi_{1}}{\partial y_{1}}(y)+\lambda_{2} \frac{\partial \Phi_{1}}{\partial y_{2}}(y) \\
u_{2} & =\Phi_{2}(y)+\lambda_{1} \frac{\partial \Phi_{2}}{\partial y_{1}}(y)+\lambda_{2} \frac{\partial \Phi_{2}}{\partial y_{2}}(y) \\
u_{3} & =\Phi_{3}(y)+\lambda_{1} \frac{\partial \Phi_{3}}{\partial y_{1}}(y)+\lambda_{2} \frac{\partial \Phi_{3}}{\partial y_{2}}(y), \quad \lambda \in \mathbb{R}^{2} .
\end{aligned}
$$

It turns out that

$$
T_{x} S=\left\{h \in \mathbb{R}^{3} \left\lvert\, h \bullet \frac{\partial \Phi}{\partial y_{1}}(y) \times \frac{\partial \Phi}{\partial y_{2}}(y)=0\right.\right\}
$$

3.4.16 Example. The submanifold $S \subset \mathbb{R}^{3}$ is the (geometric) curve in $\mathbb{R}^{3}$ given as a zero-set of a smooth submersion $F: \operatorname{dom}(F) \subseteq \mathbb{R}^{3} \rightarrow \mathbb{R}^{2}$. That is,

$$
x \in S \Longleftrightarrow F(x)=\left(f_{1}(x), f_{2}(x)\right)=0 .
$$

Then

$$
D F(x)=\left[\begin{array}{l}
\frac{\partial f_{1}}{\partial x}(x) \\
\frac{\partial f_{2}}{\partial x}(x)
\end{array}\right]
$$

and thus

$$
\operatorname{ker}(D F(x))=\left\{h \in \mathbb{R}^{3} \mid \operatorname{grad} f_{1}(x) \bullet h=\operatorname{grad} f_{2}(x) \bullet h=0\right\}
$$

The tangent space $T_{x} S$ is seen to be the line in $\mathbb{R}^{3}$ through the origin, formed by intersection of two planes

$$
\left\{h \in \mathbb{R}^{3} \mid \operatorname{grad} f_{1}(x) \bullet h=0\right\} \quad \text { and } \quad\left\{h \in \mathbb{R}^{3} \mid \operatorname{grad} f_{2}(x) \bullet h=0\right\}
$$

$\diamond$ Exercise 169 Let $S$ be the hyperboloid of two sheets

$$
S=\left\{\left(a \sinh y_{1} \cos y_{2}, b \sinh y_{1} \sin y_{2}, c \cosh y_{1}\right) \mid y=\left(y_{1}, y_{2}\right) \in \mathbb{R}^{2}\right\}, \quad a, b, c>0 .
$$

(a) Show that $S$ is a smooth submanifold of $\mathbb{R}^{3}$ of dimension 2.
(b) Determine the geometric tangent space of $S$ at an arbitrary point $p$ of $S$ in three ways, by successively considering $S$ as a zero-set, a parametrized set and a graph.
$\diamond$ Exercise 170 Let $Q \subset \mathbb{R}^{m}$ be a nondegenerate quadric given by

$$
Q=\left\{x \in \mathbb{R}^{m} \mid x^{\top} A x+b^{\top} x+c=0\right\} .
$$

Let $x \in Q \backslash\left\{-\frac{1}{2} A^{-1} b\right\}$.
(a) Prove that

$$
T_{x} Q=\left\{h \in \mathbb{R}^{m} \mid(2 A x+b) \bullet h=0\right\} .
$$

(b) Prove that

$$
\begin{aligned}
x+T_{x} Q & =\left\{h \in \mathbb{R}^{m} \mid(2 A x+b) \bullet(x-h)=0\right\} \\
& =\left\{h \in \mathbb{R}^{m} \left\lvert\,(A x) \bullet h+\frac{1}{2} b \bullet(x+h)+c=0\right.\right\} .
\end{aligned}
$$

