Geometry

Johan Martens*

Version of 12th December 2017

Abstract

Differential geometry is the study of geometry using methods of calculus and linear algebra. It has numerous applications in science and mathematics. This course is an introduction to this classical subject in the context of curves and surfaces in euclidean space. The course begins with curves in euclidean space; these have no intrinsic geometry and are fully determined by the way they bend and twist (curvature and torsion). The rest of the course will then develop the classic theory of surfaces. This will be done in the modern language of differential forms. Surfaces possess a notion of intrinsic geometry and many of the more advanced aspects of differential geometry can be demonstrated in this simpler context. One of the main aims will be to quantify the notions of curvature and shape of surfaces. The culmination of the course will be a sketch proof of the Gauss-Bonnet theorem, a profound result which relates the curvature of surfaces to their topology.

* ✉ johan.martens@ed.ac.uk

Contents

Section 1: Introduction

Differential geometry is the study of geometry using the methods of calculus and linear algebra. This was originally developed in the 18^{th} and 19^{th} centuries to describe surfaces in three dimensional Euclidean space. It has since grown into a large and rich subject whose basic objects are differentiable *manifolds*, which make the notion of a higher dimensional "curved space" precise. In this course, we will develop the classic theory of *curves* and *surfaces* in Euclidean space (1-dimensional and 2-dimensional manifolds respectively), using modern notation and concepts.

It is worth mentioning that differential geometry has many applications in science and mathematics. In particular, it is the natural language in which to formulate fundamental theories in physics: Maxwell's theory of electromagnetism, Yang-Mills theory (particle physics) and perhaps most notably Einstein's theory of General Relativity (gravity).

1.1 Aims and scope of this course

The fundamental objects we wish to study in this course are surfaces. One of our main aims is to find quantities defined on such surfaces that describe their "shape" (e.g. bowl-shaped, saddle-shaped etc.). As we will see, for suitably well behaved surfaces we may choose "coordinates" and develop a means of doing calculus on the surface itself. Eventually, this will allow us to define various quantities that do not depend on the coordinates and allow us describe the geometry of such surfaces, such as their curvature. To do this we will introduce the concept of vector fields and develop the calculus of differential forms in arbitrary "coordinate systems".

1.2 Conventions

- We understand *smooth* to mean C^∞, i.e. infinitely differentiable.
- $\mathbb{R}^n$ is the set of ordered n-tuples $x = (x^1, x^2, \ldots, x^n)$ such that $x_i \in \mathbb{R}$ for $1 \leq i \leq n$. $\mathbb{R}^n$ is often called **Cartesian space** and its elements are referred to as **points**.
- An (open and connected[1]) subset $D \subseteq \mathbb{R}^n$ is called a **domain**. We will mostly work on such domains which sometimes we refer to as **coordinate space**.
- In a variety of contexts we will consider maps $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$. Unless otherwise stated, we will assume such maps to be **smooth** in all cases. This means the component functions $f^1(x^1, \ldots, x^n), \ldots, f^m(x^1, \ldots, x^n)$ and their partial derivatives of every order exist and are continuous on D.
- We define n-dimensional **Euclidean space** $\mathbb{E}^n$ as the set $\mathbb{R}^n$ *together* with its natural vector space operations *and* the standard inner product given by the "dot product" (called the Euclidean structure). So

$$\mathbb{E}^n = \left\{ \boldsymbol{x} = \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix} \quad \text{such that } x^i \in \mathbb{R} \text{ for all } 1 \leq i \leq n \right\}$$

[1] These are terms from topology which we will not define in this course.

and

$$\boldsymbol{x} \cdot \boldsymbol{y} = \sum_{k=1}^{n} x^k y^k \qquad\qquad |\boldsymbol{x}| = \sqrt{\boldsymbol{x} \cdot \boldsymbol{x}}\ .$$

Note the use of the usual bold letters – underlined when hand-written – for vectors in Euclidean space (written as columns), to distinguish them from mere points in $\mathbb{R}^n$ (written as rows). We will preserve this distinction throughout.

1.3 Example

An important example is the surface of a sphere in Euclidean space, defined by

$$S^2 = \{\boldsymbol{x} \in \mathbb{E}^3 \mid |\boldsymbol{x}| = 1\} \subset \mathbb{E}^3.$$

Unlike, $\mathbb{E}^3$, the surface S^2 is not a vector space. This raises the question: how do you define vectors on S^2? There is a very elegant, geometrical answer to this.

At each point on S^2, one can construct a whole plane which is tangent to the sphere, called the **tangent plane**. This plane $\mathbb{R}^2$ is the two dimensional vector space of lines tangent to the sphere at the given point, called tangent vectors. However, as is clear from this picture, each point on the sphere defines a different tangent plane. This leads to the notion of a vector field which is a rule for smoothly assigning a tangent vector to each point on S^2. This simple geometrical picture provides the basis for defining vectors on general curved surfaces and indeed on more general curved spaces known as manifolds.

Notice that the above description of the surface S^2 is valid everywhere on the surface: this is referred to as a **global** description. However, usually it is not possible to have such a simple description of a surface. In general, the way to deal with this is to parameterise a number of "patches" of the surface using coordinates, in such a way that the patches cover the whole surface. In this way one can always identify patches of the surface with domains $D \subset \mathbb{R}^2$. In our example, we may parameterise with the standard spherical polar coordinates map

$$(\theta, \phi) \mapsto \boldsymbol{x}(\theta, \phi) = \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix}$$

so that most of the surface of the sphere becomes identified with the rectangular domain

$$D = \{(\theta, \phi) \mid 0 < \theta < \pi, 0 < \phi < 2\pi\} \subset \mathbb{R}^2\ .$$

This defines a (smoth) map $\boldsymbol{x} : D \to S^2 \subset \mathbb{E}^3$, which gives a useful description of the surface, in this case valid everywhere except for a line of longitude (including the two poles). Such a description is known as **local**, referring to the fact it does not apply everywhere on the surface.

There are a few important points to observe about this. The $\mathbb{R}^2$ in which this domain D is contained is not to be thought of as the "vector space $\mathbb{R}^2$". The vector space operations of addition and scalar multiplication have no meaning in D. Furthermore, the usual Euclidean distance between two points in this $\mathbb{R}^2$ is also not the correct notion of distance on a sphere. This is why in general we need to carefully distinguish between the coordinate space $D \subseteq \mathbb{R}^n$ and Euclidean space $\mathbb{E}^n$. Due to this lack of a vector space structure, it is not so clear how one defines vectors on such domains. The general answer to this, as we will see, is inspired by the above construction of the tangent plane on a sphere, however needs no reference to structures "outside" the domain/surface.

Another important point is that the standard polar coordinates are not the only way of parameterising a sphere. For example, we could choose polar coordinates based at a different "North pole", or even choose different types of coordinates altogether. As mentioned above, the standard polar coordinates do not cover quite the whole sphere, so one really must be prepared to use other coordinates to cover these "gaps". For these reasons, we need our calculus to deal with general coordinate systems and behave in a well defined way under change of coordinates. As we will see **differential forms** provide a very elegant language for doing this. This will allow us to develop a general theory for the local description of surfaces based on arbitrary coordinate systems.

1.4 Outline of this course

The content of this course can be broadly organised into the following parts.

Curves in Euclidean space We develop the general theory of curves in Euclidean space and show that they are completely classified by two quantities: curvature and torsion.

Vector fields and the calculus of differential forms We define vector fields and the calculus of differential forms on $D \subseteq \mathbb{R}^n$. We also develop the concept of moving frames and the structure equations. These tools will be used to define the curvature of surfaces.

Surfaces in Euclidean space In this part of the course we will develop a general local theory of surfaces in $\mathbb{E}^3$. In particular, we will introduce the notion of curvature for such objects and interpret its meaning. We will also discuss geodesic curves on surfaces.

Integration on surfaces and applications We discuss a general theory of integration over surfaces. This will give us the tools to discuss the celebrated Gauss-Bonnet theorem.

These lecture notes are based on notes by James Lucietti. In turn those were based on notes by José Figueroa-O'Farrill and Toby Bailey.

Section 2: Curves in Euclidean space

In this section we will develop the general theory of curves in Euclidean space. This will help illustrate some (but not all!) of the concepts encountered in the theory of surfaces.

2.1 Regular curves

A **(parametrised) curve in Euclidean space** is a (smooth) map $\boldsymbol{x} : (a, b) \to \mathbb{E}^n$,

$$t \mapsto \boldsymbol{x}(t) = \begin{pmatrix} x^1(t) \\ x^2(t) \\ \vdots \\ x^n(t) \end{pmatrix} .$$

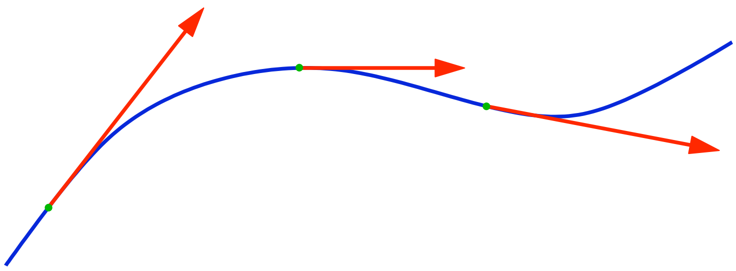
An **orientation** of the curve is a prescribed direction of travel along the curve. We will always take this to be the direction in which the parameter t is increasing.

Definition 2.1. *The* **velocity (or tangent) vector** *of the curve $\boldsymbol{x}(t)$ is given by*

$$\boldsymbol{x}'(t) = \begin{pmatrix} (x^1)'(t) \\ (x^2)'(t) \\ \vdots \\ (x^n)'(t) \end{pmatrix}$$

A curve $\boldsymbol{x}(t)$ is **regular** *if its velocity $\boldsymbol{x}'(t) \neq 0$ for all $t \in (a, b)$. The* **tangent line** *to a regular curve $\boldsymbol{x}(t)$ at $\boldsymbol{x}(t^0)$ is the line $\{\boldsymbol{x}(t^0) + \lambda\boldsymbol{x}'(t^0) \mid \lambda \in \mathbb{R}\}$.*

For each t, we can think of $\boldsymbol{x}'(t)$ as a vector, sitting at $\boldsymbol{x}(t)$, tangent to the curve $\boldsymbol{x}(t)$. This is depicted in the figure below, where the arrows denote the velocities of the curve at the indicated points.



Definition 2.2. *The norm of the velocity*

$$v(t) = \sqrt{\boldsymbol{x}'(t) \cdot \boldsymbol{x}'(t)}$$

is the **speed** *of the curve at $\boldsymbol{x}(t)$. Note $v(t) > 0$ for all t if and only if the curve is regular. A parametrisation of a regular curve $\boldsymbol{x}(t)$ such that $v(t) = 1$ everywhere is called a* **unit-speed parametrisation***.*

Definition 2.3. *The* **arc-length** *of a regular curve $\boldsymbol{x} : (a, b) \to \mathbb{E}^n$ from $\boldsymbol{x}(t^0)$ to $\boldsymbol{x}(t)$ is*

$$s(t) = \int_{t^0}^{t} v(t)\mathrm{d}t .$$

For a unit-speed parameterisation $s = t - t^0$, hence it is also called an **arc-length parametrisation***.*

Example 2.4. The helix in $\mathbb{E}^3$ defined by

$$\boldsymbol{x}(s) = \frac{1}{\sqrt{2}} \begin{pmatrix} \cos s \\ \sin s \\ s \end{pmatrix}$$

is arc-length parameterised.

Definition 2.5. *Let $g : (k, l) \to (a, b), \tau \mapsto t$ be a smooth map with $g'(\tau) \neq 0$ everywhere. The curve $\tilde{\boldsymbol{x}}(\tau) = \boldsymbol{x}(g(\tau))$ is a* **reparametrisation** *of $\boldsymbol{x}(t)$. If $g' > 0$ it preserves orientation (the curve is traversing the same path in the same direction, but at a different speed).*

Theorem 2.6. *For any regular curve $\boldsymbol{x} : (a, b) \to \mathbb{E}^n$, there exists a reparameterisation which is unit-speed.*

Proof. The derivative of the arc-length along the curve is $s'(t) = v(t) > 0$, so by the inverse function theorem we can write $t = g(s)$ for some function g as above. Now, consider the reparameterisation $\tilde{\boldsymbol{x}}(s) = \boldsymbol{x}(g(s))$. By the chain rule,

$$\tilde{\boldsymbol{x}}'(s) = g'(s)\boldsymbol{x}'(g(s)) = \frac{dt}{ds}\boldsymbol{x}'(t) = \frac{1}{v(t)}\boldsymbol{x}'(t) \ .$$

Thus, $\tilde{\boldsymbol{x}}'(s) \cdot \tilde{\boldsymbol{x}}'(s) = v^{-2}\boldsymbol{x}'(t) \cdot \boldsymbol{x}'(t) = 1$, as required. □

Therefore unit speed parametrisations always exist. However, they are not usually easy to compute explicitly. Consequently they are mainly a theoretical tool.

Definition 2.7. *A* **vector field along the curve** *$\boldsymbol{x}(t)$ in Euclidean space is a vector function $\boldsymbol{v} : (a, b) \to \mathbb{E}^n$. (Think of $\boldsymbol{v}(t)$ as a vector "sitting at" $\boldsymbol{x}(t)$. We will make this precise later.)*

Definition 2.8. *The* **unit tangent** *vector field along a regular curve is $\boldsymbol{x}(t)$ is*

$$\boldsymbol{T}(t) = \frac{\boldsymbol{x}'(t)}{v(t)}$$

Thus, for a unit-speed curve $\boldsymbol{x}(s)$ it is simply $\boldsymbol{T}(s) = \boldsymbol{x}'(s)$.

Definition 2.9. *For a unit-speed curve $\boldsymbol{x}(s)$ the curvature $\kappa(s)$ is defined by*

$$\kappa(s) = |\boldsymbol{T}'(s)| \ \ (= |\boldsymbol{x}''(s)|)$$

For a general parameterisation this becomes $\kappa(t) = |\boldsymbol{T}'(t)|/v(t)$ (chain rule).

Remark 2.10. The curvature at a point on a curve measures how fast the curve is *bending* away from its tangent line at that point.

Example 2.11. For the helix in example 2.4

$$\boldsymbol{x}''(s) = -\frac{1}{\sqrt{2}} \begin{pmatrix} \cos s \\ \sin s \\ 0 \end{pmatrix}$$

and hence $\kappa = 1/\sqrt{2}$.

2.2 The Frenet-Serret frame and the structure equations

Until further notice, we specialise to the case $n = 3$ so we are considering curves in Euclidean space $\mathbb{E}^3$. These are sometimes referred to as **space-curves**. The $n = 2$ case is a special case which are called **plane curves**.

Definition 2.12. *A unit-speed curve $\boldsymbol{x}(s)$ is* **biregular** *if $\kappa(s) \neq 0$ for all values of s. (Note that a unit-speed curve is necessarily regular.)*

Definition 2.13. *The* **principal normal** *along a unit-speed biregular curve $\boldsymbol{x}(s)$ is*

$$\boldsymbol{N}(s) = \frac{\boldsymbol{T}'(s)}{|\boldsymbol{T}'(s)|} = \frac{\boldsymbol{T}'(s)}{\kappa}$$

The **binormal** *vector field along $\boldsymbol{x}(s)$ is*

$$\boldsymbol{B} = \boldsymbol{T} \times \boldsymbol{N}$$

Proposition 2.14. *The vector fields $\{\boldsymbol{T}(s), \boldsymbol{N}(s), \boldsymbol{B}(s)\}$ along a biregular curve $\boldsymbol{x}(s)$ are an orthonormal basis for $\mathbb{E}^3$ for each s. This is called the* **Frenet-Serret frame** *of $\boldsymbol{x}(s)$.*

Proof. By definition of the unit tangent, $\boldsymbol{T}(s) \cdot \boldsymbol{T}(s) = 1$. Differentiate this with respect to s to find $\boldsymbol{T}(s) \cdot \boldsymbol{T}'(s) = 0$ for all s. Thus, the principal normal satisfies $\boldsymbol{N} \cdot \boldsymbol{N} = 1$ and $\boldsymbol{T} \cdot \boldsymbol{N} = \boldsymbol{T} \cdot \boldsymbol{T}'/\kappa = 0$. By definition of the binormal we also have $\boldsymbol{B} \cdot \boldsymbol{B} = 1$, $\boldsymbol{B} \cdot \boldsymbol{T} = 0$ and $\boldsymbol{B} \cdot \boldsymbol{N} = 0$. Hence $\boldsymbol{T}, \boldsymbol{N}, \boldsymbol{B}$ form an orthonormal basis. □

Thus, any vector field $\boldsymbol{v}(s)$ along a biregular curve may be expanded in this basis so,

$$\boldsymbol{v}(s) = (\boldsymbol{v} \cdot \boldsymbol{T})\, \boldsymbol{T}(s) + (\boldsymbol{v} \cdot \boldsymbol{N})\, \boldsymbol{N}(s) + (\boldsymbol{v} \cdot \boldsymbol{B})\, \boldsymbol{B}(s)$$

where we have used the fact the frame is orthonormal to express the coefficients in terms of $\boldsymbol{v}$ (take dot products to check!). Now consider expanding $\boldsymbol{B}'$ in the Frenet-Serret frame. Observe $\boldsymbol{B}' = \boldsymbol{T}' \times \boldsymbol{N} + \boldsymbol{T} \times \boldsymbol{N}' = \boldsymbol{T} \times \boldsymbol{N}'$ and hence $\boldsymbol{B}' \cdot \boldsymbol{T} = 0$. Also, since $\boldsymbol{B}$ is a unit vector, by the same argument as above $\boldsymbol{B}' \cdot \boldsymbol{B} = 0$. Thus, $\boldsymbol{B}' = (\boldsymbol{B}' \cdot \boldsymbol{N})\, \boldsymbol{N}$.

Definition 2.15. *The* **torsion** *of a biregular uni-speed curve $\boldsymbol{x}(s)$ is defined by*

$$\boldsymbol{B}' = -\tau \boldsymbol{N}$$

or equivalently $\tau = -\boldsymbol{B}' \cdot \boldsymbol{N}$.

Theorem 2.16. *Let $\boldsymbol{x}$ be a unit-speed biregular curve in $\mathbb{E}^3$. The Frenet-Serret frame $\boldsymbol{T}(s), \boldsymbol{N}(s), \boldsymbol{B}(s)$ along $\boldsymbol{x}(s)$ satisfies*

$$\begin{aligned} \boldsymbol{T}' &= \kappa \boldsymbol{N} \\ \boldsymbol{N}' &= -\kappa \boldsymbol{T} + \tau \boldsymbol{B} \\ \boldsymbol{B}' &= -\tau \boldsymbol{N}\ . \end{aligned}$$

These are called the **structure equations for a unit-speed space curve***, or sometimes the "Frenet-Serret equations".*

Proof. We have already established the first and third formulas. To get the second equation, first note that since $(\boldsymbol{T}(s), \boldsymbol{N}(s), \boldsymbol{B}(s))$ is an oriented (right-handed) orthonormal basis, $\boldsymbol{N} = \boldsymbol{B} \times \boldsymbol{T}$. Now differentiate

$$\boldsymbol{N}' = \boldsymbol{B}' \times \boldsymbol{T} + \boldsymbol{B} \times \boldsymbol{T}' = -\tau \boldsymbol{N} \times \boldsymbol{T} + \kappa \boldsymbol{B} \times \boldsymbol{N} = \tau \boldsymbol{B} - \kappa \boldsymbol{T}$$

where the first and third equations have been used. □

Example 2.17. Consider the helix of example (2.4). The Frenet-Serret frame is

$$\boldsymbol{T}(s) = \frac{1}{\sqrt{2}} \begin{pmatrix} -\sin s \\ \cos s \\ 1 \end{pmatrix}, \qquad \boldsymbol{N}(s) = \begin{pmatrix} -\cos s \\ -\sin s \\ 0 \end{pmatrix}, \qquad \boldsymbol{B}(s) = \frac{1}{\sqrt{2}} \begin{pmatrix} \sin s \\ -\cos s \\ 1 \end{pmatrix}.$$

Computing $\boldsymbol{B}'$ and using the structure equation we find $\tau = 1/\sqrt{2}$.

For explicit examples it is often more convenient to work in non unit-parameterisations.

Theorem 2.18. *For a general parametrisation $\boldsymbol{x}(t)$ of a biregular space curve the structure equations become*

$$\begin{aligned} \boldsymbol{T}' &= v\kappa \boldsymbol{N} \\ \boldsymbol{N}' &= -v\kappa \boldsymbol{T} + v\tau \boldsymbol{B} \\ \boldsymbol{B}' &= -v\tau \boldsymbol{N}, \end{aligned}$$

where v is the speed of the curve.

Proof. If s is a unit-speed parameter, $\frac{ds}{dt} = v$, and the result follows from the chain rule. □

The structure equations can be used to derive a variety of properties of space curves. The following gives an interpretation for the torsion.

Theorem 2.19. *A biregular curve $\boldsymbol{x}(t)$ is a plane curve if and only if $\tau = 0$ everywhere.*

Proof. If $\boldsymbol{x}(t)$ lies in a plane, then $\boldsymbol{T}$ and $\boldsymbol{N}$ are tangent to the plane and so $\boldsymbol{B}$ must be a unit normal to this plane and hence constant. The structure equations then imply $\tau = 0$. On the other hand, if $\tau = 0$ then $\boldsymbol{B}' = 0$ and so $\boldsymbol{B}$ is constant. Then

$$(\boldsymbol{x} \cdot \boldsymbol{B})' = \boldsymbol{x}' \cdot \boldsymbol{B} = v\boldsymbol{T} \cdot \boldsymbol{B} = 0$$

and so $\boldsymbol{x}(t)$ is contained in a plane $\boldsymbol{x} \cdot \boldsymbol{B} =$ constant. □

Remark 2.20. The **osculating plane** at a point on a curve is the plane spanned by $\boldsymbol{T}$ and $\boldsymbol{N}$. The torsion measures how fast the curve is *twisting* out of this plane.

For computational purposes the following general expressions are sometimes useful.

Theorem 2.21. *The curvature and torsion of a biregular space curve in any parameterisation can be computed by*

$$v^3\kappa = |\boldsymbol{x}' \times \boldsymbol{x}''|, \quad v^6\kappa^2\tau = (\boldsymbol{x}' \times \boldsymbol{x}'') \cdot \boldsymbol{x}'''.$$

Proof. For simplicity consider a unit-speed curve. Differentiating $\boldsymbol{x}' = \boldsymbol{T}$ and using the structure equations gives

$$\begin{aligned} \boldsymbol{x}'' &= \kappa \boldsymbol{N} \\ \boldsymbol{x}''' &= -\kappa^2 \boldsymbol{T} + \kappa' \boldsymbol{N} + \kappa\tau \boldsymbol{B}. \end{aligned}$$

The claim then follows from a few short computations. □

2.3 The equivalence problem

The "equivalence problem" is the problem of classifying all curves up to "rigid motions".

Definition 2.22. *An* **isometry** *of* $\mathbb{E}^3$ *is a map* $\mathbb{E}^3 \to \mathbb{E}^3$ *given by*

$$\boldsymbol{x} \mapsto A\boldsymbol{x} + \boldsymbol{b}$$

where A *is an orthogonal matrix (i.e.* $A^T A = 1$*) and* $\boldsymbol{b}$ *is a fixed vector. If* $\det A = 1$*, so that* A *is a rotation matrix, then the isometry is said to be a* **Euclidean motion** *or a* **rigid motion***. If* $\det A = -1$ *the isometry is orientation-reversing.*

By definition, an isometry preserves the Euclidean distance between two points $|\boldsymbol{x} - \boldsymbol{y}|$. This is equivalent to the inner product being preserved

$$(A\boldsymbol{x}) \cdot (A\boldsymbol{y}) = \boldsymbol{x} \cdot \boldsymbol{y} \ .$$

In general $(A\boldsymbol{x}) \times (A\boldsymbol{y}) = (\det A)\, A(\boldsymbol{x} \times \boldsymbol{y})$ and therefore for an orthogonal A we have

$$(A\boldsymbol{x}) \times (A\boldsymbol{y}) = \pm A(\boldsymbol{x} \times \boldsymbol{y}) \ ,$$

where the sign depends on $\det A = \pm 1$.

Proposition 2.23. *The curvature and torsion of a biregular curve are unchanged by a Euclidean motion. In other words, taking a unit speed parameterisation, if*

$$\hat{\boldsymbol{x}}(s) = A\boldsymbol{x}(s) + \boldsymbol{b},$$

where A *is a fixed rotation matrix and* $\boldsymbol{b}$ *is a fixed vector, then*

$$\hat{\kappa}(s) = \kappa(s), \quad \hat{\tau}(s) = \tau(s).$$

If instead $\det A = -1$ *then*

$$\hat{\kappa}(s) = \kappa(s), \quad \hat{\tau}(s) = -\tau(s).$$

Proof. Because A is just a constant matrix, differentiating $\hat{\boldsymbol{x}}'(s) = A\boldsymbol{x}'(s)$ and similarly for higher derivatives. It then follows that if $\det A = 1$ (i.e. a rotation matrix) then

$$\hat{\boldsymbol{T}} = A\boldsymbol{T}, \quad \hat{\boldsymbol{N}} = A\boldsymbol{N}, \quad \hat{\boldsymbol{B}} = A\boldsymbol{B}.$$

Substituting into the structure equations gives the result. If A has $\det A = -1$ then A takes oriented bases to unoriented bases

$$\hat{\boldsymbol{T}} = A\boldsymbol{T}, \quad \hat{\boldsymbol{N}} = A\boldsymbol{N}, \quad \hat{\boldsymbol{B}} = -A\boldsymbol{B}$$

and the structure equations give the result. □

Theorem 2.24. *Let* $\kappa(s), \tau(s)$ *be given, with* $\kappa(s)$ *everywhere positive. Then there exists a unique unit-speed biregular curve* $\boldsymbol{x}(s)$ *with these as curvature and torsion and such that* $\boldsymbol{x}(0) = \boldsymbol{b}$ *and* $\{\boldsymbol{T}(0), \boldsymbol{N}(0), \boldsymbol{B}(0)\}$ *is any fixed oriented orthonormal basis of* $\mathbb{E}^3$*.*

Proof. It is convenient to express the structure equations in matrix form

$$\frac{d}{ds}\begin{pmatrix}\boldsymbol{T} & \boldsymbol{N} & \boldsymbol{B}\end{pmatrix} = \begin{pmatrix}\boldsymbol{T} & \boldsymbol{N} & \boldsymbol{B}\end{pmatrix}\begin{pmatrix}0 & -\kappa & 0\\ \kappa & 0 & -\tau\\ 0 & \tau & 0\end{pmatrix} \tag{1}$$

By ODE theory, for given $\kappa(s), \tau(s)$ and initial condition $(\boldsymbol{T}(0)\ \boldsymbol{N}(0)\ \boldsymbol{B}(0))$, there exists a unique solution $(\boldsymbol{T}(s)\ \boldsymbol{N}(s)\ \boldsymbol{B}(s))$ to the ODE system (1) and hence it must coincide with the Frenet-Serret frame. There is then a unique curve $\boldsymbol{x}(s)$ satisfying

$$\boldsymbol{x}' = \boldsymbol{T}, \quad \boldsymbol{x}(0) = \boldsymbol{b}$$

again by the usual theorem on existence and uniqueness for ODEs. □

We are now ready to deduce our main result.

Theorem 2.25. (**The fundamental theorem of curves**). *If two biregular space curves have the same curvature and torsion then they differ at most by a Euclidean motion.*

Proof. We have established that for given $\kappa(s), \tau(s)$, the only freedom in fixing the curve is given by the initial conditions. These are given by a constant vector $\boldsymbol{x}(0)$ and a 3×3 matrix $(\boldsymbol{T}(0)\ \boldsymbol{N}(0)\ \boldsymbol{B}(0))$ which is orthogonal with unit determinant. Thus any two sets of initial conditions differ by a translation and a rotation. □

This theorem completely classifies all biregular curves (up to rigid motion) in terms of just two functions.

> ❝ Biregular curves are determined by their curvature and torsion (up to Euclidean motions) ❞

Corollary 2.26. *Every biregular curve with κ, τ both constant is a helix, unless $\tau = 0$ in which case it is a circle.*

Proof. We know from calculations that helices and circles have this property. The uniqueness in the Theorem then implies that every curve with these properties is a helix. □

2.4 Exercises

Exercise 2.27. Show that the curve $\boldsymbol{x} : (-1, 1) \to \mathbb{E}^3$

$$\boldsymbol{x}(s) = \begin{pmatrix}\frac{(1+s)^{3/2}}{3}\\ \frac{(1-s)^{3/2}}{3}\\ \frac{s}{\sqrt{2}}\end{pmatrix}$$

is arc-length parameterised. Compute its Frenet-Serret frame, curvature and torsion.

Solution. We have

$$\boldsymbol{x}'(s) = \begin{pmatrix}\frac{(1+s)^{1/2}}{2}\\ -\frac{(1-s)^{1/2}}{2}\\ \frac{1}{\sqrt{2}}\end{pmatrix}$$

so $|\boldsymbol{x}'(s)|^2 = \frac{1}{4}(1+s) + \frac{1}{4}(1-s) + \frac{1}{2} = 1$, so the curve is arc-length parameterised. Hence $\boldsymbol{T} = \boldsymbol{x}'$. Next,

$$\boldsymbol{T}'(s) = \begin{pmatrix} \frac{1}{4}(1+s)^{-1/2} \\ \frac{1}{4}(1-s)^{-1/2} \\ 0 \end{pmatrix}$$

so

$$\kappa(s) = \frac{1}{2\sqrt{2}\sqrt{1-s^2}}, \qquad \boldsymbol{N}(s) = \begin{pmatrix} \frac{1}{\sqrt{2}}(1-s)^{1/2} \\ \frac{1}{\sqrt{2}}(1+s)^{1/2} \\ 0 \end{pmatrix} .$$

Therefore

$$\boldsymbol{B}(s) = \begin{pmatrix} -\frac{1}{2}(1+s)^{1/2} \\ \frac{1}{2}(1-s)^{1/2} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

so that

$$\boldsymbol{B}'(s) = -\frac{1}{2\sqrt{2}\sqrt{1-s^2}} \begin{pmatrix} \frac{1}{\sqrt{2}}(1-s)^{1/2} \\ \frac{1}{\sqrt{2}}(1+s)^{1/2} \\ 0 \end{pmatrix}$$

and hence

$$\tau(s) = \frac{1}{2\sqrt{2}\sqrt{1-s^2}} .$$

◆

Exercise 2.28. Consider the unit speed plane curve

$$\boldsymbol{x}(s) = \begin{pmatrix} x(s) \\ y(s) \\ 0 \end{pmatrix}$$

and let $\phi(s)$ be the angle between $\mathbf{T}(s)$ and the x-axis. Show that $\kappa(s) = |\phi'(s)|$.

Solution. If we let $\boldsymbol{e}_1, \boldsymbol{e}_2$ be the standard unit-vectors pointing in the x, y-axis, by definition of $\phi(s)$ we have

$$\boldsymbol{T}(s) = \cos\phi(s)\boldsymbol{e}_1 + \sin\phi(s)\boldsymbol{e}_2 .$$

Hence

$$\boldsymbol{T}'(s) = -\phi'(s)\sin\phi(s)\boldsymbol{e}_1 + \phi'(s)\cos\phi(s)\boldsymbol{e}_2 .$$

and therefore $|\boldsymbol{T}'|^2 = \phi'(s)^2$, so using $\boldsymbol{T}' = \kappa\boldsymbol{N}$ the result follows. ◆

Exercise 2.29. Let $\boldsymbol{x}(s)$ be a unit-speed biregular curve such that $\boldsymbol{T}(0), \boldsymbol{N}(0), \boldsymbol{B}(0)$ are the standard oriented orthonormal basis $\mathbf{i}, \mathbf{j}, \mathbf{k}$ of $\mathbb{E}^3$. Show that for small s

$$\boldsymbol{x}(s) = \boldsymbol{x}(0) + \left(s - \frac{1}{6}\kappa(0)^2 s^3\right)\mathbf{i} + \left(\frac{1}{2}\kappa(0)s^2 + \frac{1}{6}\kappa'(0)s^3\right)\mathbf{j} + \frac{1}{6}\kappa(0)\tau(0)s^3\boldsymbol{k} + \mathcal{O}(s^4)$$

Solution. Taylor expand the curve about $s = 0$

$$\boldsymbol{x}(s) = \boldsymbol{x}(0) + s\boldsymbol{x}'(0) + \frac{1}{2}s^2\boldsymbol{x}''(0) + \frac{1}{6}s^3\boldsymbol{x}'''(0) + \mathcal{O}(s^4) .$$

and evaluate the expressions in the proof of Theorem 2.21 for $\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}'''$ at $s = 0$. ◆

Exercise 2.30. Compute the curvature and torsion of the arc-length parameterised helix in example 2.4 at $s = 0$ by Taylor expansion.

Solution. Taylor expanding

$$\boldsymbol{x}(s) = \frac{1}{\sqrt{2}} \begin{pmatrix} \cos s \\ \sin s \\ s \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 - \frac{s^2}{2} + O(s^4) \\ s - \frac{s^3}{6} + O(s^5) \\ s \end{pmatrix}$$

which gives

$$\boldsymbol{T}(0) = \boldsymbol{x}'(0) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \qquad \boldsymbol{x}''(0) = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} = \kappa(0)\boldsymbol{N}(0)$$

and hence $\kappa(0) = 1/\sqrt{2}$. Hence

$$\boldsymbol{B}(0) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$$

and thus proposition (2.29) gives $\kappa(0)\tau(0) = \boldsymbol{x}'''(0) \cdot \boldsymbol{B}(0) = 1/2$, and hence $\tau(0) = 1/\sqrt{2}$. ◆

Exercise 2.31. Prove Theorem 2.21 by explicit computation using the structure equations in a general parameterisation.

Solution. Differentiating $\boldsymbol{x}' = v\boldsymbol{T}$ gives $\boldsymbol{x}'' = v\boldsymbol{T}' + v'\boldsymbol{T} = v^2\kappa\boldsymbol{N} + v'\boldsymbol{T}$, where the second equality uses the structure equation $\boldsymbol{T}' = v\kappa\boldsymbol{N}$. Hence $\boldsymbol{x}' \times \boldsymbol{x}'' = v^3\kappa\boldsymbol{T} \times \boldsymbol{N} = v^3\kappa\boldsymbol{B}$ so $|\boldsymbol{x}' \times \boldsymbol{x}''| = v^3\kappa$. Simarly for the third derivative.

◆

Exercise 2.32. Define the **Darboux** vector $\boldsymbol{\Omega} = v\tau\boldsymbol{T} + v\kappa\boldsymbol{B}$. Show that the structure equations take the elegant form

$$\boldsymbol{T}' = \boldsymbol{\Omega} \times \boldsymbol{T}\,, \qquad \boldsymbol{N}' = \boldsymbol{\Omega} \times \boldsymbol{N}\,, \qquad \boldsymbol{B}' = \boldsymbol{\Omega} \times \boldsymbol{B}$$

and $|\boldsymbol{\Omega}| = v\sqrt{\kappa^2 + \tau^2}$. If you have done any physics, you may recognise $\boldsymbol{\Omega}$ as the angular velocity vector of a particle moving along the curve.

Solution. Recall $\boldsymbol{T}, \boldsymbol{N}, \boldsymbol{B}$ is an oriented orthonormal frame along the curve. We thus have $\boldsymbol{\Omega} \times \mathbf{T} = v\kappa\boldsymbol{B} \times \boldsymbol{T} = v\kappa\boldsymbol{N}$, so that the structure equation $\boldsymbol{T}' = v\kappa\boldsymbol{N}$, can indeed be rewritten as above. The others follows similarly from the other structure equations. Also since $\boldsymbol{T}, \boldsymbol{B}$ are orthonormal $\boldsymbol{\Omega}^2 = v^2\tau^2 + v^2\kappa^2$. ◆

Exercise 2.33. Let $(\boldsymbol{T}_1(s), \boldsymbol{N}_1(s), \boldsymbol{B}_1(s))$ and $(\boldsymbol{T}_2(s), \boldsymbol{N}_2(s), \boldsymbol{B}_2(s))$ be two Frenet-Serret frames with the same curvature $\kappa(s)$ and torsion $\tau(s)$. Use the structure equations to show

$$\frac{d}{ds}\left(|\boldsymbol{T}_1 - \boldsymbol{T}_2|^2 + |\boldsymbol{N}_1 - \boldsymbol{N}_2|^2 + |\boldsymbol{B}_1 - \boldsymbol{B}_2|^2\right) = 0$$

Deduce that if the two frames coincide at a point on the curve they coincide everywhere along the curve. This gives a proof of the uniqueness part of the equivalence problem.

Solution. Differentiating the LHS one gets

$$\begin{aligned}
&\frac{d}{ds}\left(|\boldsymbol{T}_1-\boldsymbol{T}_2|^2+|\boldsymbol{N}_1-\boldsymbol{N}_2|^2+|\boldsymbol{B}_1-\boldsymbol{B}_2|^2\right)\\
&\quad=2(\boldsymbol{T}_1-\boldsymbol{T}_2)\cdot(\boldsymbol{T}_1'-\boldsymbol{T}_2')+2(\boldsymbol{N}_1-\boldsymbol{N}_2)\cdot(\boldsymbol{N}_1'-\boldsymbol{N}_2')+2(\boldsymbol{B}_1-\boldsymbol{B}_2)\cdot(\boldsymbol{B}_1'-\boldsymbol{B}_2')\\
&\quad=2\kappa(\boldsymbol{T}_1-\boldsymbol{T}_2)\cdot(\boldsymbol{N}_1-\boldsymbol{N}_2)+2(\boldsymbol{N}_1-\boldsymbol{N}_2)\cdot[-\kappa(\boldsymbol{T}_1-\boldsymbol{T}_2)+\tau(\boldsymbol{B}_1-\boldsymbol{B}_2)]\\
&\qquad-2\tau(\boldsymbol{B}_1-\boldsymbol{B}_2)\cdot(\boldsymbol{N}_1-\boldsymbol{N}_2)\\
&\quad=0
\end{aligned}$$

where we have used the structure equations to simplify the derivative terms. Therefore $|\boldsymbol{T}_1-\boldsymbol{T}_2|^2+|\boldsymbol{N}_1-\boldsymbol{N}_2|^2+|\boldsymbol{B}_1-\boldsymbol{B}_2|^2$ is a constant. Hence if it vanishes at a point, it vanishes everywhere.

◆

Section 3: Vector fields and one-forms

3.1 Vector spaces

We begin by recalling some central concepts in linear algebra, phrased in a more general context than the familiar vectors in Euclidean space.

Definition 3.1. *A (real)* **vector space** *V is a set together with two operations, 'addition' and 'scalar multiplication', which obey the following axioms; for all $u, v, w \in V$ and $a, b \in \mathbb{R}$*

$$\begin{aligned}
\mathbf{u} + (\mathbf{v} + \mathbf{w}) &= (\mathbf{u} + \mathbf{v}) + \mathbf{w} && \textit{associativity of addition} \\
\mathbf{v} + \mathbf{w} &= \mathbf{w} + \mathbf{v} && \textit{commutativity of addition} \\
\mathbf{v} + 0 &= \mathbf{v} && \textit{existence of an identity element } 0 \in V \\
\mathbf{v} + (-\mathbf{v}) &= 0 && \textit{existence of an inverse element } -v \in V \\
a(\mathbf{v} + \mathbf{w}) &= a\mathbf{v} + a\mathbf{w} && \textit{distributivity of scalar multiplication} \\
(a + b)\mathbf{v} &= a\mathbf{v} + b\mathbf{v} && \textit{distributivity of scalar multiplication} \\
a(b\mathbf{v}) &= (ab)\mathbf{v} && \textit{compatibility of scalar multiplication} \\
1v &= v && \textit{scalar multiplication identity}
\end{aligned}$$

The prototypical example of a vector space is given by $\mathbb{R}^n$ together with the 'obvious' operations of addition $\boldsymbol{x} + \mathbf{y} = (x^1 + y^1, \ldots, x^n + y^n)$ and scalar multiplication $a\boldsymbol{x} = (ax^1, \ldots, ax^n)$. However, it is important to realise that it is an abstract definition and there are many other examples of vector spaces (as we will see!).

Example 3.2. The set of real polynomials $p_n(X) = a_n X^n + \ldots a_1 X + a_0$ of order n is a real vector space under the standard operations of addition and real multiplication.

Definition 3.3. *A* **basis** *of V is a set $\{\mathbf{e}_i \in V : i = 1, \ldots, n\}$ with the following properties:*

- *Spans V: for all $\mathbf{v} \in V$, $v = \sum_{i=1}^n v^i \mathbf{e}_i$, where $\mathbf{v}_i \in \mathbb{R}$*
- *Linearly independent: $\sum_{i=1}^n a^i \mathbf{e}_i = 0 \implies a^i = 0$ for all $i = 1, \ldots n$*

The number of elements in a basis n is called the dimension of V (it's independent of basis).

For the familiar example $\mathbb{R}^n$ a natural basis is given by $\mathbf{e}_i = (0 \ldots, 0, 1, 0, \ldots, 0)$ where the 1 is in the ith position and the rest of the entries are 0. Thus $\mathbb{R}^n$ has dimension n.

Example 3.4. A basis for polynomials of order n is $\{1, X, \ldots, X^n\}$. Hence the dimension of this space is $n + 1$.

3.2 Tangent space and vector fields

Euclidean space $\mathbb{E}^n$ is a vector space, so points in it can be thought of as position vectors. However, we sometimes think of a vector as having no particular location in space. More generally, we can think of a vector as being associated with a particular point in space, for example a vector field associates a different vector to each point in space, or a vector modelling a force will often also have a "point of application". Think perhaps of a vector with its tail sitting at some particular point in space. In contrast, a surface is not a vector space (e.g. the sum of two points on a sphere will not in general remain on the sphere). In fact, surfaces and more generally manifolds, are spaces which near any point "look like a small piece" of Euclidean space (we will make this precise later).

Thus, let us recall a key concept from several variable calculus in Euclidean space $\mathbb{E}^n$. If $\boldsymbol{v}$ is a vector at $p \in \mathbb{E}^n$ and $f : \mathbb{E}^n \to \mathbb{R}$ is a differentiable function, the directional derivative of f along $\boldsymbol{v}$ at p is

$$(D_{\boldsymbol{v}} f)_p = (\boldsymbol{v} \cdot \boldsymbol{\nabla} f)(p) = \sum_{k=1}^n v^k \frac{\partial f}{\partial x^k}\Big|_p$$

More abstractly, we can define the derivative operator acting on functions f at p,

$$\sum_{k=1}^n v^k \frac{\partial}{\partial x^k}\Big|_p \quad : \quad f \mapsto \sum_{k=1}^n v^k \frac{\partial f}{\partial x^k}\Big|_p$$

The vector $\boldsymbol{v}$ determines this operator and vice-versa. We will use these observations to define what we mean by a vector in $\mathbb{R}^n$, and more generally, by a vector in some arbitrary (open) domain $D \subseteq \mathbb{R}^n$ (itself not a vector space!).

Definition 3.5. *For each $p \in D$, we define the* **tangent space** T_pD **to** D **at** p *as the set of all derivative operators at p, called* **tangent vectors** *at $p \in D$,*

$$\sum_{i=1}^n v^i \frac{\partial}{\partial x^i}\Big|_p , \qquad v^i \in \mathbb{R} .$$

A smooth **vector field** *on D is a smooth map v which assigns to each $p \in D$ an element in T_pD; in other words, it is a differential operator*

$$v = \sum_{i=1}^n v^i \frac{\partial}{\partial x^i} ,$$

where now each $v^i : D \to \mathbb{R}$ is a smooth function.

> “ vector fields are directional derivatives ”

Example 3.6. $v = 2x^1 \frac{\partial}{\partial x^1} - 3x^1x^2x^3 \frac{\partial}{\partial x^2} + \frac{\partial}{\partial x^3}$ is a vector field on $\mathbb{R}^3$.

Definition 3.7. *If v, w are vector fields*

$$v = \sum_i v^i \frac{\partial}{\partial x^i} \qquad \textit{and} \qquad w = \sum_i w^i \frac{\partial}{\partial x^i}$$

and if $\lambda \in \mathbb{R}$, then λv and $v + w$ are vector fields defined by

$$\lambda v = \sum_i \lambda v^i \frac{\partial}{\partial x^i} \qquad \textit{and} \qquad v + w = \sum_i (v^i + w^i) \frac{\partial}{\partial x^i} .$$

Furthermore if f is a function then we can define a vector field fv by

$$fv = \sum_i f v^i \frac{\partial}{\partial x^i} .$$

Theorem 3.8. *For any $D \subseteq \mathbb{R}^n$ and $p \in D$ the tangent space T_pD is a real vector space of dimension n. The set of tangent vectors $\left\{ \frac{\partial}{\partial x^1}\Big|_p, \ldots, \frac{\partial}{\partial x^n}\Big|_p \right\}$ is a basis of T_pD.*

Proof. The tangent space T_pD is a vector space under the operations of addition and scalar multiplication defined above. By definition the set $\left\{ \frac{\partial}{\partial x^1}\Big|_p, \ldots, \frac{\partial}{\partial x^n}\Big|_p \right\}$ spans T_pD and it is linearly independent because,

$$\sum_{i=1}^n a^i \frac{\partial}{\partial x^i}\Big|_p = 0 \implies \sum_{i=1}^n a^i \frac{\partial f}{\partial x^i}\Big|_p = 0 \implies a^j = 0 \quad \text{for all} \quad j = 1, \ldots n,$$

where the last implication follows from choosing $f = x^j$.

□

Definition 3.9. *Given a vector field v and a smooth function f, we can define a function*

$$v(f) = \sum_i v^i \frac{\partial f}{\partial x^i} \ .$$

Example 3.10. Let v be a vector field as in Example 3.6 and $f = x^1(x^3)^2$. Then $v(f) = 2x^1(x^3)^2 + 2x^1x^3$.

Proposition 3.11. *For all vector fields v, w and functions f, g, the following identities hold:*

$$\begin{aligned} (v+w)(f) &= v(f) + w(f) \\ (fv)(g) &= fv(g) \\ v(f+g) &= v(f) + v(g) \\ v(fg) &= gv(f) + fv(g) \ . \end{aligned}$$

Proof. These follow from the usual rules of differentiation. □

Remark 3.12. It can be shown that any map $f \mapsto v(f)$ from the set of smooth functions on D to itself which obeys the above identities defines a vector field. This gives an algebraic definition of a vector field.

The above definitions are very useful for practical computations, although may seem a bit abstract. There is an equivalent, more intuitive, way to think about tangent vectors, which directly generalises the concept used in Euclidean space. A (regular) **curve** in D is a (smooth) map $c : (a, b) \to D$, given by

$$c(t) = (x^1(t), x^2(t), \ldots, x^n(t)) \tag{2}$$

where each $x^i : (a, b) \to \mathbb{R}$ is a smooth function, such that its velocity

$$c'(t) = ((x^1)'(t), (x^2)'(t), \ldots, (x^n)'(t))$$

is non-vanishing (as an element of $\mathbb{R}^n$) for all $t \in (a, b)$. We say that a curve c **passes through** $p \in D$ if, say, $c(0) = p$ (without loss of generality one can always take the parameter value at p to be 0).

Proposition 3.13. *Let $c : (a,b) \to D$ be a curve that passes through p. There exists a unique $c'_p \in T_pD$ such that for any smooth function $f : D \to \mathbb{R}$*

$$c'_p(f) = \frac{d}{dt}\Big|_{t=0} f(c(t)) \tag{3}$$

Proof. The composition $f \circ c : (a,b) \to \mathbb{R}$ is a smooth function $(f \circ c)(t) = f(c(t))$. The rate of change of f along the curve c at p is

$$\frac{d}{dt}\Big|_{t=0} f(c(t)) = \sum_{i=1}^{n} (x^i)'(0) \frac{\partial f}{\partial x^i}\Big|_p$$

where we have used (2) and the chain rule for partial differentiation. The RHS is determined uniquely by the velocity of the curve at p , $c'(0)$, and defines a tangent vector at p

$$c'_p \equiv \sum_{i=1}^{n} (x_i)'(0) \frac{\partial}{\partial x^i}\Big|_p . \tag{4}$$

□

Corollary 3.14. *There is a one-to-one correspondence between velocities of curves that pass through $p \in D$ and tangent vectors in T_pD. By (standard) abuse of notation sometimes we denote c'_p by the corresponding velocity $c'(0)$.*

Proof. If $v = \sum_i v^i \frac{\partial}{\partial x^i}\big|_p$ is any tangent vector, then there is a curve with components $x^i(t) = p^i + v^i t$, so that $c'_p = v$. □

❝ tangent vectors are velocities of curves ❞

3.3 One-forms

Definition 3.15. *A* 1**-form at** *$p \in D$ is a linear map $\alpha : T_pD \to \mathbb{R}$. This means, for all $\mathbf{v}, \mathbf{w} \in T_pD$ and $a \in \mathbb{R}$,*

$$\alpha(\mathbf{v} + \mathbf{w}) = \alpha(\mathbf{v}) + \alpha(\mathbf{w}), \qquad \alpha(a\mathbf{v}) = a\alpha(\mathbf{v}) .$$

The set of 1-forms at $p \in D$, denoted by T^*_pD, is called the **dual vector space** *of T_pD.*

❝ One-forms are dual to vectors ❞

Definition 3.16. *We define 1-forms $dx^j|_p$ at each $p \in D$ by their action on the basis $\{\frac{\partial}{\partial x^i}\big|_p\}$:*

$$dx^j|_p \left(\frac{\partial}{\partial x^k}\Big|_p \right) = \begin{cases} 1, & \text{if } j = k \\ 0, & \text{otherwise.} \end{cases}$$

Equivalently, $dx^{|}_p i$ are defined by their action on an arbitrary tangent vector $\mathbf{v} = \sum_{i=1}^n v^i \frac{\partial}{\partial x^i}\big|_p$:

$$dx^j|_p(\mathbf{v}) = v^j .$$

(We will see the reason for the notation dx^i shortly.)

Remark 3.17. T_p^*D is a vector space of dimension n with basis $\{dx^1|_p, \dots, dx^n|_p\}$, called the **dual basis** to the basis $\left\{ \frac{\partial}{\partial x^1}\big|_p, \dots, \frac{\partial}{\partial x^n}\big|_p \right\}$ for T_pD.

Definition 3.18. *A* **differential** 1**-form** *on D is a smooth map α which assigns to each $p \in D$ a 1-form in T_p^*D; it can be written as*

$$\boxed{\alpha = \sum_{i=1}^{n} \alpha_i dx^i}$$

where $\alpha_i : D \to \mathbb{R}$ are smooth functions. Given 1-forms $\alpha = \sum_i \alpha_i dx^i$ and $\beta = \sum_i \beta_i dx^i$ and a function f on D, addition and scalar multiplication of 1-forms are given by

$$\alpha + \beta = \sum_i (\alpha_i + \beta_i) dx^i \,, \qquad f\alpha \;=\; \sum_i (f\alpha_i) dx^i \,.$$

Remark 3.19. Implicit in this definition is that dx^i is itself the 1-form that assigns $dx^i|_p \in T_p^*D$ to each $p \in D$.

Lemma 3.20. *If $\alpha = \sum_i \alpha_i dx^i$ and $\mathbf{v} = \sum_i v^i \frac{\partial}{\partial x^i}$, then $\alpha(\mathbf{v})$ is the smooth function*

$$\alpha(\mathbf{v}) = \sum_{i=1}^{n} \alpha_i v^i \;.$$

Proof. The action of the differential 1-form α on the vector field v is

$$\alpha(v) = \sum_{i=1}^{n} \alpha_i dx^i(v) = \sum_{i=1}^{n} \alpha_i v^i$$

where we have used the definition of dx^i. □

Example 3.21. An example of a 1-form in $\mathbb{R}^3$ is $\alpha = 3dx^1 + 2x^3 dx^2 - dx^3$. In this case, taking v as in Example 3.6, we have

$$\alpha(v) = 3(2x^1) + 2x^3(-3x^1x^2x^3) - 1(1) = 6x^1 - 6x^1x^2(x^3)^2 - 1.$$

Definition 3.22. *Given a smooth function f on D, its* **exterior derivative** *(or differential) is the* 1*-form df defined by*

$$(df)(v) = v(f)$$

for any vector field v. Equivalently,

$$\boxed{df = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i} dx^i \;.}$$

(We note that dx^i as defined earlier really is d of the coordinate function x^i.)

Remark 3.23. Notice that df encodes the same information as ∇f of ordinary vector calculus. Thus we see it emerges naturally as a 1-form rather than a vector field.

Proposition 3.24. *For any functions f, g on D, we have*

- $d(f+g) = df + dg$
- $d(fg) = fdg + gdf$
- $df = 0$ *if and only if f is constant.*

Proof. Using the coordinate formula for df above these identities follow from the standard properties of partial derivatives. □

3.4 Line integrals

> ❝ 1-forms can be integrated on (oriented) curves ❞

Given any curve $c : (a,b) \to D$, for each $t \in (a,b)$ there is a unique tangent vector $c'(t)$ to D at $c(t)$ given by

$$c'(t) = \sum_{i=1}^{n} (x^i)'(t) \frac{\partial}{\partial x^i}\Big|_{c(t)}$$

Sometimes we denote the tangent vector field along the curve simply by c'.

Definition 3.25. *Let $c : [a,b] \to D$ be a curve (the end points are included to ensure the integrals exist[2]) and α on 1-form on D. The* **integral of α over the curve** *c is*

$$\boxed{\int_c \alpha = \int_a^b \alpha(c'(t))dt \ .}$$

where $c'(t)$ is the tangent vector field to the curve.

Working in coordinates, the result of applying the 1-form α on $c'(t)$ gives the expression

$$\alpha(c') = \sum_{i=1}^{n} (x^i)'(t)\alpha_i(c(t))$$

at any point on the curve, which is useful for evaluating line integrals in specific examples.

Proposition 3.26. *Let $c : [a,b] \to D$ be a curve in D and let f a function on D. Then*

$$\int_c df = f(c(b)) - f(c(a)) \ .$$

[2]In order to avoid confusion about the existence of derivatives at the end-points of closed intervals, we shall stick to the following convention: whenever we consider a curve on a closed interval $[a,b]$, we always assume the definition of the curve can be extended smoothly to a larger open interval $(\tilde{a}, \tilde{b}) \supset [a,b]$.

Proof. Recall that $df(c') = c'(f)$. Since along the curve $c'(f) = \sum_i (x^i)'(t)\frac{\partial f}{\partial x^i} = \frac{d}{dt}f(c(t))$, the integral becomes

$$\int_c df = \int_a^b (df)(c'(t))dt = \int_a^b \frac{df(c(t))}{dt}dt = f(c(b)) - f(c(a)) \ .$$

$\square$

Corollary 3.27. *If c is a closed curve, i.e. $c(a) = c(b)$, then $\int_c df = 0$ for any function f.*

Example 3.28. Let D be $\mathbb{R}^2$ with the origin removed and consider the 1-form

$$\omega = \frac{x^1 dx^2 - x^2 dx^1}{(x^1)^2 + (x^2)^2} \ .$$

Let c be a circle of radius a about the origin and parameterise by

$$c(t) = (a\cos t, a\sin t) \qquad t \in [0, 2\pi] \ .$$

Then the tangent vector field along the curve is

$$c'(t) = -a\sin t\frac{\partial}{\partial x^1} + a\cos t\frac{\partial}{\partial x^2}.$$

Noting that $x^1(t)^2 + x^2(t)^2 = a^2$ on the curve one gets

$$\omega(c'(t)) = (-a\sin t)\left(-\frac{a\sin t}{a^2}\right) + (a\cos t)\left(\frac{a\cos t}{a^2}\right) = 1.$$

Hence $\int_c \omega = \int_0^{2\pi} dt = 2\pi$. In fact the polar angle $\theta = \arctan(x^2/x^1)$ satisfies $d\theta = \omega$. However, this does not contradict the above corollary because θ is *not* a smooth (or even continuous) function on D: it increases by 2π as you go round the origin explaining the value of the integral.

Proposition 3.29. *If $\tilde{c}$ is a reparametrisation of c and α is a 1-form then*

$$\int_c \alpha = \pm\int_{\tilde{c}} \alpha,$$

where the plus sign is taken when the reparametrisation is orientation-preserving, and the minus if it is orientation-reversing.

Proof. Suppose the reparametrisation is $\tilde{c}(t) = c(g(t))$ where $c : [a, b] \to D$ and $\tilde{c} : [k, l] \to D$. We have

$$\tilde{c}'(t) = g'(t)c'(g(t)).$$

Thus

$$\int_{\tilde{c}} \alpha = \int_k^l \alpha(\tilde{c}'(t))dt = \int_k^l \alpha(g'(t)c'(g(t)))dt = \int_k^l g'(t)\alpha(c'(g(t)))dt.$$

(In the last equality we are using the fact that a 1-form is a *linear* function of a vector.) Now, change variables by $\tau = g(t)$, and assuming for a moment that we are in the orientation-preserving case so that $g(k) = a$ and $g(l) = b$ this becomes

$$\int_a^b \alpha(c'(\tau))d\tau = \int_c \alpha$$

as desired. In the orientation-reversing case, the only difference is that $g(k) = b$ and $g(l) = a$ so that we get

$$\int_b^a \alpha(c'(\tau))d\tau = -\int_a^b \alpha(c'(\tau))d\tau = -\int_c \alpha$$

instead. □

Remark 3.30. Because of the reparametrisation invariance, the integral of 1-forms along an oriented curve $c : [a, b] \to D$ only depends on the image of that curve in D,

$$c([a, b]) = \{c(t) \mid a \le t \le b\} \ .$$

Definition 3.31. *The* **pull-back** *of α by the map $c : [a, b] \to D$ is the 1-form on the interval $[a, b]$ defined by*

$$c^*\alpha := \alpha(c'(t))dt \ . \tag{5}$$

The integral of α over c may now be thought of as the integral of the pull-back $c^*\alpha$ over the interval $[a, b]$.

Remark 3.32. Note how the identity $\int_a^b f(x)dx = -\int_b^a f(x)dx$ for integrals over the real line, finds a natural explanation in this new language. One is integrating the 1-form $f(x)dx$ over the oriented curve corresponding to the interval $[a, b]$. There is a natural orientation on the interval $[a, b]$ (that of increasing x), but if one chooses the opposite orientation one gets a sign reversal.

Remark 3.33. It is interesting how the change of variable formula (i.e. integration by substitution) for integrals on the real line emerges in terms of 1-forms. Consider a 1-form $\alpha = f(\tau)d\tau$ and its line integral over an interval $\int_{[a,b]} \alpha = \int_a^b f(\tau)d\tau$. Now "change coordinates" on $[a, b]$ so $\tau = g(t)$ with $g : [k, l] \to [a, b]$ and $g' > 0$. Since $d\tau = g'(t)dt$ we have $\alpha = f(g(t))g'(t)dt$ and its line integral $\int_{[k,l]} \alpha = \int_k^l f(g(t))g'(t)dt$. Thus, the behaviour of integrals under change of variables is incorporated into the definition of a 1-form.

3.5 Exercises

Exercise 3.34. Define two vector fields on $\mathbb{R}^3$ by

$$u = x^1\frac{\partial}{\partial x^1} + x^1x^2\frac{\partial}{\partial x^2} - \frac{\partial}{\partial x^3} \qquad \text{and} \qquad v = -x^2\frac{\partial}{\partial x^1} + (x^2)^3\frac{\partial}{\partial x^2}$$

and smooth functions on $\mathbb{R}^3$ by $f(x) = x^1x^2x^3$ and $g(x) = (x^2)^2$.

1. Compute the vector fields fu, gv , $u - v$, $fu + gv$.
2. Compute the functions $u(f)$, $u(g)$, $u(fg)$. Check explicitly the product rule for $u(fg)$.

Solution.

1. $$fu = (x^1)^2x^2x^3\frac{\partial}{\partial x^1} + (x^1)^2(x^2)^2x^3\frac{\partial}{\partial x^2} - x^1x^2x^3\frac{\partial}{\partial x^3}$$
   $$u - v = (x^1 + x^2)\frac{\partial}{\partial x^1} + (x^1x^2 - (x^2)^3)\frac{\partial}{\partial x^2} - \frac{\partial}{\partial x^3}.$$
   and so on.

2.

$$\begin{aligned} u(f) &= \left(x^1 \frac{\partial}{\partial x^1} + x^1 x^2 \frac{\partial}{\partial x^2} - \frac{\partial}{\partial x^3}\right)(x^1 x^2 x^3) \\ &= x^1 x^2 x^3 + (x^1)^2 x^2 x^3 - x^1 x^2 = x^1 x^2 (x^3 + x^1 x^3 - 1) \end{aligned}$$

and so on.

◆

Exercise 3.35. Let $u = \sum u^k \frac{\partial}{\partial x^k}, v = \sum v^k \frac{\partial}{\partial x^k}$ be two vector fields on (a domain in) $\mathbb{R}^n$. Show that there is a unique vector field z such that for all functions f we have

$$z(f) = u(v(f)) - v(u(f)) \tag{6}$$

and find an expression for the components of z in terms of the components u^k, v^k and their derivatives $\frac{\partial u^j}{\partial x^k}, \frac{\partial v^j}{\partial x^k}$. The vector field z is called the **Lie bracket** of u and v. It is usually written as $[u, v]$ and it is very important in advanced differential geometry.

Compute the Lie bracket of the vector fields in Exercise 3.34.

Show that for all vector fields u, v, w and functions f we have

- $[u, v] = -[v, u]$
- $[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0$ (The "Jacobi identity")
- $[fu, v] = f[u, v] - v(f)u$
- $[u, fv] = f[u, v] + u(f)v$

The best way to do these is to work from "abstract" properties like equation (6) above rather than from the coordinate formula, and show that the two sides agree applied to an arbitrary function.

Solution. Computing, for any function f we have

$$\begin{aligned} u(v(f)) &= u\left(\sum_k v^k \frac{\partial}{\partial x^k}\right) \\ &= \sum_j u^j \frac{\partial}{\partial x^j}\left(\sum_k v^k \frac{\partial f}{\partial x^k}\right) \\ &= \sum_j \sum_k u^j \frac{\partial}{\partial x^j}\left(v^k \frac{\partial f}{\partial x^k}\right) \\ &= \sum_j \sum_k \left(u^j v^k \frac{\partial^2 f}{\partial x^j x^k} + u^j \frac{\partial v^k}{\partial x^j}\frac{\partial f}{\partial x^k}\right) \\ &= \sum_j \sum_k u^j v^k \frac{\partial^2 f}{\partial x^j x^k} + \sum_k u[v^k]\frac{\partial f}{\partial x^k} \end{aligned}$$

Now, swapping u and v and subtracting the first terms cancel and the second combine to give

$$\sum_k (u(v^k) - v(u^k))\frac{\partial f}{\partial x^k}$$

and so

$$z = [u, v] = \sum_k (u(v^k) - v(u^k)) \frac{\partial}{\partial x^k}.$$

Applying this to

$$u = x^1 \frac{\partial}{\partial x^1} + x^1 x^2 \frac{\partial}{\partial x^2} - \frac{\partial}{\partial x^3}, \quad v = -x^2 \frac{\partial}{\partial x^1} + (x^2)^3 \frac{\partial}{\partial x^2}$$

we get

$$[u, v] = (-x^1 x^2 + x^2) \frac{\partial}{\partial x^1} + (2x^1(x^2)^3 + (x^2)^2) \frac{\partial}{\partial x^2}.$$

- For all f we have $[u, v](f) = u(v(f)) - v(u(f)) = -(v(u(f)) - u(v(f))) = -[v, u](f)$. Thus $[u, v] = -[v, u]$.
- We have

  $$\begin{aligned} [u, [v, w]](f) &= u([v, w](f)) - [v, w](u(f)) \\ &= u(v(w(f)) - w(v(f))) - v(w(u(f]] + w(v(u(f]]] \\ &= u(v(w(f))) - u(w(v(f))) - v(w(u(f))) + w(v(u(f))) \end{aligned}$$

  Now write this down twice more with u, v, w cyclically permuted and add: everything cancels!
- For all functions g we have

  $$\begin{aligned} &[fu, v](g) = fu(v(g)) - v(fu(g)) \\ &\qquad = f(u(v(g)) - fv(u(g)) - v(f)u(g) = (f[u, v])(g) - (v(f)u)(g). \end{aligned}$$

- Very similar.

◆

Exercise 3.36. In $\mathbb{R}^3$, consider the vector fields

$$u = 2x^1 \frac{\partial}{\partial x^1} - x^3 \frac{\partial}{\partial x^2} + \frac{\partial}{\partial x^3}, \quad v = (x^2)^2 \frac{\partial}{\partial x^1} + (x^3)^2 \frac{\partial}{\partial x^2} - \frac{\partial}{\partial x^3}$$

and the two 1-forms

$$\alpha = (x^1)^2 dx^1 + dx^2 - 3x^1 x^2 x^3 dx^3, \quad \beta = -x^1 dx^1 + 2dx^2 - 3(x^1)^2 dx^3.$$

Evaluate $\alpha(u), \alpha(v), \beta(u), \beta(v)$.

Solution. To give just one example,

$$\alpha(u) = (x^1)^2(2x^1) + 1(-x^3) - 3x^1 x^2 x^3(1) = 2(x^1)^3 - x^3 - 3x^1 x^2 x^3.$$

◆

Exercise 3.37. Consider the functions f, g on $\mathbb{R}^3$ given by

$$f(x^1, x^2, x^3) = x^1 x^2 (x^3)^2 \qquad \text{and} \qquad g(x^1, x^2, x^3) = (x^2)^4.$$

Calculate df and dg. Check that $d(fg) = gdf + fdg$. Let u be the vector field in Exercise 3.36. Check explicitly that $(df)(u) = u(f)$.

Solution. We have

$$df = x^2(x^3)^2dx^1 + x^1(x^3)^2dx^2 + 2x^1x^2x^3dx^3, \quad dg = 4(x^2)^3dx^2.$$

The function fg is given by $x^1(x^2)^5(x^3)^2$ and so

$$d(fg) = (x^2)^5(x^3)^2dx^1 + 5x^1(x^2)^4(x^3)^2dx^2 + 2x^1(x^2)^4x^3dx^3.$$

On the other hand,

$$\begin{aligned} gdf + fdg &= (x^2)^4(x^2(x^3)^2dx^1 + x^1(x^3)^2dx^2 + 2x^1x^2x^3dx^3) + x^1x^2(x^3)^2(4(x^2)^3dx^2) \\ &= (x^2)^5(x^3)^2dx^1 + 5x^1(x^2)^4(x^3)^2dx^2 + 2x^1(x^2)^4x^3dx^3. \end{aligned}$$

Of course, the fact that $df(u) = u(f)$ is the *definition* of df, but we can check this holds for our formulae:

$$\begin{aligned} df(u) &= (x^2(x^3)^2dx^1 + x^1(x^3)^2dx^2 + 2x^1x^2x^3dx^3)\left(2x^1\frac{\partial}{\partial x^1} - x^3\frac{\partial}{\partial x^2} + \frac{\partial}{\partial x^3}\right) \\ &= 2x^1x^2(x^3)^2 - x^1(x^3)^3 + 2x^1x^2x^3\ . \end{aligned}$$

On the other hand,

$$u(f) = \left(2x^1\frac{\partial}{\partial x^1} - x^3\frac{\partial}{\partial x^2} + \frac{\partial}{\partial x^3}\right)x^1x^2(x^3)^2 = 2x^1x^2(x^3)^2 - x^1(x^3)^3 + 2x^1x^2x^3\ .$$

◆

Exercise 3.38. Show that $d(f^2) = 2fdf$ for any function f on D. Show in general that if $g : \mathbb{R} \to \mathbb{R}$ is a smooth function then

$$d(g(f(x)) = g'(f(x))df.$$

(We are abusing notation here, because $g(f(x))$ is not a function, but its value. We usually write df and *not* $d(f(x))$. The correct way of writing this result is $d(g \circ f) = (g' \circ f)df$.)

Solution. Firstly,

$$d(f^2) = \sum \frac{\partial f^2}{\partial x^k}dx^k = \sum 2f\frac{\partial f}{\partial x^k}dx^k = 2fdf.$$

In general,

$$d(g(f(x)) = \sum \frac{\partial g(f(x))}{\partial x^k}dx^k = \sum g'(f(x))\frac{\partial f}{\partial x^k}dx^k = g'(f(x))df.$$

You can also do these without calculation by translating them into results about $v(f)$. For example,

$$d(f^2)(v) = v(f^2) = 2fv(f) = 2fdf(v).$$

This holds for all v so $d(f^2) = 2fdf$. ◆

Exercise 3.39. Let $f(x^1, x^2) = 0$ define a curve in $\mathbb{R}^2$. Suppose that $t \mapsto x(t)$ is a regular parametrised curve whose image is the curve $f(x^1, x^2) = 0$. Show that $df(x') = 0$ on the curve. (Work from abstract properties: use the definition of d so that $df(x') = x'(f)$.)

Solution. $f(x(t))$ is identically zero along the curve. Hence the directional derivative of f in the direction given by the tangent vector $x'(t)$ must be zero since f is not changing in that direction, in other words $x'(f) = 0$ for all t. But $x'(f) = df(x')$ and so the result follows. Thus, df is serving as a sort of "normal" to the curve $f = 0$, but note we are not assuming or using any inner product space structure on $\mathbb{E}^2$. ◆

Exercise 3.40. Consider the parametrised graph of a function f:

$$c(t) = (t, f(t)) \qquad t \in [a, b]$$

as a curve in the plane. Show that $\int_c x^2 dx^1$ is just the usual integral $\int_a^b f(x^1) dx^1$.

Solution. Parametrising as given,

$$(x^1)'(t) = 1, \quad (x^2)'(t) = f'(t).$$

so the field of tangent vectors to the curve is $c'(t) = \frac{\partial}{\partial x^1} + f'(t)\frac{\partial}{\partial x^2}$. Thus,

$$x^2 dx^1(c') = x^2$$

and so

$$\int_c x^2 dx^1 = \int_a^b f(t) dt.$$

◆

Intermezzo: the Einstein summation convention and Ricci calculus

Indices

Now that we have introduced tangent spaces, cotangent spaces, vector fields and one-forms, it is perhaps time to comment on our convention to have the indices for coordinates as superscripts (like $(x^1, \ldots, x^n)$), unlike the more conventional subscripts (like $(x_1, \ldots, x_n)$).

In fact, this is all part of a very convenient notational scheme (essentially for linear algebra), part of which was introduced by no other than Albert Einstein.

The starting point for this is the following observation: suppose that I have a vector space V, and dual vector space V^*. A choice of a basis $\mathbf{e_1}, \ldots, \mathbf{e_n}$ for V induces a dual basis $\mathbf{f^1}, \ldots, \mathbf{f^n}$ for V^*, determined by the rule $\mathbf{f}^i(\mathbf{e}_j) = \delta^i_j$, where δ^i_j is the Kronecker delta.

Any element $\mathbf{a}$ of V or $\mathbf{b}$ of V^*can be written as a linear combination of these basis vectors:

$$\mathbf{a} = \sum_i a^i \mathbf{e}_i \text{ and } \mathbf{b} = \sum_j b_j \mathbf{f}^j,$$

and we have $\mathbf{b}(\mathbf{a}) = \sum_i a^i b_i$.

Now one can make the following observation:

> ❝ If we do a change of basis for V, which induces a change of basis for V^*, then the coefficients of a vector in V transform in the same way as the basis vectors of V^* and vice versa, the coefficients of a vector in V^* transform in the same way as the basis vectors for V. ❞

Indeed, suppose a new basis for V is given by $\widetilde{\mathbf{e}}_\mathbf{1}, \ldots, \widetilde{\mathbf{e}}_\mathbf{n}$, with

$$\widetilde{\mathbf{v}}_i = \sum_j h^j_i \mathbf{v_j},$$

where the h^j_i are the coefficients of the invertible change-of basis matrix, and g^k_l are the coefficients of its inverse (i.e. $\sum_j h^j_i g^l_j = \delta^l_i$). If we then denote the new induced dual basis for V^* by $\widetilde{\mathbf{f}}^\mathbf{1}, \ldots, \widetilde{\mathbf{f}}^\mathbf{n}$, we have

$$\widetilde{\mathbf{f}}^i = \sum_j g^i_j \mathbf{f^j}.$$

Moreover, for any elements $\mathbf{a}$ of V and $\mathbf{b}$ of V which we can write as

$$\mathbf{e} = \sum_i a^i \mathbf{e_i} = \sum_i \widetilde{a}^i \widetilde{\mathbf{e}}_i \text{ and } \mathbf{b} = \sum_j b_j \mathbf{f}^i = \sum_j \widetilde{b}_j \widetilde{\mathbf{f}}^j$$

we have

$$\widetilde{a}^i = \sum_j g^i_j a^j \text{ and } \widetilde{b}_i = \sum_j h^j_i b_j.$$

Physicists often use the terminology of *co-variance* and *contra-variance*: the entities b_i and $\mathbf{e_i}$ are co-variant, and the entities a^i and $\mathbf{f^i}$ are contra-variant. Moreover, physicists often don't bother to use the basis vectors, and only use the coefficients. So for them a vector field is the assignment of a set of smooth functions $(v_1(x^1,\dots,x^n),\dots,v_n(x^1,\dots,x^n))$ to each coordinate system $(x^1,\dots,x^n)$, such that these functions transform in the contra-variant way when you switch coordinates. They also sometimes refer to one-forms as *co-vectors*, because their coefficients transform in a co-variant way.

Now the first part of this notational convention (sometimes referred to as *Ricci calculus* keeps track of which entities are co-variant and which are contra-variant: those with indices as sub-scripts (such as the basis vectors $\mathbf{e}_i$ for V and the coefficients b_i for V^*) are co-variant, and those with indices as super-scripts (such as the basis vectors $\mathbf{f}^i$ for V^* and the coefficients a^i for V) are contra-variant.

What does this have to do with the indices for coordinates? Well, we have seen that if you choose coordinates $(x^1,\dots,x^n)$ on an open set $D \subset \mathbb{R}^n$ containing a point p, then you get a basis for the tangent space T_pD at p

$$\left\{ \frac{\partial}{\partial x^1}\Big|_p, \dots, \frac{\partial}{\partial x^n}\Big|_p \right\},$$

and so the "superscript indices in the denominator" are understood to be lower indices, and indicate a co-variant entity (remark that p is not an index here, but just denotes the point we are at). Similary we get a basis for the cotangent space T_p^*D at p

$$\left\{ dx^1|_p, \dots, dx^n|_p \right\},$$

and again the indices are super-script, so indicate a contra-variant entity.

Now, since the dual of the dual $(V^*)^*$ of a finite-dimensional vector space is canonically isomorphic to the original vector space $V \cong (V^*)^*$, one can object to this and say it is the story of the chicken and the egg: why not swap the roles of co-variance and contra-variance? The reason why we fix this notation in differential geometry is that the co-variant entities transform like the coordinates do: if $(\tilde{x}^1,\dots,\tilde{x}^n)$ are new coordinates, giving rise to new basis

$$\left\{ \frac{\partial}{\partial \tilde{x}^1}\Big|_p, \dots, \frac{\partial}{\partial \tilde{x}^n}\Big|_p \right\}$$

for T_pD, then by the chain rule the change of basis is given by

$$\frac{\partial}{\partial \tilde{x}}\Big|_p = \sum_j \frac{\partial x^j}{\partial \tilde{x}^i}(p)\,\frac{\partial}{\partial x^j}\Big|_p,$$

where we think of the old coordinates $(x^1,\dots,x^n)$ as functions of the new coordinates $(\tilde{x}^1,\dots,\tilde{x}^n)$. Similarly, we have

$$d\widetilde{x}^k|_p = \sum_l \frac{\partial \widetilde{x}^k}{\partial x^l}(p)\, dx^l|_p.$$

Summations

Now there is another part to this notational scheme: the Einstein summation convenction. Since we very often have to let elements of T_p^*D act on elements of T_pD, or write a vector expressed with respect to a basis, this involves a lot of summation signs over the relevant indices. The final part of the Einstein summation convention therefor says to stop writing these summation signs, and stick to the following:

> ❝ Any time an index in a formula occurs twice, once as sub-script and once as super-script, it is implicitly assumed that we sum over all instances of that index. ❞

Moreover, an unspecified index (such as u^i, as opposed to a particular instance, such as u^3) can only occur either once or twice per term in a formula.

If it occurs twice in a term, it should occur once as a lower index and once as an upper index – we refer to those indices as *dummy indices* or *summation indices.* They are always summed over the whole range of the index, so the index does not have a meaning outside the formula. E.e. when we write $v^i \frac{\partial}{\partial x^i}$ on a surface, this really means $v^1 \frac{\partial}{\partial x^1} + \frac{\partial v^2}{\partial x^2}$, and the index is 'all used up'.

If it occurs once we refer to it as a *free index*, and those have a meaning outside the formula. So for instance when we write $e^i f_j g^j$ on a surface, i is a free index, and j is a dummy index. This really means $e^i f_1 g^1 + e^i f_2 g^2$, and the whole expression only depends on i.

Remark that this only applies to indices that are unspecified, so we are still allowed to write $e^3 f^3 g_4 h^3$.

So from now on we would just write things like

$$\mathbf{a} = a^i \mathbf{e}_i \text{ or } \widetilde{\mathbf{v}}_i = h_i^j \mathbf{v}_j$$

etc, where the summation over the repeated index is implicitly assumed. We will continue with this notation even when talking about more general things than vector fields or 1-forms, such as differential forms in the next section, where we use multi-indices.

Example: polar coordinates

Let's put this to work in an example: carthesian versus polar coordinates on the Euclidean plane $\mathbb{E}^2$. So our first set of coordinates are $x^1 = x, x^2 = y$ – the usual Carthesian coordinates. Our second set of coordinates are the polar coordinates, $\widetilde{x}^1 = r$ and $\widetilde{x}^2 = \theta$. Note that these polar coordinates only offer a coordinate system on the open subset D of $\mathbb{E}^2$ obtained by removing the negative x-axis[3], i.e. we take $r \in (0, \infty)$ and $\theta \in (-\pi, \pi)$. So from now on we will just work on D.

Each of these coordinates can be written as a function of the others:

$$x = r\cos(\theta) \quad \text{and} \quad y = r\sin(\theta)$$

and vice-versa

$$r = \sqrt{x^2 + y^2} \quad \text{and} \quad \theta = \begin{cases} \arctan\left(\frac{y}{x}\right) & \text{if } x > 0 \\ \arctan\left(\frac{y}{x}\right) + \pi & \text{if} x < 0 \\ \frac{\pi}{2} & \text{if } x = 0 \text{ and } y > 0 \\ -\frac{\pi}{2} & \text{if } x = 0 \text{ and } y < 0. \end{cases}$$

So, suppose we have a point $p \in D$, with Carthesian coordinates $(x(p), y(p))$ and polar coordinates $(r(p), \theta(p))$. Any tangent vector $\mathbf{v} \in T_pD$ can be written as (remember the

[3] One could also choose to remove another closed half-line.

Einstein summation convention)

$$\mathbf{v} = v^i \frac{\partial}{\partial x^i}\Big|_p \quad \text{and} \quad \mathbf{v} = \widetilde{v}^j \frac{\partial}{\partial \widetilde{x}^j}\Big|_p$$

and similarly any one form at p, $\alpha \in T_p^*D$ can be written as

$$\alpha = \alpha_k dx^k|_p \quad \text{and} \quad \alpha = \widetilde{\alpha}_l d\widetilde{x}^l|_p.$$

Now, how are the coordinates v^i and $\widetilde{v}^j$ related to each other, and similarly what about α_k and $\widetilde{\alpha}_l$?

To find out, we begin by writing down the change of basis formulas:

$$\frac{\partial}{\partial \widetilde{x}^k}\Big|_p = \frac{\partial x^i}{\partial \widetilde{x}^k}(p)\frac{\partial}{\partial x^i}\Big|_p \quad \text{and} \quad d\widetilde{x}^k|_p = \frac{\partial \widetilde{x}^k}{\partial x^l}(p)dx^l|_p,$$

which become here

$$\begin{aligned}\frac{\partial}{\partial r}\Big|_p = \frac{\partial x}{\partial r}(p)\frac{\partial}{\partial x}\Big|_p + \frac{\partial y}{\partial r}(p)\frac{\partial}{\partial y}\Big|_p &= \cos(\theta(p))\frac{\partial}{\partial x}\Big|_p + \sin(\theta(p))\frac{\partial}{\partial y}\Big|_p \\ &= \frac{x}{\sqrt{x^2+y^2}}(p)\frac{\partial}{\partial x}\Big|_p + \frac{y}{\sqrt{x^2+y^2}}(p)\frac{\partial}{\partial y}\Big|_p\end{aligned}$$

and

$$\begin{aligned}\frac{\partial}{\partial \theta}\Big|_p = \frac{\partial x}{\partial \theta}(p)\frac{\partial}{\partial x}\Big|_p + \frac{\partial y}{\partial \theta}\frac{\partial}{\partial y}\Big|_p &= -r(p)\sin(\theta(p))\frac{\partial}{\partial x}\Big|_p + r(p)\cos(\theta(p))\frac{\partial}{\partial y}\Big|_p \\ &= -y(p)\frac{\partial}{\partial x}\Big|_p + x(p)\frac{\partial}{\partial y}\Big|_p\end{aligned}$$

for the bases for T_pD. Likewise we have

$$dr|_p = \frac{\partial r}{\partial x}(p)dx|_p + \frac{\partial r}{\partial y}(p)dy|_p = \frac{x}{\sqrt{x^2+y^2}}(p)\,dx|_p + \frac{y}{\sqrt{x^2+y^2}}(p)\,dy|_p$$

and

$$d\theta|_p = \frac{\partial \theta}{\partial x}(p)\,dx|_p + \frac{\partial \theta}{\partial y}(p)\,dy|_p = \frac{-y}{x^2+y^2}(p)\,dx|_p + \frac{x}{x^2+y^2}(p)\,dy|_p$$

for the bases for T_p^*D.

Now, from this we also figure out how the coefficients of $\mathbf{v}$ and α transform: the v^i transform like the $dx^i|_p$, and the α_i transform like the $\frac{\partial}{\partial x^i}\Big|_p$. So we get

$$\widetilde{v}^1 = \frac{\partial r}{\partial x}(p)\,v^1 + \frac{\partial r}{\partial y}(p)\,v^2 = \frac{x}{\sqrt{x^2+y^2}}(p)\,v^1 + \frac{y}{\sqrt{x^2+y}}(p)\,v^2$$

and

$$\widetilde{v}^2 = \frac{\partial \theta}{\partial x}(p)v^1 + \frac{\partial \theta}{\partial y}(p)\,v^2 = \frac{-y}{x^2+y^2}(p)\,v^1 + \frac{x}{x^2+y^2}(p)\,v^2,$$

and similarly

$$\widetilde{\alpha}_1 = \frac{\partial x}{\partial r}(p)\,\alpha_1 + \frac{\partial y}{\partial r}(p)\,\alpha_2 = \frac{x}{\sqrt{x^2+y^2}}(p)\,\alpha_1 + \frac{y}{\sqrt{x^2+y^2}}(p)\,\alpha_2$$

and

$$\widetilde{\alpha}_2 = \frac{\partial x}{\partial \theta}(p)\,\alpha_1 + \frac{\partial y}{\partial \theta}(p)\,\alpha_2 = -y(p)\alpha_1 + x(p)\alpha_2.$$

Note that these transformation rules also work for 2 forms here (and in general p-forms). In particular we get

$$\begin{aligned}
dr \wedge d\theta &= \left(\frac{\partial r}{\partial x}dx + \frac{\partial r}{\partial y}dy\right) \wedge \left(\frac{\partial \theta}{\partial x}dx + \frac{\partial \theta}{\partial y}dy\right) \\
&= \left(\frac{x}{\sqrt{x^2+y^2}}dx + \frac{y}{\sqrt{x^2+y^2}}dy\right) \wedge \left(\frac{-y}{x^2+y^2}dx + \frac{x}{x^2+y^2}dy\right) \\
&= \frac{x^2+y^2}{(x^2+y^2)^{\frac{3}{2}}}dx \wedge dy \\
&= \frac{1}{\sqrt{x^2+y^2}}dx \wedge dy
\end{aligned}$$

or equivalently

$$dx \wedge dy = r\,dr \wedge d\theta.$$

Section 4: Differential forms

In this section we will construct an "exterior algebra", whose elements are called differential forms on a domain $D \subseteq \mathbb{R}^n$. There will be two operations in this algebra: a type of multiplication ($\wedge$) and differentiation (d) satisfying a number of important identities.

4.1 The exterior algebra of differential forms

Definition 4.1. *A 2-form at $p \in D$ is a map $\alpha : T_pD \times T_pD \to \mathbb{R}$ which is linear in each argument and alternating $\alpha(v, w) = -\alpha(w, v)$ for all $v, w \in T_pD$. More generally, a k-form at $p \in D$ is a map of k vectors in T_pD to $\mathbb{R}$ which is* **multilinear** *(linear in each argument) and* **alternating** *(changes sign under a swap of any two arguments).*

Definition 4.2. *The* **wedge product** *$\alpha \wedge \beta$ of 1-forms α and β is a 2-form defined by the following bilinear and alternating map,*

$$(\alpha \wedge \beta)(v, w) = \alpha(v)\beta(w) - \alpha(w)\beta(v) = \begin{vmatrix} \alpha(v) & \alpha(w) \\ \beta(v) & \beta(w) \end{vmatrix}.$$

More generally, the wedge product of k 1-forms $\alpha_1, \alpha_2, \ldots, \alpha_k$ can be defined as a map acting on k vectors $v_1, v_2, \ldots, v_k$

$$(\alpha^1 \wedge \cdots \wedge \alpha^k)(v_1, \ldots, v_k) = \begin{vmatrix} \alpha^1(v_1) & \ldots & \alpha^1(v_k) \\ \vdots & \ddots & \vdots \\ \alpha^k(v_1) & \ldots & \alpha^k(v_k) \end{vmatrix}.$$

From the properties of the determinant it follows that the resulting map is linear in each vector separately and changes sign if any pair of vectors is exchanged (this corresponds to exchanging two columns in the determinant). Hence it defines a k-form.

Now consider the wedge product of the basis 1-forms dx^i. We have

$$\boxed{dx^i \wedge dx^j = -dx^j \wedge dx^i \qquad \text{for all } i, j.} \tag{7}$$

Notice in particular that $dx^i \wedge dx^i = 0$ for all i. Using equation (7) we can reorder any wedge product $dx^i \wedge dx^j \wedge \cdots \wedge dx^k$ (up to signs) in such a way that $i < j < \cdots < k$. In $D \subseteq \mathbb{R}^n$ we have the following (together with their number in the first column):

$$\begin{array}{lll} n & dx^i & 1 \le i \le n \\ \binom{n}{2} & dx^i \wedge dx^j & 1 \le i < j \le n \\ \binom{n}{3} & dx^i \wedge dx^j \wedge dx^k & 1 \le i < j < k \le n \\ & \vdots & \\ 1 & dx^1 \wedge dx^2 \wedge \cdots \wedge dx^n & \end{array}$$

Example 4.3. In $\mathbb{E}^3$, let

$$\varepsilon = dx \wedge dy \wedge dz.$$

If $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ are vector fields in $\mathbb{E}^3$ then

$$\varepsilon(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = \boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w})$$

which is the volume of the parallelepiped spanned by the three vectors. (In fact, ε is often called the "standard volume form on $\mathbb{E}^3$".)

Example 4.4. In $\mathbb{E}^3$ consider the 2-form $\beta = dx \wedge dy$. Then given two vectors $\boldsymbol{u}, \boldsymbol{v}$

$$\beta(\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{k} \cdot (\boldsymbol{u} \times \boldsymbol{v})$$

where $\boldsymbol{k}$ is the usual unit vector in the z direction. From this one sees that $\beta(\boldsymbol{u}, \boldsymbol{v})$ measures the projection of the (signed) area of the parallelogram spanned by $\boldsymbol{u}, \boldsymbol{v}$ onto the (x, y)-plane.

To avoid a proliferation of indices, it is convenient to introduce a more compact notation.

Definition 4.5. *By a* **multi-index** I **of length** $|I| = k$ *we shall mean an increasing sequence* $I = (i_1, \ldots, i_k)$ *of integers* $1 \le i_1 < i_2 < \cdots < i_k \le n$. *We will write*

$$dx^I = dx^{i_1} \wedge \cdots \wedge dx^{i_k}.$$

Example 4.6. For example, if $I = (1, 2, 4)$ in $\mathbb{R}^4$, then

$$dx^I = dx^1 \wedge dx^2 \wedge dx^4.$$

Remark 4.7. The set of k-forms at p is a vector space of dimension $\binom{n}{k}$ for $0 \le k \le n$ with basis

$$\left\{ dx^I|_p : |I| = k \right\}.$$

Definition 4.8. *A* **differential** k**-form** *or a* **differential form of degree** k *on* D *is a smooth map* α *which assigns to each* $p \in D$ *a* k*-form at* p*; it can be written as*

$$\alpha = \alpha_I dx^I$$

where $\alpha_I : D \to \mathbb{R}$ *are smooth functions, and the sum (remember the Einstein summation convention!) happens over all multi-indices* I *with* $|I| = k$. *Given two differential* k*-forms* α, β *and a function* f *the differential* k*-forms* $\alpha + \beta$ *and* $f\alpha$ *are*

$$\alpha + \beta = (\alpha_I + \beta_I) dx^I \ , \qquad f\alpha = f\alpha_I dx^I \ .$$

The set of k*-forms on* D *is denoted* $\Omega^k(D)$. *By convention, a* **zero-form** *is a function. If* $k > n$ *then* $\Omega^k(D) = \varnothing$ *(every form has a repeated index).*

Example 4.9.

$$2x^1x^2dx^1 \wedge dx^2 - (x^1)^3 dx^1 \wedge dx^3 + x^1 dx^2 \wedge dx^3$$

is a 2-form in $\mathbb{R}^3$.

Example 4.10. In $\mathbb{R}^3$ we have the following:

General 0-form: a function f

General 1-form: $\alpha = \alpha_1 dx^1 + \alpha_2 dx^2 + \alpha_3 dx^3$

General 2-form: $\beta = \beta_{12} dx^1 \wedge dx^2 + \beta_{13} dx^1 \wedge dx^3 + \beta_{23} dx^2 \wedge dx^3$

General 3-form: $\gamma \, dx^1 \wedge dx^2 \wedge dx^3$

(We wrote γ rather than γ_{123} for simplicity since a 3-form in $\mathbb{R}^3$ has a single component.) In general, an n-form in $\mathbb{R}^n$ will have a single component, and a form of higher degree than that is necessarily zero. For $0 \leq k \leq n$ a k-form in $\mathbb{R}^n$ has $\binom{n}{k}$ components.

Definition 4.11. *We extend $\wedge$ linearly in order to define the* **wedge product** *of a k-form α and an ℓ-form β. Explicitly,*

$$\left(\alpha_I dx^I\right) \wedge \left(\beta_J dx^J\right) = \alpha_I \beta_J dx^I \wedge dx^J \ .$$

Here the sum is happening over all multi-indices I and J with $|I| = k$ and $|J| = \ell$ Now two things can happen: either $I \cap J \neq \varnothing$, in which case $dx^I \wedge dx^J = 0$ since there will be a repeated index, or else $I \cap J = \varnothing$ in which case $dx^I \wedge dx^J = \pm dx^K$, for some multi-index K of length $|K| = k+\ell$. The sign is due to having to reorder them to be increasing. Therefore, the wedge product defines a (bilinear) map

$$\wedge : \Omega^k(D) \times \Omega^\ell(D) \to \Omega^{k+\ell}(D)$$

Example 4.12. For example,

$$\begin{aligned}(x^1dx^2 - dx^3) \wedge ((x^1)^2 dx^1 \wedge dx^2 + x^3 dx^1 \wedge dx^3) &= (x^1)^3 dx^2 \wedge dx^1 \wedge dx^2 \\ &+ x^1x^3 dx^2 \wedge dx^1 \wedge dx^3 - (x^1)^2 dx^3 \wedge dx^1 \wedge dx^2 - x^3 dx^3 \wedge dx^1 \wedge dx^3.\end{aligned}$$

In the above, every term with a repeated dx^i is zero as a consequence of the defining relations for $\wedge$, so this reduces to

$$x^1x^3 dx^2 \wedge dx^1 \wedge dx^3 - (x^1)^2 dx^3 \wedge dx^1 \wedge dx^2.$$

Finally, we can do one swap of the dx^i in the first term (at the expense of a minus sign) and two in the second (no net effect) to put both terms in the same order to get

$$-x^1(x^1 + x^3) dx^1 \wedge dx^2 \wedge dx^3.$$

Proposition 4.13. *For all $\alpha \in \Omega^k(D)$ and $\beta \in \Omega^\ell(D)$,*

$$\boxed{\alpha \wedge \beta = (-1)^{k\ell} \beta \wedge \alpha \ .} \tag{8}$$

Note that if $f \in \Omega^0(D)$ is a function, then we define

$$f \wedge \beta = f\beta$$

and the formula still applies.

Proof. Because of linearity of the wedge product, it is enough to check this on the basis $dx^I \wedge dx^J$, for which the only nontrivial case is when $I \cap J = \varnothing$. Consider then

$$dx^{i_1} \wedge \cdots \wedge dx^{i_k} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_\ell} \ .$$

Moving dx^{j_1} to the left, we get a sign every time it passes through a dx^{i_m}. There are k of them, whence we arrive at

$$(-1)^k dx^{j_1} \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} \wedge dx^{j_2} \wedge \cdots \wedge dx^{j_\ell} \ .$$

We now move dx^{j_2} to the left and get another $(-1)^k$ after passing through all the dx^{i_m}'s, to arrive at

$$(-1)^{2k} dx^{j_1} \wedge dx^{j_2} \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} \wedge dx^{j_3} \wedge \cdots \wedge dx^{j_\ell} \ .$$

After passing all ℓ dx^{j_p} through the dx^{i_m}, we arrive at

$$(-1)^{k\ell} dx^{j_1} \wedge \cdots \wedge dx^{j_\ell} \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} \ ,$$

which proves the result. □

4.2 The exterior derivative

We defined the exterior derivative $df \in \Omega^1(D)$ of a function $f \in \Omega^0(D)$. This defines a linear map $d : \Omega^0(D) \to \Omega^1(D)$, which we will now extend to forms of any degree.

Definition 4.14. *If $\alpha = \alpha_I dx^I \in \Omega^k(D)$, then its* **exterior derivative** *$d\alpha \in \Omega^{k+1}(D)$ is*

$$d\alpha = d\alpha_I \wedge dx^I$$

where $d\alpha_I$ denotes the exterior derivative of the function α_I (which we defined earlier). Note that a consequence of this definition is $d(dx^I) = d(1) \wedge dx^I = 0$.

Example 4.15. Letting $\alpha = x^1 x^2 dx^1 + x^3 dx^2 - dx^3 \in \Omega^1(\mathbb{R}^3)$ we have

$$\begin{aligned} d\alpha &= d(x^1 x^2) \wedge dx^1 + dx^3 \wedge dx^2 \\ &= (x^1 dx^2 + x^2 dx^1) \wedge dx^1 + dx^3 \wedge dx^2 \\ &= -x^1 dx^1 \wedge dx^2 - dx^2 \wedge dx^3 \end{aligned}$$

in standard form. Notice further that $d(d\alpha) = 0$.

Theorem 4.16. *The exterior derivative $d : \Omega^k(D) \to \Omega^{k+1}(D)$ is a linear map satisfying the following properties:*

1. *d obeys the graded derivation property, for any $\alpha \in \Omega^k(D)$*

$$\boxed{d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^k \alpha \wedge d\beta} \qquad (9)$$

2. *$d(d\alpha) = 0$ for any $\alpha \in \Omega^k(D)$, or more compactly $d^2 = 0$.*

Proof. Linearity of d is clear from (4.14): $d(\alpha + \beta) = d\alpha + d\beta$ and $d(\lambda\alpha) = \lambda d\alpha$, $\lambda \in \mathbb{R}$. Next

$$\begin{aligned} d(\alpha \wedge \beta) &= d(\alpha_I \beta_J) \wedge dx^I \wedge dx^J \\ &= \beta_J d\alpha_I \wedge dx^I \wedge dx^J + \alpha_I d\beta_J \wedge dx^I \wedge dx^J \\ &= d\alpha \wedge \beta + (-1)^k \alpha_I dx^I \wedge d\beta^J \wedge dx^J = d\alpha \wedge \beta + (-1)^k \alpha \wedge d\beta \end{aligned}$$

where in the last line we have used $d\beta_J \wedge dx^I = (-1)^k dx^I \wedge d\beta_J$ which follows from (8) and the fact that β_J is a function (0-form). Finally, note that by using linearity and the graded derivation property $d(d\alpha) = d(d\alpha_I \wedge dx^I) = d(d\alpha_I) \wedge dx^I$, where we have also used $d(dx^I) = 0$. Hence it is enough to prove $d(df) = 0$ for any function f. We have

$$\begin{aligned} d(df) &= d\left(\frac{\partial f}{\partial x^j}\right) \wedge dx^j = \frac{\partial^2 f}{\partial x^i \partial x^j} dx^i \wedge dx^j \\ &= \left(\frac{\partial^2 f}{\partial x^i \partial x^j} - \frac{\partial^2 f}{\partial x^j \partial x^i}\right) dx^i \wedge dx^j = 0 \ . \end{aligned}$$

We have rewritten the last line using (7) and relabelled i and j appropriately. The final equality follows from commutativity of the second partial derivatives of f. □

Remark 4.17. The sign $(-1)^k$ in the derivation property is necessary for compatibility with the sign in (8).

Definition 4.18. *A form $\alpha \in \Omega^k(D)$ is said to be* **closed** *if $d\alpha = 0$ and it is said to be* **exact** *if $\alpha = d\beta$ for some $\beta \in \Omega^{k-1}(D)$.*

Remark 4.19. Notice that every exact form is closed (since $d^2 = 0$). However, the converse need not be true – it depends on the topology of D!

Example 4.20. Consider example 3.28. Explicit computation shows that $d\omega = 0$, so ω is a closed form on D. However ω is *not* an exact form on D. To see this suppose that it is exact: then there exists a smooth function f such that $\omega = df$, so $\int_c \omega = \int_c df = 0$ for any closed curve c in D. However, in example 3.28 we showed $\int_c \omega = 2\pi$ if c is a circle centred at the origin. It follows that ω can not be exact, as claimed.

Notice in the above example our domain contains a "hole" (the origin). It is a result of great importance that if D is a region in $\mathbb{R}^n$ containing no "holes", then every closed form is exact. We will state the result without proof for $D = \mathbb{R}^n$.

Theorem 4.21 (Poincaré Lemma)**.** *Every closed differential form on $\mathbb{R}^n$ is exact.*

Remark 4.22. The Poincaré Lemma generalises to any *contractible* domain $D \subseteq \mathbb{R}^n$. Intuitively a contractible domain is one that can be continuously shrunk to a point. (This is made precise in topology). The domain $D = \mathbb{R}^n - \{0\}$ is not contractible, and as we saw in the example (4.20) closed forms are not always exact on such a domain.

4.3 Integration in $\mathbb{R}^n$

We saw that 1-forms integrate over curves (or domains in $\mathbb{R}$). Similarly, we now show that differential n-forms naturally integrate over (domains in) $\mathbb{R}^n$. First, observe that any nowhere-vanishing differential n-form on $\mathbb{R}^n$ can be written as

$$f(x^1, \ldots, x^n) dx^1 \wedge \cdots \wedge dx^n \ ,$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a nowhere-vanishing smooth function. Hence either $f > 0$ everywhere or $f < 0$ everywhere. This defines two equivalence classes of nowhere-vanishing n-forms on $\mathbb{R}^n$. These equivalence classes are the two possible **orientations** of $\mathbb{R}^n$.

Definition 4.23. *The* **standard orientation** *(which we always assume) is defined by*

$$dx^1 \wedge dx^2 \wedge \cdots \wedge dx^n \ .$$

Coordinates $(y^1, \ldots, y^n)$ *(an ordered set!) are said to be* **oriented** *on* D *iff* $dy^1 \wedge \cdots \wedge dy^n$ *is a positive multiple of* $dx^1 \wedge \cdots \wedge dx^n$ *for all* $x \in D \subseteq \mathbb{R}^n$*.*

Example 4.24. The usual polar coordinates (r, θ) are oriented coordinates for the plane (away from the origin). Explicitly, we have $x = r\cos\theta$ and $y = r\sin\theta$ and therefore

$$dx = -r\sin\theta d\theta + \cos\theta dr, \qquad dy = r\cos\theta d\theta + \sin\theta dr.$$

Hence

$$dx \wedge dy = rdr \wedge d\theta$$

so (r, θ) are oriented coordinates for $r > 0$, i.e. away from the origin in polar coordinates.

Proposition 4.25. *Let* $(x^1, \ldots, x^n)$ *be oriented coordinates for* $\mathbb{R}^n$*. Let* $(y^1, \ldots, y^n)$ *be smooth functions on* $\mathbb{R}^n$*. Then*

$$dy^1 \wedge \cdots \wedge dy^n = \frac{\partial(y^1, \ldots, y^n)}{\partial(x^1, \ldots, x^n)}\, dx^1 \wedge \cdots \wedge dx^n,$$

where the factor on the RHS is the Jacobian of the coordinate transformation (i.e. the determinant of the matrix whose ij *component is* $\frac{\partial y^i}{\partial x^j}$*.)*

Proof. The case $n = 2$ is as follows. We have $dy^i = \frac{\partial y^i}{\partial x^j}dx^j = \frac{\partial y^i}{\partial x^1}dx^1 + \frac{\partial y^i}{\partial x^2}dx^2$ and therefore

$$dy^1 \wedge dy^2 = \left(\frac{\partial y^1}{\partial x^1}\frac{\partial y^2}{\partial x^2} - \frac{\partial y^1}{\partial x^2}\frac{\partial y^2}{\partial x^1}\right) dx^1 \wedge dx^2 = \det\begin{pmatrix} \frac{\partial y^1}{\partial x^1} & \frac{\partial y^1}{\partial x^2} \\ \frac{\partial y^2}{\partial x^1} & \frac{\partial y^2}{\partial x^2} \end{pmatrix} dx^1 \wedge dx^2$$

$\square$

Remark 4.26. The functions $(y^1, y^2, \ldots, y^n)$ are oriented coordinates on some domain if and only if the Jacobian is positive there.

Definition 4.27. *Let* $(x^1, \ldots, x^n)$ *be oriented coordinates on* $D \subseteq \mathbb{R}^n$ *and write*

$$\omega = f(x^1, \ldots, x^n)dx^1 \wedge \cdots \wedge dx^n \in \Omega^n(D).$$

Then the **integral of** ω **over** D *is defined by*

$$\int_D \omega = \int_D f(x^1, \ldots, x^n)dx^1 \ldots dx^n \ ,$$

where the right-hand side is now the usual multi-integral of several variable calculus (provided it exists).

Proposition 4.28. *The integral of an* n*-form* ω *over* $D \subseteq \mathbb{R}^n$ *is independent of the choice of oriented coordinates (and changes sign if the opposite orientation is chosen).*

Proof. Recall the formula for change of variables for integration in several variable calculus,

$$\int f(x^1,\dots,x^n)dx^1\dots dx^n = \int f(x^1(y),\dots,x^n(y))\left|\frac{\partial(x^1,\dots,x^n)}{\partial(y^1,\dots,y^n)}\right| dy^1\dots dy^n$$

The LHS is $\int\omega$ evaluated in the x coordinates. If the Jacobian is positive (or negative), (4.25) shows the RHS is $\int\omega$ (or $-\int\omega$) evaluated in the y coordinates. □

Example 4.29. We may integrate $\alpha = dx\wedge dy$ over the unit disc $D=\{x^2+y^2\le 1\}$ in the plane using two different coordinate systems. Working directly in Cartesian coordinates

$$\int_D dx\wedge dy = \int_D 1dxdy = \pi$$

since it is simply the area of the unit disc. Now, convert α into polar coordinates and evaluate the integral using these coordinates. From above $\alpha = rdr\wedge d\theta$. So,

$$\int_D rdr\wedge d\theta = \int_{r=0}^{1}\int_{\theta=0}^{2\pi} rdrd\theta = \pi.$$

Remark 4.30. Later in the course we will learn that just like a 1-form can be integrated on an oriented curve, a 2-form can be integrated on an oriented surface.

4.4 Exercises

Exercise 4.31. Consider the following differential forms in $\mathbb{R}^3$:

$$\begin{aligned}\alpha &= 2x^1dx^1 + x^1x^2dx^2 + x^2dx^3\\ \beta &= x^3dx^1 + (x^1)^3dx^2 - dx^3,\\ \gamma &= x^1dx^1\wedge dx^2 - (x^2)^4dx^1\wedge dx^3 - dx^2\wedge dx^3\\ \delta &= -dx^1\wedge dx^2 + x^1x^2x^3dx^1\wedge dx^3 - (x^1)^3dx^2\wedge dx^3.\end{aligned}$$

Work out every possible wedge product of two of the forms, in each possible order. Express the results in standard form. Check in each case $\omega\wedge\phi = (-1)^{\deg(\omega)\deg(\phi)}\phi\wedge\omega$ is satisfied.

Solution. I won't do all of these, but for example

$$\begin{aligned}\alpha\wedge\gamma &= (2x^1dx^1 + x^1x^2dx^2 + x^2dx^3)\wedge(x^1dx^1\wedge dx^2 - (x^2)^4dx^1\wedge dx^3 - dx^2\wedge dx^3)\\ &= -2x^1dx^1\wedge dx^2\wedge dx^3 - x^1(x^2)^5dx^2\wedge dx^1\wedge dx^3 + x^1x^2dx^3\wedge dx^1\wedge dx^2\\ &= (-2x^1 + x^1(x^2)^5 + x^1x^2)dx^1\wedge dx^2\wedge dx^3\end{aligned}$$

in standard form. Notice that in multiplying out the brackets six of the nine terms involve a repeated dx^i and so are automatically zero without even writing them down. Notice also the sign change in the second term when collecting together, because an odd number (one, in fact) of swaps were needed to put the dx^k's in standard order. ◆

Exercise 4.32.☆ Let α,β be 1-forms in $\mathbb{R}^n$, with α vanishing nowhere. Show that if $\alpha\wedge\beta = 0$ then β is proportional to α; that is, there is a function f such that $\beta = f\alpha$.

Solution. This is a slightly tricky question. Let $\alpha = \alpha_i dx^i$ and similarly for β. Then the wedge product in standard form is given by

$$\alpha \wedge \beta = (\alpha_i \beta_j - \alpha_j \beta_i) dx^i \wedge dx^j .$$

For this to be zero then we require

$$\alpha_i \beta_j - \alpha_j \beta_i = 0, \text{for all } 1 \le i < j \le n. \tag{10}$$

Let us assume first of all that there is one component of α, say α_j, which is never zero. Rearranging we get

$$\beta_i = \frac{\beta_j}{\alpha_j} \alpha_i \text{ , for all } 1 \le i \le n,$$

and so

$$\beta = \frac{\beta_j}{\alpha_j} \alpha \ .$$

Now, in general, we are only given that not all components of α vanish together, hence we have to work a little harder. For $j = 1, \ldots, n$, let $U_j \subseteq \mathbb{R}^n$ denote the (open) subset where α_j is not zero. The hypothesis says that $\mathbb{R}^n = \bigcup_{j=1}^{n} U_j$. On each U_j we find, as we did above, that $\beta = f_j \alpha$, where $f_j = \frac{\beta_j}{\alpha_j}$. To prove that $\beta = f\alpha$ everywhere it remains to show that $f_i = f_j$ on on $U_i \cap U_j$, whence f_i is the restriction to U_i of some function f, but this is precisely the content of equation (10). ◆

Exercise 4.33. Let $\alpha = dx^1 \wedge dx^2 + dx^3 \wedge dx^4$ in $\mathbb{R}^4$. Calculate $\alpha \wedge \alpha$.

Solution. You should get $2dx^1 \wedge dx^2 \wedge dx^3 \wedge dx^4$, which is a reminder that unlike with 1-forms, a 2-form wedged with itself need not give zero. ◆

Exercise 4.34. Calculate the exterior derivative of each of the forms of Exercise 4.31 and express the result in standard form. Compute also $d(\alpha \wedge \beta)$ and verify that the formula giving this in terms of $d\alpha$ and $d\beta$ holds. Verify also that $dd\beta = 0$.

Solution. To calculate $d\alpha$, for example:

$$\begin{aligned} d\alpha &= d(2x^1 dx^1 + x^1 x^2 dx^2 + x^2 dx^3) \\ &= 2dx^1 \wedge dx^1 + d(x^1 x^2) \wedge dx^2 + dx^2 \wedge dx^3 \\ &= (x^1 dx^2 + x^2 dx^1) \wedge dx^2 + dx^2 \wedge dx^3 \\ &= x^2 dx^1 \wedge dx^2 + dx^2 \wedge dx^3 . \end{aligned}$$

◆

Exercise 4.35. Let f, g be functions and define the 1-form $\alpha = f dg$. Show that $\alpha \wedge d\alpha = 0$. Can the 1-form $dx^1 + x^2 dx^3$ in $\mathbb{R}^3$ be written in the form $f dg$?

Solution. Firstly,

$$d\alpha = d(f dg) = df \wedge dg + f ddg = df \wedge dg.$$

Note that we are applying the Leibniz rule here where f is a function, i.e. a zero form. The second term could be written as "$f \wedge ddg$", but one usually omits the wedge when dealing with zero forms. To continue,

$$\alpha \wedge d\alpha = f dg \wedge df \wedge dg = -f df \wedge dg \wedge dg = 0,$$

since any 1-form wedged with itself gives zero.

Consider now the 1-form $\beta = dx^1 + x^2 dx^3$. We have

$$\beta \wedge d\beta = (dx^1 + x^2 dx^3) \wedge (dx^2 \wedge dx^3) = dx^1 \wedge dx^2 \wedge dx^3.$$

Since this is not zero, the first part of the question proves that we can not have $\beta = f dg$ for any functions f, g. ◆

Exercise 4.36. Let $\alpha = x^2 dx^1 \wedge dx^3 - dx^2 \wedge dx^3$. Calculate $\alpha(u, v)$ where u, v are as in Exercise 3.34.

Solution. Just use $dx^k(v) = v^k$, etc, and the definition. ◆

Exercise 4.37. Let $\alpha = \alpha_k dx^k$ be a 1-form. Show that

$$d\alpha = \frac{\partial \alpha_j}{\partial x^i} dx^i \wedge dx^j \ .$$

Hence show that

$$d\alpha(u, v) = u(\alpha(v)) - v(\alpha(u)) - \alpha([u, v])$$

where $[u, v]$ is the Lie bracket as in Exercise 3.35. (There's no smart way to do this with the technology we've developed: you'll have to use the coordinate form for the Lie bracket.)

Solution. The expression for $d\alpha$ follows immediately from the definition of d. Now, $dx^i(v) = v^i$ and so from the definition

$$d\alpha(u, v) = \frac{\partial \alpha_j}{\partial x^i}(u^i v^j - u^j v^i).$$

Turning to the right-hand side,

$$\begin{aligned} u(\alpha(v)) &= u\left(\alpha_j v^j\right) \\ &= \frac{\partial(\alpha_j v^j)}{\partial x^i} \\ &= \left(u^i v^j \frac{\partial \alpha_j}{\partial x^i} + u^i \alpha_j \frac{\partial v^j}{\partial x^i}\right) \end{aligned}$$

Subtracting off the same with u and v swapped over, the first term of each combine to give $d\alpha(u, v)$. Thus it remains to evaluate $\alpha([u, v])$ and see that it gives the remaining terms. Now, from the Exercise on Lie derivative,

$$[u, v] = \left(u^i \frac{\partial v^j}{\partial x^i} - v^i \frac{\partial u^j}{\partial x^i}\right) \frac{\partial}{\partial x^j}$$

and so

$$\alpha([u, v]) = \left(u^i \frac{\partial v^j}{\partial x^i} - v^i \frac{\partial u^j}{\partial x^i}\right) \alpha_j$$

which really does give the required terms. ◆

Exercise 4.38. Check that cylindrical polar coordinates (ρ, θ, z) are oriented coordinates for $\mathbb{R}^3$ away from the x^3-axis. Recall these are defined by $x^1 = \rho \cos\theta$, $x^2 = \rho \sin\theta$, $x^3 = z$.

Solution.

$$dx^1 = \cos\theta d\rho - \rho\sin\theta d\theta \qquad dx^2 = \sin\theta d\rho + \rho\cos\theta d\theta \qquad dx^3 = dz \ ,$$

$$dx^1 \wedge dx^2 \wedge dx^3 = \rho d\rho \wedge d\theta \wedge dz$$

which shows that they are oriented for $\rho > 0$, hence away from the x^3-axis. Indeed, inverting this equation we find

$$d\rho \wedge d\theta \wedge dz = \frac{1}{\sqrt{(x^1)^2 + (x^2)^2}} dx^1 \wedge dx^2 \wedge dx^3 \ .$$

◆

Exercise 4.39.☆ This exercise sets up the basic notions of de Rham cohomology. Let $Z^k(D)$ and $B^k(D)$ denote the closed and exact closed k-forms on D, respectively. Show that $B^k(D) \subseteq Z^k(D) \subset \Omega^k(D)$ are vector subspaces. Show that the wedge product satisfies:

$$\wedge : Z^k(D) \times Z^\ell(D) \to Z^{k+\ell}(D) \qquad \text{and} \qquad \wedge : B^k(D) \times Z^\ell(D) \to B^{k+\ell}(D) \ .$$

In other words, the wedge product of two closed forms is closed, and that of a closed form with an exact form is exact. Define the **k-th de Rham cohomology of D** to be the quotient vector space

$$H^k(D) = Z^k(D)/B^k(D) \ .$$

In other words, $H^k(D)$ consists of equivalence classes of closed k-forms, where two such forms are equivalent if their difference is an exact form. Show that $H^0(D) \cong \mathbb{R}$ and consists of the constant functions. Prove that the wedge product of forms induces a well-defined product on cohomology:

$$\wedge : H^k(D) \times H^\ell(D) \to H^{k+\ell}(D) \ .$$

It is called the **cup product** and it makes $\bigoplus_{k=0}^n H^k(D)$ into a graded algebra.

Solution. Since $d : \Omega^k(D) \to \Omega^{k+1}(D)$ is a linear map, $Z^k(D) = \ker d : \Omega^k(D) \to \Omega^{k+1}(D)$ and $B^k(D) = \operatorname{im} d : \Omega^{k-1}(D) \to \Omega^k(D)$ are both subspaces of $\Omega^k(D)$. The fact that $d^2 = 0$ means that $B^k(D) \subseteq Z^k(D)$.

Now let $\alpha \in Z^k(D)$ and $\beta \in Z^\ell(D)$. Their wedge product $\alpha \wedge \beta \in \Omega^{k+\ell}(D)$ and, using equation (9), we see that

$$d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^k \alpha \wedge d\beta = 0 \ ,$$

whence $\alpha \wedge \beta \in Z^{k+\ell}(D)$. If now $\alpha = d\gamma$, then

$$\alpha \wedge \beta = d\gamma \wedge \beta = d(\gamma \wedge \beta) \ ,$$

whence $\alpha \wedge \beta \in B^{k+\ell}(D)$. It is clear that $B^0(D) = 0$ since there are no (-1)-forms, whereas $Z^0(D)$ consists of those functions f such that $df = 0$. Since D is connected, this means that f is constant. Therefore $H^0(D) = Z^0(D) \cong \mathbb{R}$.

Let $\alpha \in Z^k(D)$ and $\beta \in Z^\ell(D)$ define classes $[\alpha] \in H^k(D)$ and $[\beta] \in H^\ell(D)$. Since $\alpha \wedge \beta \in Z^{k+\ell}(D)$, it defines a class $[\alpha \wedge \beta] \in H^{k+\ell}(D)$. This suggests defining

$$[\alpha] \wedge [\beta] = [\alpha \wedge \beta] \ .$$

We need only check that it is well-defined, but this follows from the fact that the wedge product of a closed form with an exact form is exact. ◆

Exercise 4.40. Let r, θ, ϕ be spherical polar coordinates on $\mathbb{R}^3$. Compute $dr, d\theta, d\phi$ in terms of dx, dy, dz and *vice versa*. Hence show that the standard orientation

$$dx \wedge dy \wedge dz = r^2 \sin\theta dr \wedge d\theta \wedge d\phi \ .$$

Deduce (r, θ, ϕ) are oriented coordinates (away from the z-axis).

Solution. We have $x = r\sin\theta\cos\phi$ and so

$$dx = \sin\theta\cos\phi dr + r\cos\theta\cos\phi d\theta - r\sin\theta\sin\phi d\phi.$$

Do the same for y, z and then wedge everything together! ◆

Exercise 4.41. Check that

$$dy^1 \wedge \cdots \wedge dy^n = \det\left(\tfrac{\partial(y^1,\ldots,y^n)}{\partial(x^1,\ldots,x^n)}\right) dx^1 \wedge \cdots \wedge dx^n$$

in the case $n = 3$.

Solution. For $i = 1, 2, 3$ we have

$$dy^i = \frac{\partial y^i}{\partial x^1}dx^1 + \frac{\partial y^i}{\partial x^2}dx^2 + \frac{\partial y^i}{\partial x^3}dx^3$$

and so

$$\begin{aligned} dy^1 \wedge dy^2 &= \left(\frac{\partial y^1}{\partial x^1}\frac{\partial y^2}{\partial x^2} - \frac{\partial y^1}{\partial x^2}\frac{\partial y^2}{\partial x^1}\right) dx^1 \wedge dx^2 + \left(\frac{\partial y^1}{\partial x^1}\frac{\partial y^2}{\partial x^3} - \frac{\partial y^1}{\partial x^3}\frac{\partial y^2}{\partial x^1}\right) dx^1 \wedge dx^3 \\ &+ \left(\frac{\partial y^1}{\partial x^2}\frac{\partial y^2}{\partial x^3} - \frac{\partial y^1}{\partial x^3}\frac{\partial y^2}{\partial x^2}\right) dx^2 \wedge dx^3 \end{aligned}$$

and so wedging with $dy^3 = \frac{\partial y^3}{\partial x^1}dx^1 + \frac{\partial y^3}{\partial x^2}dx^2 + \frac{\partial y^3}{\partial x^3}dx^3$ gives the result. ◆

Section 5: Surfaces

In this section we will define the notion of a regular local surface and introduce some basic examples. Throughout this section, $D \subseteq \mathbb{R}^2$ is a domain in the plane whose Cartesian coordinates will be denoted (u^1, u^2), unless otherwise stated. We consider maps from D into $\mathbb{E}^3$, which we consider to be equipped with the standard orientation, so that we can take cross products of vectors. We will use coordinates (x^1, x^2, x^3) on $\mathbb{E}^3$.

5.1 Regular surfaces

Intuitively, a surface in Euclidean space is a subset which "locally looks like a subset of $\mathbb{R}^2$". This idea an be made precise as follows.

Definition 5.1. *A* **local surface** *in $\mathbb{E}^3$ is a smooth, injective, map $\boldsymbol{x} : D \to \mathbb{E}^3$ with a continuous inverse $\boldsymbol{x}^{-1} : \boldsymbol{x}(D) \to D$. Sometimes we denote the image $\boldsymbol{x}(D)$ by S.*

The assumption that $\boldsymbol{x}$ is injective means that points in the image $\boldsymbol{x}(D)$ are uniquely labelled by points in D. That the inverse map is continuous is a technical assumption to prevent "near self-intersection" (we will not discuss this further). A local surface is sometimes called a **parametrised surface**. The idea is that a surface is a subset $\Sigma \subset \mathbb{E}^3$ which can be 'covered' by local surfaces. First, though we need to impose a condition to ensure there are no singularities.

Definition 5.2. *Given a local surface we define*

$$\boldsymbol{x}_{u^1} = \frac{\partial x^i}{\partial u^1}\frac{\partial}{\partial x^i}, \quad \boldsymbol{x}_{u^2} = \frac{\partial x^i}{\partial u^2}\frac{\partial}{\partial x^i}.$$

For every point $p \in D$, these are vectors in $T_{\boldsymbol{x}(p)}\mathbb{E}^3$, which we will identify with $\mathbb{E}^3$ itself. We say that a local surface $\boldsymbol{x}$ is **regular at** *$p \in D$ if $\boldsymbol{x}_{u^1}(p)$ and $\boldsymbol{x}_{u^2}(p)$ are linearly independent. A local surface is* **regular** *if it is regular at p for all $p \in D$.*

Definition 5.3. *$\Sigma \subset \mathbb{E}^3$ is a* **regular surface** *if for each $\boldsymbol{p} \in \Sigma$ there exists a regular local surface $\boldsymbol{x} : D \to \mathbb{E}^3$ such that $\boldsymbol{p} \in \boldsymbol{x}(D)$ and $\boldsymbol{x}(D) = U \cap \Sigma$ for some open set $U \subset \mathbb{R}^3$.*

Remark 5.4. A map $\boldsymbol{x}$ that defines a local surface which is part of some surface Σ, is sometimes called a **coordinate chart** on Σ.

In this course we will mostly deal with local surfaces. We will assume all our surfaces are regular unless stated otherwise.

Definition 5.5. *At a regular point on a local surface, the plane spanned by $\boldsymbol{x}_{u^1}(p)$ and $\boldsymbol{x}_{u^1}(p)$ is the* **tangent plane to the surface at $\boldsymbol{x}(p)$**, *which we denote by $T_{\boldsymbol{x}(p)}S$. At a regular point, the* **unit normal** *to the surface is*

$$\boldsymbol{N}(p) = \frac{\boldsymbol{x}_{u^1}(p) \times \boldsymbol{x}_{u^2}(p)}{|\boldsymbol{x}_{u^1}(p) \times \boldsymbol{x}_{u^2}(p)|}$$

Clearly, $\boldsymbol{N}(p)$ is orthogonal to the tangent plane $T_{\boldsymbol{x}(p)}S$.

Definition 5.6. *Locally, a surface has two possible unit normals (one pointing "inwards" and one "outwards"). A choice of one of these is a choice of* **orientation** *for the surface. We will usually assume that our surfaces are oriented and that parametrisations are chosen so that* $\boldsymbol{N}$ *agrees with the chosen orientation. (This is always possible, since by swapping* u *and* v *we can change the sign of* $\boldsymbol{N}$ *if need be.) Surfaces for which it is not possible to make a 'global' choice of unit normal are said to be "non-orientable" (e.g. the Möbius band). We can still work locally on such a surfaces.*

Proposition 5.7. *Given a local surface the map* $\boldsymbol{N} : D \to \mathbb{E}^3$ *is a smooth function whose image lies in a unit sphere* $S^2 \subset \mathbb{E}^3$. *The map* $\boldsymbol{N}$ *is called the local* **Gauss map**.

Proof. A regular surface has a well-defined unit normal at each point which defines the map $\boldsymbol{N} : D \to \mathbb{E}^3$. By definition $\boldsymbol{N} \cdot \boldsymbol{N} = 1$, so the image lies on a unit sphere in $\mathbb{E}^3$. □

Example 5.8. The **cylinder** is the surface defined by

$$\boldsymbol{x}(\phi, z) = \begin{pmatrix} a\cos\phi \\ a\sin\phi \\ z \end{pmatrix} \qquad \text{with } a > 0.$$

We have

$$\boldsymbol{x}_\phi = \begin{pmatrix} -a\sin\phi \\ a\cos\phi \\ 0 \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{x}_z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} ,$$

whence

$$\boldsymbol{N} = \begin{pmatrix} \cos\phi \\ \sin\phi \\ 0 \end{pmatrix} .$$

Notice that the image of the Gauss map for a cylinder is a unit circle (which of course lies in a unit sphere).

Example 5.9. Consider the spherical polar coordinate parametrisation of the **sphere** of radius $a > 0$:

$$\boldsymbol{x}(\alpha, \phi) = \begin{pmatrix} a\sin\alpha\cos\phi \\ a\sin\alpha\sin\phi \\ a\cos\alpha \end{pmatrix} .$$

The unit normal is

$$\boldsymbol{N} = \begin{pmatrix} \sin\alpha\cos\phi \\ \sin\alpha\sin\phi \\ \cos\alpha \end{pmatrix}$$

wherever it is defined. The proof of this is left to exercise 5.20. Notice that the image of the Gauss map of a sphere is the whole of the unit sphere.

5.2 Standard surfaces

In this section we introduce various "standard" types of surface.

Definition 5.10. *Let $g : D \to \mathbb{R}$ be a smooth function. The* **graph of** *g is the local surface defined by*

$$\boldsymbol{x}(u^1, u^2) = \begin{pmatrix} u^1 \\ u^2 \\ g(u^1, u^2) \end{pmatrix} .$$

Proposition 5.11. *Graphs are always regular, and the unit normal is given by*

$$\boldsymbol{N} = \frac{1}{\sqrt{1 + \left(\frac{\partial g}{\partial u^1}\right)^2 + \left(\frac{\partial g}{\partial u^2}\right)^2}} \begin{pmatrix} -\frac{\partial g}{\partial u^1} \\ -\frac{\partial g}{\partial u^2} \\ 1 \end{pmatrix} .$$

Proof. Indeed,

$$\boldsymbol{x}_u = \begin{pmatrix} 1 \\ 0 \\ \frac{\partial g}{\partial u^1} \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{x}_v = \begin{pmatrix} 0 \\ 1 \\ \frac{\partial g}{\partial u^2} \end{pmatrix} ,$$

which are clearly linearly independent. One finds

$$\boldsymbol{x}_{u^1} \times \boldsymbol{x}_{u^2} = \begin{pmatrix} -\frac{\partial g}{\partial u^1} \\ -\frac{\partial g}{\partial u^2} \\ 1 \end{pmatrix} ,$$

which is normalised to give $\boldsymbol{N}$. □

Definition 5.12. *An* **implicitly defined surface** *Σ is the zero set of a (smooth) function $f : \mathbb{E}^3 \to \mathbb{R}$, i.e., $\Sigma = f^{-1}(0)$.*

Proposition 5.13. *An implicitly defined surface $\Sigma = f^{-1}(0)$, such that $df \neq 0$ everywhere on Σ, is a regular surface.*

Proof. For any $\boldsymbol{p} \in \Sigma$ we have $df|_{\boldsymbol{p}} \neq 0$. Thus, at least one partial derivative of f must be non-vanishing at $\boldsymbol{p}$, say $(\partial f/\partial x^3)|_{\boldsymbol{p}} \neq 0$. In that case, by the Implicit Function Theorem, for all $\boldsymbol{x}$ in an open set containing $\boldsymbol{p}$ such that $f(\boldsymbol{x}) = 0$, there exists a smooth function g such that $x^3 = g(x^1, x^2)$. Thus, near $\boldsymbol{p}$, the surface can be parametrised as the graph

$$\boldsymbol{x}(u^1, u^2) = \begin{pmatrix} u^1 \\ u^2 \\ g(u^1, u^2) \end{pmatrix}$$

By Proposition 5.11 this is a regular local surface. □

Example 5.14. The sphere of unit radius is the surface implicitly defined by

$$f(\boldsymbol{x}) = |\boldsymbol{x}| - 1 .$$

Computing, we find $df = (\boldsymbol{x}/|\boldsymbol{x}|) \cdot d\boldsymbol{x} = \boldsymbol{x} \cdot d\boldsymbol{x}$ on the sphere. This only vanishes at $\boldsymbol{x} = \boldsymbol{0}$, which is not on the sphere, hence the sphere is regular.

Definition 5.15. *A* **surface of revolution** *with profile curve $f(u)$ is a local surface of the form*

$$\boldsymbol{x}(u,\phi) = \begin{pmatrix} f(u)\cos\phi \\ f(u)\sin\phi \\ u \end{pmatrix}$$

A surface of revolution can be constructed by rotating a curve $x_1 = f(x_3)$ around the x_3-axis in $\mathbb{R}^3$. It thus has cylindrical symmetry.

Example 5.16. The cylinder in Example 5.8 is surface of revolution with profile function $f(u) = a$ and $u \in \mathbb{R}$.

Example 5.17. The sphere in 5.9 is surface of revolution with profile function $f(u) = \sqrt{1-u^2}$ where $u = \cos\alpha$. In this parameterisation it is regular only for $u \in (-1,1)$ (not at the endpoints); this is a coordinate artefact.

Definition 5.18. *A* **ruled surface** *is a surface of the form*

$$\boldsymbol{x}(u,v) = \boldsymbol{z}(u) + v\boldsymbol{p}(u).$$

Notice that curves of constant u are straight lines in $\mathbb{E}^3$ through $\boldsymbol{z}(u)$ in the direction $\boldsymbol{p}(u)$.

Example 5.19. The cylinder in Example 5.8 is an example of a ruled surface where

$$\boldsymbol{z}(u) = \begin{pmatrix} a\cos u \\ a\sin u \\ 0 \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{p}(u) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

On the other hand, the sphere is not a ruled surface.

5.3 Exercises

Exercise 5.20. Consider the spherical polar coordinate parametrisation of the sphere of radius a as in example 5.9. Compute the unit normal $\boldsymbol{N}$ and determine where the surface is regular.

Solution. You should get

$$\boldsymbol{x}_\alpha = \begin{pmatrix} a\cos\alpha\cos\phi \\ a\cos\alpha\sin\phi \\ -a\sin\alpha \end{pmatrix}, \quad \boldsymbol{x}_\phi = \begin{pmatrix} -a\sin\alpha\sin\phi \\ a\sin\alpha\cos\phi \\ 0 \end{pmatrix}.$$

Thus

$$\boldsymbol{x}_\alpha \times \boldsymbol{x}_\phi = \begin{pmatrix} a^2\sin^2\alpha\cos\phi \\ a^2\sin^2\alpha\sin\phi \\ a^2\sin\alpha\cos\alpha \end{pmatrix}.$$

This is zero iff $\sin\alpha = 0$, i.e. when $\alpha = 0, \pi$, which is at the North and South poles.

Away from the poles, $\boldsymbol{N}$ is the unit vector in this direction:

$$\boldsymbol{N} = \begin{pmatrix} \sin\alpha\cos\phi \\ \sin\alpha\sin\phi \\ \cos\alpha \end{pmatrix}.$$

In fact this is well defined even at the poles, hinting that there is nothing wrong with the surface at these points. This is of course the case, as the choice of North and South pole on a sphere is arbitrary, and here is an artefact of our choice of coordinates. Therefore, the Gauss map for a sphere of radius a is the whole unit sphere. ◆

Exercise 5.21. Show that

$$\boldsymbol{x}(z,\phi) = \begin{pmatrix} \sqrt{1+z^2}\cos\phi \\ \sqrt{1+z^2}\sin\phi \\ z \end{pmatrix}$$

parametrises the "hyperboloid of one sheet"

$$(x^1)^2 + (x^2)^2 - (x^3)^2 = 1.$$

Compute $\boldsymbol{N}$ and deduce that the parametrisation is regular. Sketch the z and ϕ coordinate curves on the hyperboloid.

Solution. Completely standard. Compute $\boldsymbol{x}_z \times \boldsymbol{z}_\phi$ and show that it is nowhere vanishing and then $\boldsymbol{N}$ is the unit vector in that direction. ◆

Exercise 5.22. Compute the unit normal $\boldsymbol{N}$ for the surface of revolution.

Solution. Standard. The answer is

$$\boldsymbol{N} = \frac{1}{\sqrt{1+(f')^2}} \begin{pmatrix} -\cos\phi \\ -\sin\phi \\ f'(u) \end{pmatrix}.$$

◆

Exercise 5.23. Find conditions on $\boldsymbol{z}(u)$ and $\boldsymbol{p}(u)$ that ensure that the ruled surface $\boldsymbol{x}(u,v) = \boldsymbol{z}(u) + v\boldsymbol{p}(u)$ is regular and compute the normal field $\boldsymbol{N}$.

Solution. We calculate

$$\boldsymbol{x}_u = \boldsymbol{z}'(u) + v\boldsymbol{p}'(u), \quad \boldsymbol{x}_v = \boldsymbol{p}(u).$$

Thus the surface is regular where

$$(\boldsymbol{z}'(u) + v\boldsymbol{p}'(u)) \times \boldsymbol{p}(u) \neq 0.$$

◆

Section 6: The fundamental forms

In this section we will introduce the first and second fundamental forms of a local surface.

6.1 Symmetric tensors

In this subsection $D \subseteq \mathbb{R}^n$ and $(x^1, x^2, \dots, x^n) \in D$ are coordinates.

Definition 6.1. *A symmetric bilinear form on T_pD is a bilinear map $B : T_pD \times T_pD \to \mathbb{R}$ such that $B(v, w) = B(w, v)$ for all $v, w \in T_pD$. Given two 1-forms α, β at $p \in D$ define a symmetric bilinear form $\alpha\beta$ on T_pD by*

$$(\alpha\beta)(v, w) = \tfrac{1}{2}\left(\alpha(v)\beta(w) + \alpha(w)\beta(v)\right)$$

where $v, w \in T_pD$. Note that $\alpha\beta = \beta\alpha$ and we write α^2 for $\alpha\alpha$.

Given the basis 1-forms dx_i, we can take the symmetric product of these according to the above rule to get the symmetric bilinear forms $dx^i dx^j$. Note that $dx^i dx^j = dx^j dx^i$.

Definition 6.2. *A symmetric tensor on D is a map which assigns to each $p \in D$ a symmetric bilinear form on T_pD; it can be written as*

$$B = B_{ij} dx^i dx^j$$

where B_{ij} are smooth functions on D.

Example 6.3. The usual dot product in $\mathbb{E}^n$ is given by the symmetric form on $\mathbb{R}^n$ (in orthonormal coordinates)

$$g = (dx^1)^2 + (dx^2)^2 + \cdots + (dx^n)^2 = \delta_{ij} dx^i dx^j.$$

Here

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

is the Kronecker delta. This is called the **Euclidean metric**. To see this, let $\boldsymbol{v} = v^i \frac{\partial}{\partial x_i}$, $\boldsymbol{w} = w^i \frac{\partial}{\partial x_i}$, so

$$g(v, w) = dx^i(\boldsymbol{v}) dx^i(\boldsymbol{w}) = \delta_{ij} v^i w^j = \boldsymbol{v} \cdot \boldsymbol{w}\ .$$

$\mathbb{E}^n$ can be defined as $\mathbb{R}^n$ equipped with the Euclidean metric.

Definition 6.4. *A (Riemannian)* **metric** *on D is a symmetric tensor $g = g_{ij} dx^i dx^j$ which is positive definite at each point; $g(v, v) \geq 0$ for all $v \in T_pD$, with equality if and only if $v = 0$. Equivalently, it is a choice for each $p \in D$ of an inner product on T_pD.*

6.2 The first fundamental form

We now return to regular local surfaces so $D \subseteq \mathbb{R}^2$ and $(u^1, u^2) \in D$ are our coordinates.

Proposition 6.5. *Consider a regular local surface defined by $\boldsymbol{x} : D \to \mathbb{E}^3$. The linear map $d\boldsymbol{x}|_p : T_pD \to T_{\boldsymbol{x}(p)}S$ is a bijection.*

Proof. Let w be a vector field on D,

$$w = a\,\frac{\partial}{\partial u^1} + b\,\frac{\partial}{\partial u^2}\ .$$

Also, we have the $\mathbb{E}^3$-valued 1-form on D:

$$d\boldsymbol{x} = \boldsymbol{x}_{u^1} du^1 + \boldsymbol{x}_{u^2} du^2\ .$$

Then,

$$d\boldsymbol{x}(w) = a\,\boldsymbol{x}_{u^1} + b\,\boldsymbol{x}_{u^2}.$$

Evaluating this at each $p \in D$ it is clear this is onto $T_{\boldsymbol{x}(p)}S$. Further, since the surface is regular $d\boldsymbol{x}(w) = 0$ implies $w = 0$, so the map is one-to-one. □

Remark 6.6. The above proposition can be used to give a coordinate free definition of regularity of a local surface.

Thus, $d\boldsymbol{x}$ is the "glue" that identifies tangent vector fields on D with vector fields in the tangent plane to the surface in Euclidean space (i.e linear combinations of $\boldsymbol{x}_u, \boldsymbol{x}_v$). This can be used to "pull-back" the Euclidean metric in $\mathbb{E}^3$ to a metric on D.

Definition 6.7. *Given a regular local surface $\boldsymbol{x} : D \to \mathbb{E}^3$, the* **first fundamental form** *is defined by*

$$I = d\boldsymbol{x} \cdot d\boldsymbol{x}\ ,$$

where we have introduced the notation $d\boldsymbol{x} \cdot d\boldsymbol{x} = (dx^1)^2 + (dx^2)^2 + (dx^3)^2$.

Proposition 6.8. *The first fundamental form is a metric on D.*

Proof. By definition it is a symmetric tensor. It is positive definite since

$$\mathrm{I}(w, w) = d\boldsymbol{x}(w) \cdot d\boldsymbol{x}(w) \geq 0$$

with equality iff $d\boldsymbol{x}(w) = 0$, and hence by proposition 6.5 the result follows. □

Proposition 6.9. *The first fundamental form of a local surface $\boldsymbol{x} : D \to \mathbb{E}^3$ is*

$$I = E(du^1)^2 + 2F du^1\, du^2 + G(du^2)^2$$

where E, F, G are functions on D given by

$$E = \boldsymbol{x}_{u^1} \cdot \boldsymbol{x}_{u^1}, \qquad F = \boldsymbol{x}_{u^1} \cdot \boldsymbol{x}_{u^2}, \qquad G = \boldsymbol{x}_{u^2} \cdot \boldsymbol{x}_{u^2}.$$

Proof. Simply note that $d\boldsymbol{x} = \boldsymbol{x}_{u^1} du^1 + \boldsymbol{x}_{u^2} du^2$ and expand out $d\boldsymbol{x} \cdot d\boldsymbol{x}$. □

Remark 6.10. Recall the general inequality $|\boldsymbol{x}_{u^1} \cdot \boldsymbol{x}_{u^2}| \leq |\boldsymbol{x}_{u^1}||\boldsymbol{x}_{u^2}|$ is saturated if and only if $\boldsymbol{x}_{u^1}$ and $\boldsymbol{x}_{u^2}$ are linearly dependent. For a regular surface this inequality is thus $EG - F^2 > 0$. This condition states that the matrix of I in the basis $(\frac{\partial}{\partial u^1}, \frac{\partial}{\partial u^2})$ is positive definite.

Remark 6.11. I is sometimes called the induced metric on the surface.

Example 6.12. Consider the cylinder in Example 5.8. We find

$$E = \boldsymbol{x}_z \cdot \boldsymbol{x}_z = 1 \qquad F = \boldsymbol{x}_z \cdot \boldsymbol{x}_\phi = 0 \qquad G = \boldsymbol{x}_\phi \cdot \boldsymbol{x}_\phi = a^2 \ ,$$

whence the first fundamental form is

$$\mathrm{I} = dz^2 + a^2 d\phi^2 \ .$$

Example 6.13. Consider the sphere in 5.9. A computation reveals (exercise!)

$$\mathrm{I} = a^2(d\alpha)^2 + a^2 \sin^2\alpha (d\phi)^2 \ .$$

6.3 The second fundamental form

The $\mathbb{E}^3$-valued 1-form $d\boldsymbol{N}$ is a linear map which which may have a non-trivial kernel. It is convenient to use the isomorphism $d\boldsymbol{x}$ to rewrite the map $d\boldsymbol{N}$ as a symmetric form.

Definition 6.14. *Given a local surface $\boldsymbol{x} : D \to \mathbb{E}^3$, the* **second fundamental form** *is defined by*

$$I\!I = -d\boldsymbol{x} \cdot d\boldsymbol{N},$$

with the dot product interpreted as above.

Remark 6.15. The second fundamental form is not necessarily positive (or negative) definite. This is because the linear map $d\boldsymbol{N}$ can have a non-trivial kernel.

Proposition 6.16. *The second fundamental form is given by*

$$I\!I = l(du^1)^2 + 2m du^1\, du^2 + n(du^2)^2$$

where l, m, n are functions on D given by

$$l = -\boldsymbol{x}_{u^1} \cdot \boldsymbol{N}_{u^1}, \quad m = -\boldsymbol{x}_{u^1} \cdot \boldsymbol{N}_{u^2} = -\boldsymbol{x}_{u^2} \cdot \boldsymbol{N}_{u^1}, \quad n = -\boldsymbol{x}_{u^2} \cdot \boldsymbol{N}_{u^2}.$$

These can be also written as

$$l = \boldsymbol{x}_{u^1u^1} \cdot \boldsymbol{N}, \qquad m = \boldsymbol{x}_{u^1u^2} \cdot \boldsymbol{N}, \qquad n = \boldsymbol{x}_{u^2u^2} \cdot \boldsymbol{N}.$$

Proof. We have

$$\begin{aligned} \mathrm{I\!I} &= -d\boldsymbol{x} \cdot d\boldsymbol{N} \\ &= -(\boldsymbol{x}_{u^1} du^1 + \boldsymbol{x}_{u^2} du^2) \cdot (\boldsymbol{N}_{u^1} du^1 + \boldsymbol{N}_{u^2} du^2) \end{aligned}$$

and so the first result follows by multiplying out. To establish the identity for l, differentiate $\boldsymbol{N} \cdot \boldsymbol{x}_{u^1} = 0$ to get

$$0 = \frac{\partial}{\partial u^1}(\boldsymbol{N} \cdot \boldsymbol{x}_{u^1}) = \boldsymbol{N}_{u^1} \cdot \boldsymbol{x}_{u^1} + \boldsymbol{N} \cdot \boldsymbol{x}_{u^1u^1}.$$

The other two are similar. $\square$

Remark 6.17. It may be shown that, for all vector fields y, z on D,

$$\mathbb{I}(y, z) = -d\boldsymbol{x}(y) \cdot d\boldsymbol{N}(z) \ .$$

This is how $\mathbb{I}$ is often defined, although this does not make it obvious that it is symmetric.

Example 6.18. For the cylinder,

$$\boldsymbol{N}_z = \boldsymbol{0} \qquad \text{and} \qquad \boldsymbol{N}_\phi = \begin{pmatrix} -\sin\phi \\ \cos\phi \\ 0 \end{pmatrix} \ ,$$

whence $l = 0$, $m = 0$ and $n = -\boldsymbol{x}_\phi \cdot \boldsymbol{N}_\phi = -a$. The second fundamental form of the cylinder is thus

$$\mathbb{I} = -a d\phi^2 \ .$$

Example 6.19. Consider again the sphere 5.9. A computation reveals

$$\mathbb{I} = -a(d\alpha)^2 - a\sin^2\alpha(d\phi)^2$$

which is left as an exercise.

Remark 6.20. Since $\boldsymbol{N}$ is unit normalised it follows that $\boldsymbol{N} \cdot \boldsymbol{N}_{u^1} = 0$ and $\boldsymbol{N} \cdot \boldsymbol{N}_{u^2} = 0$ (by differentiating $\boldsymbol{N} \cdot \boldsymbol{N} = 1$ by u^1 and u^2 respectively). Hence, $\boldsymbol{N}_{u^1}(p)$ and $\boldsymbol{N}_{u^2}(p)$ must belong to the tangent plane $T_{\boldsymbol{x}(p)}S$. In other words, $d\boldsymbol{N}|_p : T_pD \to T_{\boldsymbol{x}(p)}S$.

Remark 6.21. Later on we will see that the first fundamental form encodes information about sizes and is in some sense "intrinsic" to the surface, whereas the second fundamental form encodes information about the shape of the surface and is in some sense "extrinsic" to the surface and depends on the way it is embedded in space (for this reason $\mathbb{I}$ is sometimes called the extrinsic curvature).

6.4 Exercises

Exercise 6.22. Let $\alpha = du^1 - u^1du^2$ and $\beta = u^2du^1 + u^1du^2$. Calculate $\alpha\beta$. What is

$$(\alpha\beta)\left(2u^2\frac{\partial}{\partial u^1}, \frac{\partial}{\partial u^1} + v^2\frac{\partial}{\partial u^2}\right) .$$

Solution. We have

$$\alpha\beta = (du^1 - u^1du^2)(u^2du^1 + u^1du^2) = u^2(du^1)^2 + u^1(1-u^2)du^1du^2 - (u^1)^2(du^2)^2.$$

Letting $y = 2u^2\frac{\partial}{\partial u^1}$ and $z = \frac{\partial}{\partial u^1} + (u^2)^2\frac{\partial}{\partial u^1}$ we have

$$\begin{aligned}
(\alpha\beta)(y,z) &= \tfrac{1}{2}(\alpha(y)\beta(z) + \alpha(z)\beta(y)) \\
&= \tfrac{1}{2}((2u^2)(u^2 + u^1(u^2)^2) + (1 - u^1(u^2)^2)(2(u^2)^2)) \\
&= (u^2)^2(2 + u^1u^2 - u^1(u^2)^2)
\end{aligned}$$

◆

Exercise 6.23. Consider two vector fields

$$y = \frac{\partial}{\partial u^1} - \frac{\partial}{\partial u^2}, \quad z = \frac{\partial}{\partial u^1} + \frac{\partial}{\partial u^2}\ .$$

Evaluate

$$(du^1 du^2)(y, y), \quad (du^1 du^2)(y, z), \quad (du^1 du^2)(z, z).$$

Solution. For the first,

$$(du^1 du^2)(y, y) = \tfrac{1}{2}(du^1(y)du^2(y) + du^1(y)du^2(y)) = du^1(y)du^2(y) = -1.$$

For the second,

$$(du^1 du^2)(y, z) = \tfrac{1}{2}(du^1(y)du^2(z) + du^1(z)du^2(y)) = 0.$$

The last is similar to the first and the answer is 1. ◆

Exercise 6.24. Compute the first and second fundamental forms of the sphere of 5.9 from first principles and also using "E, F, G" and "l, m, n".

Solution. For the sphere,

$$d\boldsymbol{x} = \begin{pmatrix} a\cos\alpha\cos\phi d\alpha - a\sin\alpha\sin\phi d\phi \\ a\cos\alpha\sin\phi d\alpha + a\sin\alpha\cos\phi d\phi \\ -a\sin\alpha d\alpha \end{pmatrix}.$$

Thus, $\mathrm{I} = (dx_1)^2 + (dx_2)^2 + (dx_3)^2$ which gives

$$\mathrm{I} = a^2(d\alpha)^2 + a^2\sin^2\alpha(d\phi)^2.$$

You should get the same answer by computing $E = \boldsymbol{x}_\alpha \cdot \boldsymbol{x}_\alpha = a^2$ and similarly for F, G.

For the sphere,

$$\boldsymbol{x}_{\alpha\alpha} = \begin{pmatrix} -a\sin\alpha\cos\phi \\ -a\sin\alpha\sin\phi \\ -a\cos\alpha \end{pmatrix}, \quad \boldsymbol{x}_{\alpha\phi} = \begin{pmatrix} -a\cos\alpha\sin\phi \\ a\cos\alpha\cos\phi \\ 0 \end{pmatrix}, \quad \boldsymbol{x}_{\phi\phi} = \begin{pmatrix} -a\sin\alpha\cos\phi \\ -a\sin\alpha\sin\phi \\ 0 \end{pmatrix}$$

and

$$\boldsymbol{N} = \begin{pmatrix} \sin\alpha\cos\phi \\ \sin\alpha\sin\phi \\ \cos\alpha \end{pmatrix}.$$

So we get

$$l = \boldsymbol{x}_{\alpha\alpha} \cdot \boldsymbol{N} = -a, \quad m = \boldsymbol{x}_{\alpha\phi} \cdot \boldsymbol{N} = -a, \quad n = \boldsymbol{x}_{\phi\phi} \cdot \boldsymbol{N} = -a\sin^2\alpha$$

and so

$$\mathrm{II} = -a(d\alpha)^2 - a\sin^2\alpha(d\phi)^2.$$

One can also calculate II from first principles by calculating $d\boldsymbol{N}$ and then using the definition $\mathrm{II} = -d\boldsymbol{N} \cdot d\boldsymbol{x}$. In fact for the sphere $\boldsymbol{N} = a^{-1}\boldsymbol{x}$ and hence $\mathrm{II} = -a^{-1}d\boldsymbol{x} \cdot d\boldsymbol{x} = -a^{-1}\mathrm{I}$. ◆

Exercise 6.25. Show that the local surface defined by

$$\boldsymbol{x}(u^1,u^2) = \begin{pmatrix} \cosh u^1 \cos u^2 \\ \cosh u^1 \sin u^2 \\ \sinh u^1 \end{pmatrix}$$

is regular and parameterises the one-sheeted hyperboloid $x_1^2 + x_2^2 - x_3^2 = 1$. Compute the first and second fundamental forms of this local surface.

Solution. The fact that this local surface parametrises the one-sheeted hyperboloid is readily verified by checking that $x_1^2+x_2^2-x_3^2 = 1$ using various hyperbolic and trig identities. Now compute

$$\boldsymbol{x}_{u^1} = \begin{pmatrix} \sinh u^1 \cos u^2 \\ \sinh u^1 \sin u^2 \\ \cosh u^1 \end{pmatrix} \qquad \boldsymbol{x}_{u^2} = \begin{pmatrix} -\cosh u^1 \sin u^2 \\ \cosh u^1 \cos u^2 \\ 0 \end{pmatrix}$$

which are linearly independent and non-vanishing for all (u^1,u^2).

The coefficients of the first fundamental form are $E = |\boldsymbol{x}_{u^1}|^2 = \cosh^2 u^1 + \sinh^2 u^1 = 1 + 2\sinh^2 u^1$, $F = \boldsymbol{x}_{u^1} \cdot \boldsymbol{x}_{u^2} = 0$ and $G = |\boldsymbol{x}_{u^2}|^2 = \cosh^2 u^1$ so

$$\mathrm{I} = (1 + 2\sinh^2 u^1)(du^1)^2 + \cosh^2 u\,(du^2)^2\,.$$

For the second fundamental form we need the unit normal, which is this case is

$$\boldsymbol{N} = \frac{1}{\sqrt{1+2\sinh^2 u^1}} \begin{pmatrix} -\cosh u^1 \cos u^2 \\ -\cosh u^1 \sin u^2 \\ \sinh u^1 \end{pmatrix}\,.$$

Also we need

$$\boldsymbol{x}_{u^1u^1} = \begin{pmatrix} \cosh u^1 \cos u^2 \\ \cosh u^1 \sin u^2 \\ \sinh u^1 \end{pmatrix} \qquad \boldsymbol{x}_{u^1u^2} = \begin{pmatrix} -\sinh u^1 \sin u^2 \\ \sinh u^1 \cos u^2 \\ 0 \end{pmatrix} \qquad \boldsymbol{x}_{u^2u^2} = \begin{pmatrix} -\cosh u^1 \cos u^2 \\ -\cosh u^1 \sin u^2 \\ 0 \end{pmatrix}$$

and hence

$$l = \boldsymbol{x}_{u^1u^1}\cdot\boldsymbol{N} = -\frac{1}{\sqrt{1+2\sinh^2 u^1}} \qquad m = \boldsymbol{x}_{u^1u^2}\cdot\boldsymbol{N} = 0 \qquad n = \boldsymbol{x}_{u^2u^2}\cdot\boldsymbol{N} = \frac{\cosh^2 u^1}{\sqrt{1+2\sinh^2 u^1}}$$

◆

Section 7: Curvature of surfaces

Let $D \subseteq \mathbb{R}^2$ be a domain in the plane and $\boldsymbol{x} : D \to \mathbb{E}^3$ will be a regular local surface. In this section we will introduce the notion of curvature for such surfaces.

7.1 Aside on bilinear algebra

Let A, B be two symmetric bilinear forms on a real vector space V and $(e_1, \dots, e_n)$ a basis of V. The **eigenvalues of B with respect to A** are roots of the polynomial

$$\det(B - \lambda A) = 0 \ ,$$

where by abuse of notation A, B are the matrices of the symmetric forms A, B relative to the basis, i.e. the $n \times n$ matrices with entries $A_{ij} = A(e_i, e_j)$ and $B_{ij} = B(e_i, e_j)$. The eigenvectors with eigenvalue λ are vectors $v = v^i e_i$ whose matrices relative to the basis satisfy

$$(B - \lambda A) \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} = 0 \ .$$

Lemma 7.1. *The eigenvalues and eigenvectors of B with respect to A are independent of the basis chosen.*

Proof. To see this note that if P is a change of basis matrix, so $\det P \neq 0$, then the matrices A and B of the symmetric bilinear forms in the new basis are $P^T A P$ and $P^T B P$. Then

$$\det(P^T B P - \lambda P^T A P) = \det P^T \det(B - \lambda A) \det P = 0.$$

□

Remark 7.2. It is important to note that we are thinking about A and B as matrices representing a bilinear form V, not matrices representing a linear map from V to V. This matters when changing basis: if a matrix A represents a bilinear form with respect to a basis, and if the change of basis matrix is denoted by P, in the new basis the bilinear from is represented by $P^T A P$. If the same matrix A is used to denote a linear map from V to V in the old basis, this linear map will be represented by $P^{-1} A P$ in the new basis. If P is not orthogonal, $P^T A P$ and $P^{-1} A P$ will in general be different matrices!

Proposition 7.3. *If A is positive definite (i.e. an inner product) there exists a basis $\{e_1, \dots, e_n\}$ of V such that:*

1. *$\{e_1, \dots, e_n\}$ is orthonormal with respect to A,*

2. *each e_k is an eigenvector of B with respect to A with a real eigenvalue.*

Proof. Property (i) can be achieved by diagonalising A and then rescaling basis. Then, in an orthonormal basis for the inner product A, the matrix of A is just the identity. Then (ii) reduces to eigenvalues and eigenvectors of a real symmetric matrix B, so the result follows from a standard theorem of linear algebra. □

7.2 Gauss and mean curvatures

At each point $p \in D$ the first I and second II fundamental forms of a local surface

$$\boldsymbol{x} : D \to \mathbb{E}^3$$

are symmetric bilinear forms on T_pD. Furthermore I is positive-definite.

Definition 7.4. *The eigenvalues k_1, k_2 of II with respect to I are the* **principal curvatures** *of the surface. The corresponding eigenvectors are the* **principal directions** *of the surface. Hence the principal curvatures are the roots of the polynomial* $\det(\mathrm{II} - \lambda \mathrm{I}) = 0$. *Their mean is the* **mean curvature** *of the surface:*

$$H = \tfrac{1}{2}(k_1 + k_2) \ ,$$

and their product is the **Gauss curvature***:*

$$K = k_1 k_2 \ .$$

The principal curvatures may vary with position and so are (smooth) functions on D.

Proposition 7.5. *The Gauss and mean curvature are smooth functions on D given by*

$$K = \frac{ln - m^2}{EG - F^2}, \qquad H = \frac{lG + nE - 2mF}{2(EG - F^2)}$$

where as before E, F, G and l, m, n are the components of the first and second fundamental forms in local coordinates.

Proof. The matrix of the first fundamental form

$$\mathrm{I} = E(du)^2 + 2F du dv + G(dv)^2$$

in the basis $(\partial/\partial u, \partial/\partial v)$ of T_pD is

$$\begin{pmatrix} E & F \\ F & G \end{pmatrix}.$$

Similarly, the second fundamental form has matrix

$$\begin{pmatrix} l & m \\ m & n \end{pmatrix}.$$

The principal curvatures are thus the roots of

$$\det\left(\begin{pmatrix} l & m \\ m & n \end{pmatrix} - \lambda \begin{pmatrix} E & F \\ F & G \end{pmatrix}\right) = 0.$$

Expanded this becomes the quadratic equation

$$(EG - F^2)\lambda^2 - (lG + nE - 2mF)\lambda + (ln - m^2) = 0.$$

As $EG - F^2 > 0$ everywhere on D (since I is positive definite), we must be able to write the quadratic as $(\lambda - k_1)(\lambda - k_2) = \lambda^2 - 2H\lambda + K$, where k_1, k_2 are the principal curvatures and the equality follows from using the definition of H and K. Upon comparing to the actual quadratic allows us to read off the desired expressions for H, K. □

Remark 7.6. Note that the various curvatures are independent of the choice of parametrisation of the surface. This follows from the fact that a change of parametrisation is just a change of coordinates on D, and our definitions are all in terms of the "coordinate independent" calculus of forms. One can of course explicitly calculate the effect of change of parametrisation on the fundamental forms and hence prove this invariance, but it is not a very enlightening calculation.

Remark 7.7. Notice the elegant basis independent expressions

$$K = \frac{\det \mathrm{II}}{\det \mathrm{I}}, \qquad H = \frac{1}{2}\mathrm{Tr}(\mathrm{I}^{-1}\mathrm{II})$$

Thus, the Gauss curvature is positive if and only if II is positive definite.

Example 7.8. Recall, for the cylinder $\mathrm{I} = dz^2 + a^2 d\phi^2$ and $\mathrm{II} = -a d\phi^2$. Therefore, the principal curvatures are the roots of

$$\begin{vmatrix} -\lambda & 0 \\ 0 & -a - \lambda a^2 \end{vmatrix} = \lambda(a + \lambda a^2) = 0$$

Hence the principal curvatures are $k_1 = 0$ and $k_2 = -\frac{1}{a}$, with corresponding principal directions $\frac{\partial}{\partial z}$ and $\frac{\partial}{\partial \phi}$ respectively. It follows that $H = -\frac{1}{2a}$ and $K = 0$.

Example 7.9. For the sphere of radius a, parameterised as in Exercise 5.20, we found that the first and second fundamental forms are

$$\mathrm{I} = a^2(d\alpha^2 + \sin^2\alpha\, d\phi^2) \qquad \mathrm{II} = -a(d\alpha^2 + \sin^2\alpha\, d\phi^2).$$

Therefore $\mathrm{I} = -a\mathrm{II}$. Hence the only eigenvalue is $\lambda = -1/a$ and thus the principle curvatures are $k_1 = k_2 = -1/a$ and every vector is a principal direction. It follows that $H = -1/a$ and $K = 1/a^2$. Notice that as you increase the radius the curvatures decrease.

7.3 Exercises

Exercise 7.10. Consider bilinear forms A, B on a 2-dimensional vector space with matrices A, B respectively with respect to some given basis, where

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix}$$

Calculate the eigenvalues and eigenvectors of B with respect to A. Check that A is an inner product and verify that the eigenvectors are orthogonal with respect to A.

Solution. The characteristic equation is

$$0 = \begin{vmatrix} -1-\lambda & -1-\lambda \\ -1-\lambda & 1-2\lambda \end{vmatrix} = \lambda^2 - \lambda - 2$$

and so the eigenvalues are $\lambda = 2, -1$.

For $\lambda = 2$ to get the eigenvectors, solve

$$\begin{pmatrix} -3 & -3 \\ -3 & -3 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 0$$

to get

$$\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

similarly for $\lambda = -1$ an eigenvector is

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

The bilinear form determined by A is positive definite (recall it is sufficient that $a_{11} > 0$ and $\det A > 0$). To check the orthogonality of the eigenvectors, observe that

$$\begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = 0.$$

◆

Exercise 7.11. Compute the principal, Gauss and mean curvatures of the one-sheeted hyperboloid in the parametrisation defined in exercise 6.25.

Solution. In the basis $(\frac{\partial}{\partial u^1}, \frac{\partial}{\partial u^2})$ the first and second fundamental form are represented by the matrices

$$\mathrm{I} = \begin{pmatrix} 1 + 2\sinh^2 u & 0 \\ 0 & \cosh^2 u^1 \end{pmatrix} \qquad \mathrm{II} = \frac{1}{\sqrt{1 + 2\sinh^2 u^1}} \begin{pmatrix} -1 & 0 \\ 0 & \cosh^2 u^1 \end{pmatrix}$$

so the eigenvalues of II with respect to I are given by

$$\det \begin{pmatrix} -\frac{1}{\sqrt{1+2\sinh^2 u^1}} - \lambda(1 + 2\sinh^2 u^1) & 0 \\ 0 & \cosh^2 u^1 \left(\frac{1}{\sqrt{1+2\sinh^2 u^1}} - \lambda \right) \end{pmatrix} = 0$$

and hence the principal curvatures are

$$k_1 = -\frac{1}{(1 + 2\sinh^2 u^1)^{3/2}} \qquad k_2 = \frac{1}{\sqrt{1 + 2\sinh^2 u^1}}$$

with corresponding principal directions $\frac{\partial}{\partial u^1}$ and $\frac{\partial}{\partial u^2}$ respectively. Hence the Gauss and mean curvatures are

$$K = -\frac{1}{(1 + 2\sinh^2 u^1)^2} \qquad H = \frac{\sinh^2 u^1}{(1 + 2\sinh^2 u^1)^{3/2}}.$$

◆

Exercise 7.12. Calculate the mean and Gauss curvatures of a local surface defined by the graph of a function $x_3 = f(x_1, x_2)$. Deduce the partial differential equation that the function f must satisfy for such a surface to have vanishing mean curvature.

Solution. Parametrise as

$$\boldsymbol{x}(u^1,u^2) = \begin{pmatrix} u^1 \\ u^2 \\ f(u^1,u^2) \end{pmatrix}$$

Computing,

$$\boldsymbol{x}_{u^1} = \begin{pmatrix} 1 \\ 0 \\ f_{u^1} \end{pmatrix}, \quad \boldsymbol{x}_{u^2} = \begin{pmatrix} 0 \\ 1 \\ f_{u^2} \end{pmatrix}.$$

Also,

$$\boldsymbol{N} = \frac{\boldsymbol{x}_{u^1} \times \boldsymbol{x}_{u^2}}{|\boldsymbol{x}_{u^1} \times \boldsymbol{x}_{u^2}|} = \frac{1}{\sqrt{1+f_{u^1}^2+f_{u^2}^2}} \begin{pmatrix} -f_{u^1} \\ -f_{u^2} \\ 1 \end{pmatrix}.$$

So,

$$E = \boldsymbol{x}_{u^1} \cdot \boldsymbol{x}_{u^2} = 1 + f_{u^1}^2, \quad F = \boldsymbol{x}_{u^1} \cdot \boldsymbol{x}_{u^2} = f_{u^1} f_{u^2}, \quad G = \boldsymbol{x}_{u^1} \cdot \boldsymbol{x}_{u^2} = 1 + f_{u^2}^2.$$

Differentiating again,

$$\boldsymbol{x}_{u^1u^1} = \begin{pmatrix} 0 \\ 0 \\ f_{u^1u^1} \end{pmatrix}, \quad \boldsymbol{x}_{u^2u^2} = \begin{pmatrix} 0 \\ 0 \\ f_{u^2u^2} \end{pmatrix}, \quad \boldsymbol{x}_{u^1u^2} = \begin{pmatrix} 0 \\ 0 \\ f_{u^1u^2} \end{pmatrix}.$$

So

$$l = \boldsymbol{x}_{u^1u^1} \cdot \boldsymbol{N} = \frac{f_{u^1u^1}}{\sqrt{1+f_{u^1}^2+f_{u^2}^2}}, \quad m = \boldsymbol{x}_{u^1u^2} \cdot \boldsymbol{N} = \frac{f_{u^1u^2}}{\sqrt{1+f_{u^1}^2+f_{u^2}^2}},$$
$$n = \boldsymbol{x}_{u^2u^2} \cdot \boldsymbol{N} = \frac{f_{u^2u^2}}{\sqrt{1+f_{u^1}^2+f_{u^2}^2}}.$$

Using the standard formulae for H, K one gets

$$H = \frac{f_{u^1u^1}(1+f_{u^2}^2) + f_{u^2u^2}(1+f_{u^1}^2) - 2f_{u^1u^2}f_{u^1}f_{u^2}}{2(1+f_{u^1}^2+f_{u_2}^2)^{3/2}}, \qquad K = \frac{f_{u^1u^1}f_{u^2u^2} - f_{u^1u^2}^2}{(1+f_{u^1}^2+f_{u^2}^2)^2}.$$

Hence $H = 0$ if and only if

$$f_{u^1u^1}(1+f_{u^2}^2) + f_{u^2u^2}(1+f_{u^1}^2) - 2f_{u^1u^2}f_{u^1}f_{u^2} = 0.$$

This partial differential equation is sometimes called Lagrange's equation. ◆

Section 8: The meaning of curvature

In this section we will re-interpret the first and second fundamental forms in terms of curves on the surface. We will also explore what information can be extracted from the curvature of a surface.

8.1 Curves on surfaces

As usual we will let $D \subseteq \mathbb{R}^2$ be a domain in the plane with coordinates (u, v) and let $\boldsymbol{x} : D \to \mathbb{E}^3$ be a local surface. Let

$$c : [a, b] \to D \qquad t \mapsto c(t) = (u(t), v(t))$$

be a (regular) curve in D. The composition

$$\boldsymbol{x} \circ c : [a, b] \to \mathbb{E}^3 \qquad t \mapsto \boldsymbol{x}(c(t))$$

describes a curve in $\mathbb{E}^3$ lying on the surface. We will often refer to $c(t) = (u(t), v(t))$ as a curve in the surface. We will abuse notation by writing $\boldsymbol{x}'$ for the velocity vector of this curve in $\mathbb{E}^3$:

$$\boldsymbol{x}' = \frac{d}{dt}\boldsymbol{x}(c(t)).$$

In a similar way, we will define

$$\boldsymbol{N}' = \frac{d}{dt}\boldsymbol{N}(c(t)).$$

Lemma 8.1. *We have*

$$\boldsymbol{x}' = d\boldsymbol{x}(c') \qquad \text{and} \qquad \boldsymbol{N}' = d\boldsymbol{N}(c')$$

where c' is the tangent vector to the curve c in D, so that $c' = u'\frac{\partial}{\partial u} + v'\frac{\partial}{\partial v}$.

Proof. Clearly

$$\frac{d}{dt}\boldsymbol{x}(c(t)) = \boldsymbol{x}_u u' + \boldsymbol{x}_v v' = d\boldsymbol{x}\left(u'\frac{\partial}{\partial u} + v'\frac{\partial}{\partial v}\right) = d\boldsymbol{x}(c') \ .$$

The same argument applies to $\boldsymbol{N}$. □

Proposition 8.2. *The arclength of the curve $\boldsymbol{x}(c(t))$, $t \in [a, b]$, lying on a surface is*

$$s = \int_a^b \sqrt{I(c', c')}dt$$

where I is the first fundamental form of the surface.

Proof. This follows immediately from the velocity of the curve

$$\boldsymbol{x}' \cdot \boldsymbol{x}' = d\boldsymbol{x}(c') \cdot d\boldsymbol{x}(c') = \mathrm{I}(c', c') \ .$$

□

Remark 8.3. The first fundamental form measures distances on the surface. The arclength formula gives a geometrical way of deducing I, since if one can write down an integral formula for the arclength of an arbitrary curve on the surface, one can then read off I.

Example 8.4. Consider the cylinder of radius a. Imagine two infinitesimally nearby points (u, ϕ) and $(u + \delta u, \phi + \delta\phi)$ connected by a line whose parameter length is δt. The distance between these points is by Pythagoras approximately given by $\sqrt{(\delta u)^2 + a^2(\delta\phi)^2} \approx \sqrt{u'^2 + a^2\phi'^2}\ \delta t$. This becomes exact in the limit the two points are coincident. Adding up such infinitesimal arc elements gives the total arc length along any curve, which is thus given by the integral $\int_a^b \sqrt{u'^2 + a^2\phi'^2} dt$. Comparing to the arc length formula in terms of I allows us to deduce $\mathrm{I} = (du)^2 + a^2(d\phi)^2$ as required.

Proposition 8.5. *For a curve lying on a surface,*

$$\boldsymbol{N} \cdot \frac{d^2}{dt^2}\boldsymbol{x}(c(t)) = I\!I(c', c')$$

where $I\!I$ is the second fundamental form of the surface.

Proof. Since the tangent to a curve lying on the surface must always be perpendicular to the unit normal $\boldsymbol{N}$ we have

$$\begin{aligned} 0 &= \frac{d}{dt}\left(\boldsymbol{x}' \cdot \boldsymbol{N}\right) \\ &= \boldsymbol{x}'' \cdot \boldsymbol{N} + \boldsymbol{x}' \cdot \boldsymbol{N}' \\ &= \boldsymbol{x}'' \cdot \boldsymbol{N} + d\boldsymbol{x}(c') \cdot d\boldsymbol{N}(c') \\ &= \boldsymbol{x}'' \cdot \boldsymbol{N} - \mathrm{I\!I}(c', c') \end{aligned}$$

□

Remark 8.6. The second fundamental form gives the normal (to the surface) component of the acceleration of a curve which is required for it to remain in the surface.

8.2 Invariance under Euclidean motions

Theorem 8.7. *Let $\boldsymbol{x} : D \to \mathbb{E}^3$ and $\hat{\boldsymbol{x}} : D \to \mathbb{E}^3$ be two surfaces related by a Euclidean motion, so*

$$\hat{\boldsymbol{x}}(u, v) = A\boldsymbol{x}(u, v) + \boldsymbol{a}$$

for all $(u, v) \in D$ where A is an orthogonal matrix with $\det A = 1$ *and $\boldsymbol{a} \in \mathbb{E}^3$. Then (using "hat" to denote quantities pertaining to the second surface) we have*

$$\hat{I} = I \qquad \text{and} \qquad \hat{I\!I} = I\!I$$

and hence in particular

$$\hat{H} = H \qquad \text{and} \qquad \hat{K} = K\ .$$

Proof. It is clear that $d\hat{\boldsymbol{x}} = Ad\boldsymbol{x}$ since A and $\boldsymbol{a}$ are constant. Hence $\hat{\boldsymbol{x}}_u = A\boldsymbol{x}_u$ and $\hat{\boldsymbol{x}}_v = A\boldsymbol{x}_v$, which implies $\hat{\boldsymbol{x}}_u \times \hat{\boldsymbol{x}}_v = A(\boldsymbol{x}_u \times \boldsymbol{x}_v)$ since $\det A = +1$. Therefore we have

$\hat{\boldsymbol{N}} = A\boldsymbol{N}$ and so $d\hat{\boldsymbol{N}} = Ad\boldsymbol{N}$. We now have all the ingredients to calculate how the fundamental forms transform. We have

$$\hat{\mathrm{I}} = d\hat{\boldsymbol{x}} \cdot d\hat{\boldsymbol{x}} = Ad\boldsymbol{x} \cdot Ad\boldsymbol{x} = d\boldsymbol{x} \cdot d\boldsymbol{x} = \mathrm{I}$$

and an identical argument shows $\hat{\mathrm{II}} = \mathrm{II}$. Thus the fundamental forms are unchanged because orthogonal transformations preserve dot products. (Of course, usually one is thinking of vectors of numbers, not of 1-forms, but the algebra is the same.) ☐

Remark 8.8. In fact the converse is also true: two surfaces with the same first and second fundamental forms must be related by a Euclidean motion. Hence

> ❝ The fundamental forms determine the surface
> (up to Euclidean motions) ❞

This should be contrasted with the fundamental theorem of curves we saw earlier, which shows that only two functions (curvature, torsion) are needed to determine a (biregular) curve up to Euclidean motions. The statement for surfaces is known as the *fundamental theorem of surfaces*. We shall not prove it in this course.

Proposition 8.9. *Under an orientation reversing isometry (i.e.* $\det A = -1$*), the principal curvatures and the mean curvature change sign but the Gauss curvature is unaffected.*

Proof. This follows from the fact that if $\det A = -1$, then $A\boldsymbol{x}_u \times A\boldsymbol{x}_v = -A(\boldsymbol{x}_u \times \boldsymbol{x}_v)$, whence $A\boldsymbol{N} = -\hat{\boldsymbol{N}}$. Thus $\hat{\mathrm{I}} = \mathrm{I}$ and $\hat{\mathrm{II}} = -\mathrm{II}$. This implies both principal curvatures change sign and hence $\hat{H} = -H$ and $\hat{K} = K$. ☐

8.3 Taylor series

We may obtain a local description for a surface which depends explicitly on the values of the principal curvatures at that point.

Proposition 8.10. *Let* $\boldsymbol{p} \in \boldsymbol{x}(D)$ *be a point on a local surface. By a Euclidean motion, choose* $\boldsymbol{p}$ *to be at the origin and the unit normal at that point to be along the positive* x_3*-axis so* $T_{\boldsymbol{p}}S$ *is the* (x^1, x^2) *plane. Near* $\boldsymbol{p}$ *we can parameterise the surface as the graph*

$$\boldsymbol{x}(u,v) = \begin{pmatrix} u \\ v \\ f(u,v) \end{pmatrix}$$

where at at the origin

$$\begin{aligned} I_{(0,0)} &= (du)^2 + (dv)^2 \\ II_{(0,0)} &= f_{uu}(0,0)(du)^2 + 2f_{uv}(0,0)dudv + f_{vv}(0,0)(dv)^2 \ . \end{aligned}$$

Proof. Let $\mathbf{N}$ be the normal vector to $S = \boldsymbol{x}(D)$ at $\mathbf{p} = \boldsymbol{x}(p)$. One can always find a rotation matrix $A \in SO(\mathbb{E}^3)$ such that

$$A\mathbf{N} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} .$$

Therefore the Euclidean motion given by A and $\mathbf{a} = -A\mathbf{p}$ gives rise to local surface

$$\widehat{\boldsymbol{x}} = A\boldsymbol{x} + \mathbf{a},$$

such that

$$\widehat{\boldsymbol{x}}(p) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

and

$$\widehat{\mathbf{N}}_p = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} . \tag{11}$$

We still need to re-parametrize, and for this we will invoke the inverse function theorem. Let

$$P : \mathbb{E}^3 \to \mathbb{P}^2 : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \end{pmatrix}$$

be the projection from three to two dimensions given by forgetting the third coordinate. Then

$$P \circ \widehat{\boldsymbol{x}} : D \subset \mathbb{R}^2 \to \mathbb{R}^2$$

has the property that

$$J_{P \circ \widehat{\boldsymbol{x}}}(p) = \left(P(\widehat{\boldsymbol{x}}_{u^1}) \quad P(\widehat{\boldsymbol{x}}_{u^2}) \right),$$

and because of (11) this 2×2 matrix is invertible. Therefore, by the inverse function theorem, we can find a local inverse, say $\Phi : D' \to D$ with $p \in \Phi(D')$ and $P \circ \widehat{\boldsymbol{x}} \circ \Phi = \mathrm{Id}_{D'}$. We can therefore re-parametrise the surface near our point of interest by $\widehat{\boldsymbol{x}} \circ \Phi$. In local coordinates this is now given by

$$\left(\widehat{\boldsymbol{x}} \circ \Phi\right)(u, v) = \begin{pmatrix} u \\ v \\ f(u, v) \end{pmatrix}$$

for some smooth function f.

In order to keep the notation light we will at this point 're-set' the notation, and refer to this surface $\widehat{\boldsymbol{x}} \circ \Phi$ simply as $\boldsymbol{x}$. In the given parametrisation we have

$$\boldsymbol{x}_u = \begin{pmatrix} 1 \\ 0 \\ f_u \end{pmatrix} , \qquad \boldsymbol{x}_v = \begin{pmatrix} 0 \\ 1 \\ f_v \end{pmatrix} , \qquad \boldsymbol{N}(0,0) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Since $\boldsymbol{x}_u(0,0), \boldsymbol{x}_v(0,0)$ span $T_{\boldsymbol{p}}S$, which is the plane orthogonal to $\boldsymbol{N}(0,0)$, we see that $f_u(0,0) = 0 = f_v(0,0)$. Hence $\mathrm{I}_{(0,0)} = du^2 + dv^2$. Differentiating again,

$$\boldsymbol{x}_{uu} = \begin{pmatrix} 0 \\ 0 \\ f_{uu} \end{pmatrix} \qquad \boldsymbol{x}_{uv} = \begin{pmatrix} 0 \\ 0 \\ f_{uv} \end{pmatrix} \qquad \boldsymbol{x}_{vv} = \begin{pmatrix} 0 \\ 0 \\ f_{vv} \end{pmatrix} ,$$

whence at the origin

$$\begin{aligned} \mathrm{II}_{(0,0)} &= \boldsymbol{x}_{uu} \cdot \boldsymbol{N}(0,0)(du)^2 + 2\boldsymbol{x}_{uv} \cdot \boldsymbol{N}(0,0)dudv + \boldsymbol{x}_{vv} \cdot \boldsymbol{N}(0,0)(dv)^2 \\ &= f_{uu}(0,0)(du)^2 + 2f_{uv}(0,0)dudv + f_{vv}(0,0)(dv)^2 . \end{aligned}$$

□

The astute reader will notice that the above still allows for rotations about the x_3-axis. We may in fact use this to arrange the x_1 and x_2 axes to be the principle directions at $\boldsymbol{p}$.

Proposition 8.11. *In setup of the previous Proposition, suppose the x_1, x_2 axes correspond to the principal directions. Then the Taylor series of the surface near the origin is*

$$x_3 = f(x_1, x_2) = \frac{k_1}{2}x_1^2 + \frac{k_2}{2}x_2^2 + \textit{higher order terms}$$

where k_1 and k_2 are the principal curvatures at $\boldsymbol{p}$.

Proof. If the x_1, x_2 axes are the principal directions, then

$$\mathrm{II}_{(0,0)} = k_1 du^2 + k_2 dv^2,$$

where k_1 and k_2 are the principal curvatures at the origin. Comparing with the expression for $\mathrm{II}_{(0,0)}$ in the previous Proposition, we get $f_{uu}(0,0) = k_1$, $f_{uv}(0,0) = 0$, $f_{vv}(0,0) = k_2$, so

$$f(u, v) = \tfrac{1}{2}k_1 u^2 + \tfrac{1}{2}k_2 v^2 + \text{higher order terms.}$$

$\square$

It is now an easy corollary to sketch the surface around the origin for different values of the principal curvatures, simply by sketching the graph of f.

Corollary 8.12. *If $K > 0$ at a point then the surface is "bowl-shaped" around that point and if $K < 0$ then it is "saddle-shaped". If $K = 0$ at a point then there are two possibilities: if both principal curvatures are zero, then the surface is planar around that point; otherwise it's like a "bent" plane. These surfaces are sketched below.*

Example 8.13. The cylinder of radius a oriented in the above setup may be defined implicitly by $x_2^2 + (x_3 + a)^2 = a^2$ and $x_1 \in \mathbb{R}$. The upper half can be written as a graph $x_3 = f(x_1, x_2) = -a + \sqrt{a^2 - x_2^2}$. Taylor expanding gives $f(x_1, x_2) = -\frac{1}{2a}x_2^2 + \dots$ and hence $k_1 = 0$ and $k_2 = -1/a$ are the principal curvatures at the origin.

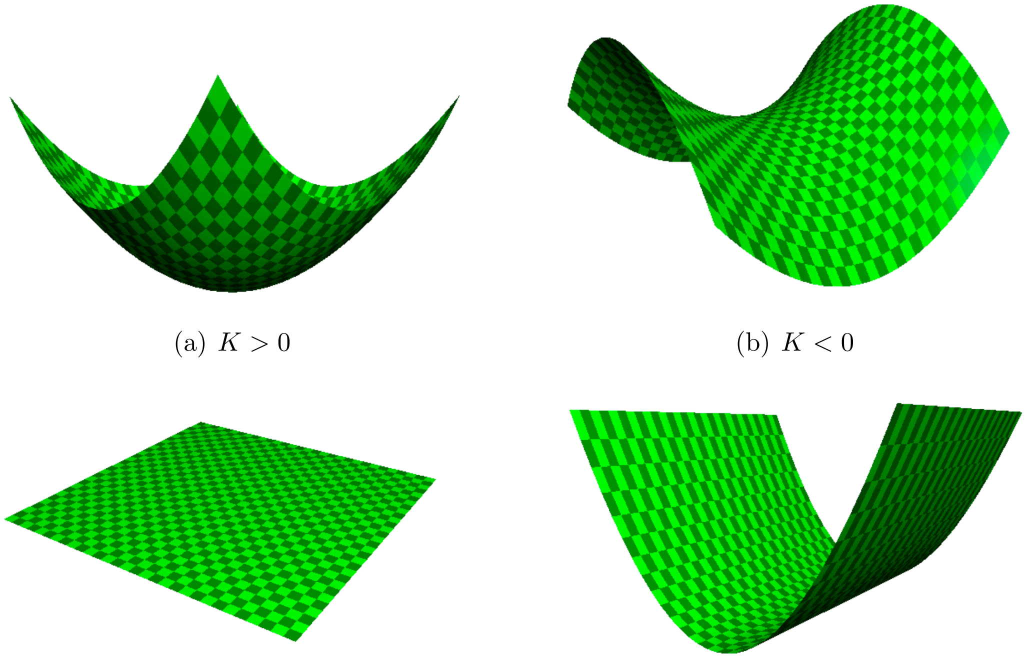
Definition 8.14. *A surface with $K = 0$ everywhere is called* **flat**.

Example 8.15. A plane has $k_1 = k_2 = 0$, whereas a cylinder has $k_1 = 0$ and $k_2 \neq 0$. Hence are both flat surfaces.

Definition 8.16. *A* **minimal surface** *is one with mean curvature $H = 0$ everywhere. A minimal surface must have $k_1 = -k_2$ and hence $K \leq 0$ at all points.*

Example 8.17. Minimal surfaces are difficult to find. The classical ones are: the plane (trivial), the **catenoid** (the most general minimal surface of revolution, given by the revolution of the profile curve $\cosh(u)$ discussed in Hand-In 3) and the **helicoid**.

Remark 8.18. It can be shown that the pressure difference across a soap film is $p = 4\sigma H$ where σ is the surface tension of the soap film and H the mean curvature of the surface defined by the soap film. Hence open soap films (so $p = 0$) want to be minimal surfaces.



(a) $K > 0$ (b) $K < 0$

(c) $K = 0$

Figure 1: The shape of the surface depending on the principal curvatures at a point.

8.4 Umbilical points and surfaces

Theorem 8.19. *A local surface has $\mathit{II} = 0$ if and only if it is (a piece of) a plane.*

Proof. For a plane $\boldsymbol{N}$ is constant, so $d\boldsymbol{N} = 0$ and hence $\mathbb{I} = 0$. Conversely, if $\mathbb{I} = 0$, then $l = m = n = 0$, hence $\boldsymbol{N}_u$ and $\boldsymbol{N}_v$ are each orthogonal to both $\boldsymbol{x}_u$ and $\boldsymbol{x}_v$. But, earlier we saw that $\boldsymbol{N}_u$ and $\boldsymbol{N}_v$ are in the tangent plane and hence they must both to vanish. □

Theorem 8.20. *A local surface has $\mathit{II} = \lambda I$ where $\lambda \neq 0$ is a constant, if and only if the surface is a (piece) of a sphere radius $1/|\lambda|$.*

Proof. Any sphere of radius $1/|\lambda|$ can be written as $\boldsymbol{x} - \boldsymbol{x}_0 = -\frac{1}{\lambda}\boldsymbol{N}$, where for $\lambda > 0$ the normal $\boldsymbol{N}$ is inward pointing and for $\lambda < 0$ the normal is outward pointing. It follows that $d\boldsymbol{N} = -\lambda d\boldsymbol{x}$ and hence $\mathbb{I} = \lambda \mathrm{I}$. Conversely, if $\mathbb{I} = \lambda\mathrm{I}$, then $l = \lambda E, m = \lambda F, n = \lambda G$ and hence by a similar argument to the previous theorem $\boldsymbol{N}_u = -\lambda \boldsymbol{x}_u, \boldsymbol{N}_v = -\lambda \boldsymbol{x}_v$, which integrates to $\boldsymbol{x} - \boldsymbol{x}_0 = -\frac{1}{\lambda}\boldsymbol{N}$. □

Definition 8.21. *An* **umbilical point** *on a surface is a point where $k_1 = k_2$. Consequently all directions are principal at such a point.*

Lemma 8.22. *For any surface $K \leq H^2$, with equality only at umbilic points.*

Proof. Follows immediately from $0 \leq (k_1 - k_2)^2 = (k_1 + k_2)^2 - 4k_1k_2$. □

Theorem 8.23. *A local surface for which every point is umbilical, is either part of a sphere or a plane.*

Proof. The principal curvatures $k_1 = k_2 = \lambda$, where λ is now a function on D, which is equivalent to $\mathbb{I} = \lambda\mathrm{I}$. Thus, as above we have $d\boldsymbol{N} = -\lambda d\boldsymbol{x}$. Take d to find $0 = d\lambda \wedge d\boldsymbol{x} = (d\lambda \wedge du)\boldsymbol{x}_u + (d\lambda \wedge dv)\boldsymbol{x}_v$ and hence by regularity of the surface $d\lambda \wedge du = d\lambda \wedge dv = 0$. Thus λ is a constant function on D. The theorem now follows from above, depending on whether $\lambda = 0$ or $\lambda \neq 0$. □

8.5 Exercises

Exercise 8.24. Consider the curve

$$z(t) = (\alpha(t), \phi(t)) = (t, 10t)$$

in the sphere of radius a. Check explicitly that $d\boldsymbol{x}(z') = \boldsymbol{x}'$ and check explicitly also that

$$\boldsymbol{N} \cdot \frac{d^2}{dt^2}\boldsymbol{x}(z(t)) = \mathbb{I}(z', z').$$

Solution. This can be a bit messy if you don't stick to the point! The curve on the sphere is

$$\boldsymbol{x}(t) = \begin{pmatrix} a \sin t \cos 10t \\ a \sin t \sin 10t \\ a \cos t \end{pmatrix}$$

Differentiating twice, we get

$$\boldsymbol{x}''(t) = a \begin{pmatrix} -101 \sin t \cos 10t - 20 \cos t \sin 10t \\ -101 \sin t \sin 10t + 20 \cos t \cos 10t \\ -\cos t \end{pmatrix}$$

Then

$$\boldsymbol{N}(t) = \begin{pmatrix} \sin t \cos 10t \\ \sin t \sin 10t \\ \cos t \end{pmatrix}$$

and a bit of calculation gives

$$\boldsymbol{N} \cdot \boldsymbol{x}'' = -a(1 + 100 \sin^2 t).$$

On the other hand, for the sphere

$$\mathbb{I} = -a((d\alpha)^2 + \sin^2 \alpha \, (d\phi)^2)$$

and for our curve,

$$z'(t) = \frac{\partial}{\partial \alpha} + 10\frac{\partial}{\partial \phi}$$

and so

$$\mathbb{I}(z', z') = -a(\alpha'^2 + \sin^2 \alpha \, \phi'^2) = -a(1 + 100 \sin^2 t).$$

◆

Exercise 8.25. Write down an integral that gives the arc-length of the curve in the previous exercise between $t = 0$ and $t = 2$.

Solution. Just doing the case of $a = 1$ we get the following. We know that $\mathrm{I} = (d\alpha)^2 + \sin^2\alpha)(d\phi)^2$. From above,

$$z'(t) = \frac{\partial}{\partial\alpha} + 10\frac{\partial}{\partial\phi}$$

and so

$$\mathrm{I}(z', z') = 1 + 100\sin^2\alpha = 1 + 100\sin^2 t.$$

Thus the arc length is

$$\int_0^2 \sqrt{1 + 100\sin^2 t}dt.$$

◆

Exercise 8.26. Consider a closed (no edge), bounded surface in $\mathbb{E}^3$. The distance from the origin in $\mathbb{E}^3$ is a continuous function on the surface, and it must therefore attain a maximum value d at some point $\boldsymbol{p}$ on the surface. By considering the Taylor expansion at $\boldsymbol{p}$, show that $K > 0$ at $\boldsymbol{p}$. (You have proved the Theorem that on any such surface there must be a point where $K > 0$.)

Solution. At the furthest point from the origin the normal vector must be in the direction of the origin. Taking the standard set-up for power series expansion at the point $\boldsymbol{p}$, with the axes being principal directions, the surface is given by

$$z = f(x, y) = k_1x^2 + k_2y^2.$$

Near to the point, the surface must be contained within the sphere of radius d centred at the origin, which has Taylor expansion

$$z = g(x, y) = d^{-1}x^2 + d^{-1}y^2.$$

Clearly this can only happen if $k_1, k_2 \geq (1/d)$ and so $K \geq 1/d^2$. ◆

Exercise 8.27. Write a piece of a sphere of radius r as the graph of a function and hence compute the Gauss and mean curvature by making a power series expansion.

Solution. The sphere through the origin of radius r and having the x, y plane as its tangent plane there is

$$x^2 + y^2 + (z - r)^2 = r^2.$$

Rearranging, we get

$$z = f(x, y) = r - \sqrt{r^2 - x^2 - y^2}.$$

(Negative square root needed to get the set-up right!) Now,

$$f_x = \frac{x}{\sqrt{r^2 - x^2 - y^2}}, \quad f_y = \frac{y}{\sqrt{r^2 - x^2 - y^2}}$$

(the second by symmetry please, not by more calculation). Then

$$f_{xx}(0, 0) = f_{yy}(0, 0) = 1/r, \quad f_{xy}(0, 0) = 0.$$

(Trivial, provided you notice the terms that are going to be zero at zero and don't write them down!) We see that at the origin $\mathrm{II} = (1/r)((du)^2 + (dv)^2)$ and so $k_1 = k_2 = 1/r$ and so $H = 1/r$, $K = 1/r^2$.

Our set up has used the inward pointing normal, if you had used the other (so that f would be exactly minus what it is above) then k_1, k_2, H would change sign but K would not. ◆

Exercise 8.28. Consider the surface of revolution with profile function f. Show (or refer to previous calculations) that if $f''(u) < 0$ then $K > 0$ and if $f''(u) > 0$ then $K < 0$. Explain how this is to be expected from the shape of the surface.

Solution. Draw a picture: if $f'' < 0$ then the profile curve is convex and the surface is bowl-shaped, etc. ◆

Section 9: Moving frames in Euclidean space

Remark 9.1. As in this chapter and next we will be referring mostly to ambient Euclidean space $\mathbb{E}^3$, we will use 'conventional' notation of vectors for this, and in particular not use the Einstein summation convention.

9.1 Euclidean maps $\mathbb{R}^n \to \mathbb{E}^m$

In this section, we will be studying smooth maps

$$\boldsymbol{x} : D \to \mathbb{E}^m$$

from some domain $D \subseteq \mathbb{R}^n$ to m-dimensional Euclidean space $\mathbb{E}^m$. We will mostly restrict to $m = 3$, although all formulae and concepts generalise easily to arbitrary dimension.

Notation

We will let $u = (u^1, \ldots, u^n)$ denote the coordinates on D and hence the above map can be written explicitly as

$$\boldsymbol{x}(u) = \boldsymbol{x}(u^1, \ldots, u^n) = \begin{pmatrix} x^1(u^1, \ldots, u^n) \\ x^2(u^1, \ldots, u^n) \\ x^3(u^1, \ldots, u^n) \end{pmatrix}$$

In the special case $n = 2$ we sometimes write $(u^1, u^2) = (u, v)$.

For $n = 1$ such maps are the curves in Euclidean space we saw earlier. The $n = 2$ case describes surfaces in Euclidean space, which we will study in detail later. For $n = 3$ such maps can be thought of as 'changing coordinates' on some domain in Euclidean space.

9.2 Moving frames in $\mathbb{E}^3$

Definition 9.2. *A* **moving frame** *for $\mathbb{E}^3$ on D is a collection of maps $\boldsymbol{e}_i : D \to \mathbb{E}^3$ for $i = 1, 2, 3$ such that for all $u \in D$ the $\boldsymbol{e}_i(u)$ form an oriented orthonormal basis of $\mathbb{E}^3$. Orthonormal means*

$$\boldsymbol{e}_i(u) \cdot \boldsymbol{e}_j(u) = \begin{cases} 1 & i = j \\ 0 & \textit{otherwise} \end{cases}$$

and oriented means that $\boldsymbol{e}_3(u) = \boldsymbol{e}_1(u) \times \boldsymbol{e}_2(u)$.

Remark 9.3. The definition of orientedness uses the usual vector cross product, which can only be defined in three dimensions. For general m there is a different definition of oriented frame which will be remarked in passing below.

Definition 9.4. *If $\boldsymbol{v} : D \to \mathbb{E}^3$, given by*

$$\boldsymbol{v}(u) = \boldsymbol{v}(u^1, \ldots, u^n) = \begin{pmatrix} v^1(u^1, \ldots, u^n) \\ v^2(u^1, \ldots, u^n) \\ v^3(u^1, \ldots, u^n) \end{pmatrix}$$

is an $\mathbb{E}^3$-valued function on D, then we write $d\boldsymbol{v}$ for its entry by entry exterior derivative:

$$d\boldsymbol{v} = \begin{pmatrix} dv^1 \\ dv^2 \\ dv^3 \end{pmatrix} .$$

We refer to an object like $d\boldsymbol{v}$ as a **matrix of 1-forms on** *D or as a* **$\mathbb{E}^3$-valued 1-form on** *D. If w is a vector field on D, then we write*

$$d\boldsymbol{v}(w) = \begin{pmatrix} dv^1(w) \\ dv^2(w) \\ dv^3(w) \end{pmatrix} : D \to \mathbb{E}^3 .$$

So $d\boldsymbol{v}$ feeds on vector fields in D and spits out vectors in $\mathbb{E}^3$, hence the term "$\mathbb{E}^3$-valued 1-form".

Example 9.5. Consider the "cylindrical coordinate map" where $(\rho, \phi, z) \in D$

$$\boldsymbol{x}(\rho, \phi, z) = \begin{pmatrix} \rho\cos\phi \\ \rho\sin\phi \\ z \end{pmatrix} .$$

Then

$$d\boldsymbol{x} = \begin{pmatrix} d\rho\cos\phi - \rho\sin\phi d\phi \\ d\rho\sin\phi + \rho\cos\phi d\phi \\ dz \end{pmatrix} ,$$

which can be written as

$$d\boldsymbol{x} = \begin{pmatrix} \cos\phi \\ \sin\phi \\ 0 \end{pmatrix} d\rho + \begin{pmatrix} -\sin\phi \\ \cos\phi \\ 0 \end{pmatrix} \rho d\phi + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} dz ,$$

where

$$\boldsymbol{e}_1 = \begin{pmatrix} \cos\phi \\ \sin\phi \\ 0 \end{pmatrix} , \qquad \boldsymbol{e}_2 = \begin{pmatrix} -\sin\phi \\ \cos\phi \\ 0 \end{pmatrix} , \qquad \boldsymbol{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

is an oriented moving frame.

Example 9.6. The Frenet-Serret frame $\boldsymbol{T}(t), \boldsymbol{N}(t), \boldsymbol{B}(t)$ of a (biregular) curve $\boldsymbol{x} : (a, b) \to \mathbb{E}^3$, is an example of a moving frame for $\mathbb{E}^3$ on (a, b).

9.3 Connection forms and the structure equations

Throughout this section we have a map $\boldsymbol{x} : D \to \mathbb{E}^3$ and a moving frame $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$.

Notation

We will assemble our frame $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ into a 3×3 matrix

$$\boldsymbol{E} = \begin{pmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 \end{pmatrix}$$

whose columns are the three vectors $\boldsymbol{e}_i$. Since the $\boldsymbol{e}_i$ are linearly independent, the matrix $\boldsymbol{E}$ is invertible, i.e. $\det \boldsymbol{E} \neq 0$. In fact, because the frame is orthonormal, the matrix $\boldsymbol{E}$ is orthogonal, i.e. $\boldsymbol{E}^T\boldsymbol{E} = \mathbf{I}$. Moreover, because the frame is oriented $\det \boldsymbol{E} = (\boldsymbol{e}_1 \times \boldsymbol{e}_2) \cdot \boldsymbol{e}_3 = +1$ and so it is a rotation matrix.

Remark 9.7. The above shows how to generalise the notion of an oriented frame to $m > 3$: we simply demand that $\boldsymbol{E} : D \to \mathrm{SO}(m)$, where $SO(m)$ is the group of $m \times m$ orthogonal matrices with unit determinant (called the special orthogonal group).

Since $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ is an orthonormal basis for $\mathbb{E}^3$, any vector $\boldsymbol{v}$ can be expanded as $\boldsymbol{v} = \sum_{i=1}^3 (\boldsymbol{e}_i \cdot \boldsymbol{v})\, \boldsymbol{e}_i$ and the same applies to a vector-valued 1-form, such as $d\boldsymbol{x}$. Therefore we define 1-forms $\theta_1, \theta_2, \theta_3$ by

$$\theta_i = \boldsymbol{e}_i \cdot d\boldsymbol{x} \in \Omega^1(D) \qquad \text{so that} \qquad d\boldsymbol{x} = \sum_{i=1}^{3} \theta_i \boldsymbol{e}_i \ .$$

Notation

If we assemble $\theta_1, \theta_2, \theta_3$ into a column matrix $\boldsymbol{\theta}$ of 1-forms,

$$\boldsymbol{\theta} = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix} ,$$

then our expression for $d\boldsymbol{x}$ can be written as

$$d\boldsymbol{x} = \begin{pmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix} = \boldsymbol{E}\boldsymbol{\theta} \ .$$

Example 9.8. Continuing with Example 9.5, the 1-forms θ_i can be read off from

$$d\boldsymbol{x} = \sum_{i=1}^{3} \theta_i \boldsymbol{e}_i \ ,$$

so that

$$\theta_1 = d\rho \qquad \theta_2 = \rho d\phi \qquad \theta_3 = dz \ .$$

We can also decompose $d\boldsymbol{e}_i$ in terms of the moving frame:

$$d\boldsymbol{e}_i = \sum_{j=1}^{3} (\boldsymbol{e}_j \cdot d\boldsymbol{e}_i)\, \boldsymbol{e}_j \ ,$$

which suggests the following definition.

Definition 9.9. *The* 1*-forms* $\omega_{ij} = \boldsymbol{e}_i \cdot d\boldsymbol{e}_j \in \Omega^1(D)$ *are called the* **connection 1-forms** *and by definition satisfy*

$$d\boldsymbol{e}_i = \sum_{j=1}^{3} \boldsymbol{e}_j \omega_{ji} \ .$$

Proposition 9.10. *The connection* 1*-forms* ω_{ij} *are related by the antisymmetry property:*

$$\omega_{ij} = -\omega_{ji}$$

for all i, j*. In particular* $\omega_{ii} = 0$ *for all* i*.*

Proof. The antisymmetry follows from

$$0 = d(\boldsymbol{e}_i \cdot \boldsymbol{e}_j) = (d\boldsymbol{e}_i) \cdot \boldsymbol{e}_j + \boldsymbol{e}_i \cdot (d\boldsymbol{e}_j) = \omega_{ji} + \omega_{ij}.$$

□

Remark 9.11. The notation may be confusing. Here, ω_{ij} for each value of i, j is a 1-form, so we have a set of 1-forms labelled by indices i, j. Earlier, given a single *2-form* ω we denoted its components in a coordinate basis by ω_{ij} which are functions. As long as one keeps track of which space an object belongs to, there will be no confusion.

Example 9.12. Continuing with Example 9.8, the connection 1-forms ω_{ij} can be calculated by differentiating the moving frame. In fact, in this case, we need only differentiate two of them. First,

$$d\boldsymbol{e}_3 = 0 \implies \omega_{13} = \omega_{23} = 0 \ ,$$

and then

$$d\boldsymbol{e}_2 = -\boldsymbol{e}_1 d\phi \implies \omega_{12} = -d\phi \ .$$

Notation

The above proposition shows we can assemble the connection forms into a 3×3 *anti-symmmetric* matrix $\boldsymbol{\omega}$ according to

$$\boldsymbol{\omega} = \begin{pmatrix} 0 & \omega_{12} & \omega_{13} \\ -\omega_{12} & 0 & \omega_{23} \\ -\omega_{13} & -\omega_{23} & 0 \end{pmatrix}$$

where we have replaced ω_{21} with $-\omega_{12}$, etc. Let us also write $d\boldsymbol{E}$ for the term-by-term exterior derivative: $d\boldsymbol{E} = \begin{pmatrix} d\boldsymbol{e}_1 & d\boldsymbol{e}_2 & d\boldsymbol{e}_3 \end{pmatrix}$. Then we have

$$d\boldsymbol{E} = \boldsymbol{E}\boldsymbol{\omega}. \tag{12}$$

We will also write

$$d\boldsymbol{\theta} = \begin{pmatrix} d\theta_1 \\ d\theta_2 \\ d\theta_3 \end{pmatrix} \ .$$

Theorem 9.13. *The* **first structure equations** *are*

$$\boxed{d\boldsymbol{\theta} + \boldsymbol{\omega} \wedge \boldsymbol{\theta} = 0}\ .$$

Here, we are multiplying matrices of forms in the usual way, but taking the wedge where we would normally be multiplying numbers. Thus, written out more fully we have

$$\begin{pmatrix} d\theta_1 \\ d\theta_2 \\ d\theta_3 \end{pmatrix} + \begin{pmatrix} 0 & \omega_{12} & \omega_{13} \\ -\omega_{12} & 0 & \omega_{23} \\ -\omega_{13} & -\omega_{23} & 0 \end{pmatrix} \wedge \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix} = 0.$$

Reading off components, we can also write this as three differential form equations:

$$d\theta_i + \sum_{j=1}^{3} \omega_{ij} \wedge \theta_j = 0, \quad i = 1, 2, 3.$$

Proof. We differentiate the equation $d\boldsymbol{x} = \boldsymbol{E}\boldsymbol{\theta}$, using that $d^2 = 0$, to obtain

$$\begin{aligned} 0 = d^2\boldsymbol{x} = d(\boldsymbol{E}\boldsymbol{\theta}) \\ &= d\boldsymbol{E} \wedge \boldsymbol{\theta} + \boldsymbol{E} d\boldsymbol{\theta} \\ &= \boldsymbol{E}\boldsymbol{\omega} \wedge \boldsymbol{\theta} + \boldsymbol{E} d\boldsymbol{\theta} && \text{(using equation (12))} \\ &= \boldsymbol{E}\left(\boldsymbol{\omega} \wedge \boldsymbol{\theta} + d\boldsymbol{\theta}\right)\ . \end{aligned}$$

Now multiply by both sides on the left with the inverse matrix $\boldsymbol{E}^{-1}$ to get the result. □

Example 9.14. Continuing with Example 9.12, we can check the first structure equations:

$$d\boldsymbol{\theta} = d\begin{pmatrix} d\rho \\ \rho d\phi \\ dz \end{pmatrix} = \begin{pmatrix} 0 \\ d\rho \wedge d\phi \\ 0 \end{pmatrix}\ ,$$

whereas

$$\boldsymbol{\omega} \wedge \boldsymbol{\theta} = \begin{pmatrix} 0 & -d\phi & 0 \\ d\phi & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \wedge \begin{pmatrix} d\rho \\ \rho d\phi \\ dz \end{pmatrix} = \begin{pmatrix} 0 \\ d\phi \wedge d\rho \\ 0 \end{pmatrix}\ ,$$

whence $d\boldsymbol{\theta} + \boldsymbol{\omega} \wedge \boldsymbol{\theta} = 0$ as expected.

Theorem 9.15. *The* **second structure equations** *are*

$$\boxed{d\boldsymbol{\omega} + \boldsymbol{\omega} \wedge \boldsymbol{\omega} = 0}\ .$$

Again, one can write the matrices in full and extract the components:

$$d\omega_{ij} + \sum_{k=1}^{3} \omega_{ik} \wedge \omega_{kj} = 0$$

which can be further written out explicitly as

$$\begin{aligned} d\omega_{12} - \omega_{13} \wedge \omega_{23} &= 0 \\ d\omega_{13} + \omega_{12} \wedge \omega_{23} &= 0 \\ d\omega_{23} - \omega_{12} \wedge \omega_{13} &= 0\ . \end{aligned}$$

Proof. We now differentiate the equation $d\boldsymbol{E} = \boldsymbol{E}\boldsymbol{\omega}$ and use that $d^2 = 0$ to obtain

$$\begin{aligned} 0 = d^2\boldsymbol{E} &= d(\boldsymbol{E}\boldsymbol{\omega}) \\ &= d\boldsymbol{E} \wedge \boldsymbol{\omega} + \boldsymbol{E} d\boldsymbol{\omega} \\ &= \boldsymbol{E}\boldsymbol{\omega} \wedge \boldsymbol{\omega} + \boldsymbol{E} d\boldsymbol{\omega} && \text{(using equation (12))} \\ &= \boldsymbol{E}(\boldsymbol{\omega} \wedge \boldsymbol{\omega} + d\boldsymbol{\omega}) \ . \end{aligned}$$

Again, multiplying by $\boldsymbol{E}^{-1}$ gives the result. □

Example 9.16. Continuing with Example 9.14, we have that clearly both $d\boldsymbol{\omega}$ and $\boldsymbol{\omega} \wedge \boldsymbol{\omega}$ are separately zero, whence the second structure equation is trivially satisfied.

Remark 9.17. The structure equations tell you everything about moving frames in the following sense. Let the domain D have no "holes". Given forms ω_{jk} satisfying the second structure equations, and given a frame $\boldsymbol{e}_1(p), \boldsymbol{e}_2(p), \boldsymbol{e}_3(p)$ at some single point $p \in D$, there exists a unique moving frame $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ on D which agrees with the given frame at p and has the ω_{jk} as connection forms. Furthermore, if one has forms θ_k satisfying the first structure equations and one specifies $\boldsymbol{x}(p)$ there is a unique map $\boldsymbol{x} : D \to \mathbb{E}^3$ with $\boldsymbol{x}(p)$ as given for which $\theta_k = \boldsymbol{e}_k \cdot d\boldsymbol{x}$. This generalises the fundamental theorem of curves we saw earlier. We will not prove (or use) this result in this course.

Remark 9.18. Notice the definition of the connection 1-forms and the second structure equation only requires the existence of a moving frame and not a map $\boldsymbol{x} : D \to \mathbb{E}^3$. The structure equations exist in the more general context of Riemannian geometry, where $\mathbf{R} = d\boldsymbol{\omega} + \boldsymbol{\omega} \wedge \boldsymbol{\omega}$ is the Riemann curvature, which in general is non-vanishing. For us this is zero because our moving frame is in Euclidean space. We will see non-vanishing examples in the context of surfaces.

Remark 9.19. In other dimensions, everything generalises in the obvious way, with sums running over $1, \ldots, n$, or equivalently all the matrices with 3 rows or columns now have n. We will write down just the case of $\mathbb{E}^2$. In that case, one has forms θ_1, θ_2 with

$$d\boldsymbol{x} = \boldsymbol{e}_1\theta_1 + \boldsymbol{e}_2\theta_2$$

There is just one connection form ω_{12} and

$$d\boldsymbol{e}_1 = -\boldsymbol{e}_2\omega_{12}, \quad d\boldsymbol{e}_2 = \boldsymbol{e}_1\omega_{12}.$$

The first structure equations become

$$d\theta_1 + \omega_{12} \wedge \theta_2 = 0, \quad d\theta_2 - \omega_{12} \wedge \theta_1 = 0$$

and the second structure equations become simply

$$d\omega_{12} = 0 \ .$$

9.4 Exercises

Exercise 9.20. Consider the "spherical polar coordinate map"

$$\boldsymbol{x}(r, \alpha, \phi) = \begin{pmatrix} r \sin\alpha \cos\phi \\ r \sin\alpha \sin\phi \\ r \cos\alpha \end{pmatrix} .$$

(We use α for the angle usually labelled as θ to avoid confusion with the 1-forms.) Calculate $d\boldsymbol{x}$. Show that

$$\boldsymbol{e}_1 = \begin{pmatrix} \sin\alpha \cos\phi \\ \sin\alpha \sin\phi \\ \cos\alpha \end{pmatrix} , \quad \boldsymbol{e}_2 = \begin{pmatrix} \cos\alpha \cos\phi \\ \cos\alpha \sin\phi \\ -\sin\alpha \end{pmatrix} , \quad \boldsymbol{e}_3 = \begin{pmatrix} -\sin\phi \\ \cos\phi \\ 0 \end{pmatrix}$$

is a moving frame. Show also that

$$\boldsymbol{x}_r = \frac{\partial \boldsymbol{x}}{\partial r} = \boldsymbol{e}_1, \quad \boldsymbol{x}_\alpha = \frac{\partial \boldsymbol{x}}{\partial \alpha} = r\, \boldsymbol{e}_2, \quad \boldsymbol{x}_\phi = \frac{\partial \boldsymbol{x}}{\partial \phi} = r \sin\alpha\, \boldsymbol{e}_3.$$

Thus the vectors in the frame are the unit vectors in the direction of increasing r, α, ϕ respectively (draw a diagram!). Hence deduce

$$d\boldsymbol{x} = \boldsymbol{x}_r dr + \boldsymbol{x}_\alpha d\alpha + \boldsymbol{x}_\phi d\phi = \boldsymbol{e}_1 dr + \boldsymbol{e}_2 r d\alpha + \boldsymbol{e}_3 r \sin\alpha d\phi \ .$$

Solution. To show it is a moving framee you need only show that the dot products are zero or one as appropriate, and that $\boldsymbol{e}_3 = \boldsymbol{e}_1 \times \boldsymbol{e}_2$ rather than minus that.

For the final part one needs to compute $d\boldsymbol{x}$, the exterior derivative of each entry of $\boldsymbol{x}$. One should get

$$d\boldsymbol{x} = \begin{pmatrix} \sin\alpha \cos\phi dr + r\cos\alpha\cos\phi d\alpha - r \sin\alpha \sin\phi d\phi \\ \sin\alpha \sin\phi dr + r \cos\alpha \sin\phi d\alpha + r \sin\alpha \cos\phi d\phi \\ \cos\alpha dr - r \sin\alpha d\alpha \end{pmatrix} .$$

It can then be verified that this satisfies the stated identities above using the various definitions given.

◆

Exercise 9.21. Continuing with Exercise 9.20, compute the forms $\theta_1, \theta_2, \theta_3$.

Solution. We have

$$\theta_1 = d\boldsymbol{x} \cdot \boldsymbol{e}_1 = dr, \quad \theta_2 = d\boldsymbol{x} \cdot \boldsymbol{e}_2 = r d\alpha, \quad \theta_3 = d\boldsymbol{x} \cdot \boldsymbol{e}_3 = r \sin\alpha d\phi,$$

immediately from Exercise 9.20. ◆

Exercise 9.22. Continuing with Exercise 9.21, compute two of $d\boldsymbol{e}_1, d\boldsymbol{e}_2, d\boldsymbol{e}_3$ and hence compute the connection forms $\omega_{12}, \omega_{13}, \omega_{23}$.

Solution. It is easy to calculate

$$d\boldsymbol{e}_1 = \begin{pmatrix} \cos\alpha \cos\phi d\alpha - \sin\alpha \sin\phi d\phi \\ \cos\alpha \sin\phi d\alpha + \sin\alpha\cos\phi d\phi \\ -\sin\alpha d\alpha \end{pmatrix} , \quad d\boldsymbol{e}_3 = \begin{pmatrix} -\cos\phi d\phi \\ -\sin\phi d\phi \end{pmatrix}$$

Thus,

$$\omega_{12} = -\boldsymbol{e}_2 \cdot d\boldsymbol{e}_1 = -d\alpha$$
$$\omega_{13} = \boldsymbol{e}_1 \cdot d\boldsymbol{e}_3 = -\sin\alpha d\phi$$
$$\omega_{23} = \boldsymbol{e}_2 \cdot d\boldsymbol{e}_3 = -\cos\alpha d\phi$$

◆

Exercise 9.23. Continuing with Exercise 9.22, check that the first structure equations are satisfied. Conversely, hide for a moment your values for the connection forms ω_{jk}. Write down the first structure equations (using your computed forms θ_j) and try and compute the connection forms from these equations.

Solution. We have to check

$$d\theta_1 + \omega_{12} \wedge \theta_2 + \omega_{13} \wedge \theta_3 = 0$$
$$d\theta_2 - \omega_{12} \wedge \theta_1 + \omega_{23} \wedge \theta_3 = 0$$
$$d\theta_3 - \omega_{13} \wedge \theta_1 - \omega_{23} \wedge \theta_2 = 0$$

are satisfied.

Taking our computed θ_k we have

$$d\theta_1 = 0, \quad d\theta_2 = dr \wedge d\alpha, \quad d\theta_3 = r\cos\alpha d\alpha \wedge d\phi + \sin\alpha dr \wedge d\phi.$$

Now we can write down the first structure equations, which become

$$\omega_{12} \wedge d\alpha + \sin\alpha\omega_{13} \wedge d\phi = 0$$
$$dr \wedge d\alpha - \omega_{12} \wedge dr + r\sin\alpha\omega_{23} \wedge d\phi = 0$$
$$r\cos\alpha d\alpha \wedge d\phi + \sin\alpha dr \wedge d\phi - \omega_{13} \wedge dr - r\omega_{23} \wedge d\alpha = 0$$

One can be systematic and write down arbitrary 1-forms for each of the ω_{jk} and see what the equations tell us. It's much easier however to just "feel your way". The simplest way to get the terms we need to cancel in the final equation is to set

$$\omega_{13} = -\sin\alpha d\phi, \quad \omega_{23} = -\cos\alpha d\phi.$$

(Note that ω_{13} could also have a "dr" term and ω_{23} a "$d\alpha$" term and the last equation would still be OK.) Now look at the second equation and note that

$$\omega_{12} = -d\alpha$$

will satisfy that. Now in fact the first equation is satisfied also and so we have the solution. Warning: you must be careful when guessing a solution in this way that all the structure equations are satisfied, otherwise one is not guaranteed a solution. ◆

Exercise 9.24. Continuing with Exercise 9.23, check that the second structure equations are satisfied.

Solution. We have to check that

$$d\omega_{12} - \omega_{13} \wedge \omega_{23} = 0$$
$$d\omega_{13} + \omega_{12} \wedge \omega_{23} = 0$$
$$d\omega_{23} - \omega_{12} \wedge \omega_{13} = 0$$

are satisfied. This is again routine. To take just the very last equation, for instance,

$$d\omega_{23} - \omega_{12} \wedge \omega_{13} = \sin\alpha d\alpha \wedge d\phi - (-d\alpha) \wedge (-\sin\alpha d\phi) = 0.$$

◆

Exercise 9.25. Compute all the paraphernalia and verify the structure equations for the polar co-ordinate map in $\mathbb{E}^2$. (This should look like a stripped down version of the cylindrical polar calculations in the examples.)

Exercise 9.26. Consider the structure equations for a map $\boldsymbol{x} : \mathbb{R}^2 \to \mathbb{E}^2$. Suppose that θ_1, θ_2 are everywhere linearly independent. Show that given θ_1, θ_2, the connection form ω_{12} is uniquely determined by the first structure equations.

Solution. Write $\omega_{12} = a\theta_1 + b\theta_2$ (possible if θ_1, θ_2 are everywhere linearly independent). The first structure equation $d\theta_1 + \omega_{12} \wedge \theta_2 = 0$ gives $d\theta_1 = -a\theta_1 \wedge \theta_2$. The second structure equation $d\theta_2 - \omega_{12} \wedge \theta_1 = 0$ gives $d\theta_2 = -b\theta_1 \wedge \theta_2$. Thus given θ_1, θ_2 these equations determine a, b uniquely (note $d\theta_i$ are 2-forms on $\mathbb{R}^2$ and so must be proportional to $\theta_1 \wedge \theta_2$.) ◆

Exercise 9.27. Let $\boldsymbol{x} : D \to \mathbb{E}^2$ be a map from a domain $D \subseteq \mathbb{R}^2$. Suppose that $\boldsymbol{x}_u, \boldsymbol{x}_v$ are everywhere non-zero and orthogonal. Set $h = |\boldsymbol{x}_u|$ and $k = |\boldsymbol{x}_v|$. Show that

$$\frac{\partial}{\partial v}\left(\frac{h_v}{k}\right) + \frac{\partial}{\partial u}\left(\frac{k_u}{h}\right) = 0\ .$$

(Hint: Consider the frame $\boldsymbol{e}_1, \boldsymbol{e}_2$ such that $\boldsymbol{x}_u = h\boldsymbol{e}_1$ and $\boldsymbol{x}_v = k\boldsymbol{e}_2$. Calculate θ_1, θ_2 and hence compute ω_{12} from the first structure equations. Then examine the consequences of the second structure equation.)

Solution. Following the hint,

$$d\boldsymbol{x} = \boldsymbol{x}_u du + \boldsymbol{x}_v dv,$$

and so

$$\theta_1 = h du, \quad \theta_2 = k dv.$$

Hence,

$$d\theta_1 = -h_v du \wedge dv, \quad d\theta_2 = k_u du \wedge dv.$$

The first structure equations then become

$$-h_v du \wedge dv + k\omega_{12} \wedge dv = 0, \quad k_u du \wedge dv - h\omega_{12} \wedge du = 0.$$

These are clearly solved by

$$\omega_{12} = \frac{h_v}{k} du - \frac{k_u}{h} dv.$$

The second structure equation then gives

$$0 = d\omega_{12} = \left(-\frac{\partial}{\partial v}\frac{h_v}{k} - \frac{\partial}{\partial u}\frac{k_u}{h}\right) du \wedge dv,$$

which gives the result. ◆

Section 10: The structure equations for a surface

In this section we will apply the method of moving frames to the important case of a surface. This will give the the structure equations for a surface. These give a very efficient proof of a fundamental result in the theory of surfaces known as Gauss's Theorema Egregium.

10.1 Adapted frames and the structure equations

Let $\boldsymbol{x} : D \to \mathbb{E}^3$ be an oriented local surface.

Definition 10.1. *A moving frame* $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ *for* $\mathbb{E}^3$ *on* D *is said to be* **adapted** *to the surface if* $\boldsymbol{e}_3 = \boldsymbol{N}$*.*

Given a local surface $\boldsymbol{x} : D \to \mathbb{E}^3$ and an adapted frame we may define 1-forms $\theta_i = \boldsymbol{e}_i \cdot d\boldsymbol{x}$ and connection 1-forms $\omega_{ij} = \boldsymbol{e}_i \cdot d\boldsymbol{e}_j$ for $i, j = 1, 2, 3$ as before. Notice that since $\boldsymbol{N} \cdot d\boldsymbol{x} = 0$ then for such an adapted frame we have $\theta_3 = 0$ and hence

$$d\boldsymbol{x} = \boldsymbol{e}_1\theta_1 + \boldsymbol{e}_2\theta_2 \ .$$

Also, we have

$$d\boldsymbol{N} = \boldsymbol{e}_1\omega_{13} + \boldsymbol{e}_2\omega_{23} \ .$$

Proposition 10.2. *The first and second structure equations for a local surface with respect to an adapted frame, give the* **structure equations for a surface***:*

$$\left.\begin{aligned} d\theta_1 + \omega_{12} \wedge \theta_2 &= 0 \\ d\theta_2 - \omega_{12} \wedge \theta_1 &= 0 \end{aligned}\right\} \qquad \text{(First structure equations)}$$

$$\omega_{13} \wedge \theta_1 + \omega_{23} \wedge \theta_2 = 0 \qquad \text{(Symmetry equation)}$$

$$d\omega_{12} - \omega_{13} \wedge \omega_{23} = 0 \qquad \text{(Gauss equation)}$$

$$\left.\begin{aligned} d\omega_{13} + \omega_{12} \wedge \omega_{23} &= 0 \\ d\omega_{23} - \omega_{12} \wedge \omega_{13} &= 0 \end{aligned}\right\} \qquad \text{(Codazzi equations)}$$

Proof. Follows immediately from $\theta_3 = 0$ in the structure equations. We have grouped the equations according to their role in surface theory. □

Lemma 10.3. $\{\theta_1, \theta_2\}$ *is a basis for* $\Omega^1(D)$*.*

Proof. Linear independence of $\{\theta_1, \theta_2\}$ follows from that of $\boldsymbol{e}_1, \boldsymbol{e}_2$ and Proposition 6.5. □

Lemma 10.4. *There are unique functions* a, b, c *on* D *such that*

$$\begin{aligned} \omega_{13} &= a\theta_1 + b\theta_2 \\ \omega_{23} &= b\theta_1 + c\theta_2 \ . \end{aligned}$$

Proof. Since θ_1, θ_2 form a basis for the 1-forms on D at each point, there is certainly a unique expansion of the form $\omega_{13} = a\theta_1 + b\theta_2$ and $\omega_{23} = f\theta_1 + c\theta_2$. The symmetry equation is then

$$0 = \omega_{13} \wedge \theta_1 + \omega_{23} \wedge \theta_2 = b\theta_2 \wedge \theta_1 + f\theta_1 \wedge \theta_2 = (f - b)\theta_1 \wedge \theta_2$$

which is satisfied if and only if $f = b$. □

Proposition 10.5. *The first and second fundamental forms are*

$$I = \theta_1^2 + \theta_2^2, \qquad I\!I = -a\theta_1^2 - 2b\theta_1\theta_2 - c\theta_2^2.$$

The Gauss curvature K and the mean curvature H are given by

$$K = ac - b^2, \qquad H = -\frac{1}{2}(a+c).$$

Proof. For the first fundamental form, we compute

$$I = d\boldsymbol{x} \cdot d\boldsymbol{x} = (\theta_1 \boldsymbol{e}_1 + \theta_2 \boldsymbol{e}_2) \cdot (\theta_1 \boldsymbol{e}_1 + \theta_2 \boldsymbol{e}_2) = \theta_1^2 + \theta_2^2 \ .$$

For the second fundamental form, we compute

$$\begin{aligned} \mathrm{I\!I} &= -d\boldsymbol{x} \cdot d\boldsymbol{e}_3 = -(\boldsymbol{e}_1\theta_1 + \boldsymbol{e}_2\theta_2) \cdot (\boldsymbol{e}_1\omega_{13} + \boldsymbol{e}_2\omega_{23}) = -\theta_1\omega_{13} - \theta_2\omega_{23} \\ &= -\theta_1(a\theta_1 + b\theta_2) - \theta_2(b\theta_1 + c\theta_2) = -a\theta_1^2 - 2b\theta_1\theta_2 - c\theta_2^2 \ . \end{aligned}$$

Thus, with respect to the basis θ_1, θ_2 (more precisely, its dual basis) we see that I is represented by the identity matrix and that $\mathrm{I\!I}$ by the matrix

$$\begin{pmatrix} -a & -b \\ -b & -c \end{pmatrix} .$$

The principal curvatures are thus the eigenvalues of this matrix and the sum of the eigenvalues is the trace of the matrix and the product is the determinant. Thus

$$k_1 + k_2 = -(a+c) = 2H, \quad k_1k_2 = ac - b^2 = K.$$

□

Proposition 10.6. *The Gauss equation is equivalent to*

$$d\omega_{12} = K\theta_1 \wedge \theta_2 \ .$$

Proof. Substitute the expansions of Proposition 10.4 into the Gauss equation to get

$$d\omega_{12} = \omega_{13} \wedge \omega_{23} = (a\theta_1 + b\theta_2) \wedge (b\theta_1 + c\theta_2) = (ac - b^2)\theta_1 \wedge \theta_2$$

□

Remark 10.7. This result shows that the Gauss curvature can be computed simply from a knowledge of θ_1, θ_2 without reference to the local description of the surface $\boldsymbol{x} : D \to \mathbb{E}^3$. Next, we will show that in fact only I is needed to compute K.

Example 10.8. For the cylinder in example (5.8) again we have $d\boldsymbol{x} = \boldsymbol{e}_1 dz + \boldsymbol{e}_2 a d\phi$ where

$$\boldsymbol{e}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \boldsymbol{e}_2 = \begin{pmatrix} -\sin\phi \\ \cos\phi \\ 0 \end{pmatrix},$$

and hence $\theta_1 = dz$ and $\theta_2 = ad\phi$. It follows that $d\theta_1 = 0 = d\theta_2$. The first structure equations then force $\omega_{12} = 0$ and hence $K = 0$, as we had seen before.

10.2 Gauss's Theorema Egregium

First, we need to establish an important converse of the first part of Proposition 10.5.

Proposition 10.9. *Let $\boldsymbol{x} : D \to \mathbb{E}^3$ be a local surface with first fundamental form I and θ_1, θ_2 be 1-forms on D such that*

$$I = \theta_1^2 + \theta_2^2 \ .$$

Then there exists a unique adapted frame such that $\theta_1 = \boldsymbol{e}_1 \cdot d\boldsymbol{x}$ and $\theta_2 = \boldsymbol{e}_2 \cdot d\boldsymbol{x}$.

Proof. Let $(\tilde{\boldsymbol{e}}_1, \tilde{\boldsymbol{e}}_2, \boldsymbol{N})$ be an arbitrary adapted frame field and let $\tilde{\theta}_i = \tilde{\boldsymbol{e}}_i \cdot d\boldsymbol{x}$ be the corresponding one-forms, so $I = \tilde{\theta}_1^2 + \tilde{\theta}_2^2$. Since $\tilde{\theta}_1, \tilde{\theta}_2$ form a basis for $\Omega^1(D)$, we can relate the one-forms θ_i to $\tilde{\theta}_i$ linearly so, in matrix notation,

$$\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} = M \begin{pmatrix} \tilde{\theta}_1 \\ \tilde{\theta}_2 \end{pmatrix} ,$$

where M is a 2×2 invertible matrix whose entries are functions on D. Since by hypothesis $I = \theta_1^2 + \theta_2^2$ we deduce that M is an orthogonal matrix, so $M^{-1} = M^T$. Also, we can always arrange $\det M = +1$ (by swapping the labels on θ_i). Thus

$$d\boldsymbol{x} = \begin{pmatrix} \tilde{\boldsymbol{e}}_1 & \tilde{\boldsymbol{e}}_2 \end{pmatrix} \begin{pmatrix} \tilde{\theta}_1 \\ \tilde{\theta}_2 \end{pmatrix} = \begin{pmatrix} \tilde{\boldsymbol{e}}_1 & \tilde{\boldsymbol{e}}_2 \end{pmatrix} M^T \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} = \begin{pmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}$$

where we have defined

$$\begin{pmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 \end{pmatrix} = \begin{pmatrix} \tilde{\boldsymbol{e}}_1 & \tilde{\boldsymbol{e}}_2 \end{pmatrix} M^T$$

Since both $\begin{pmatrix} \tilde{\boldsymbol{e}}_1 & \tilde{\boldsymbol{e}}_2 \end{pmatrix}$ and M are orthogonal matrices with unit determinant, it follows that $\begin{pmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 \end{pmatrix}$ also is. Thus, the required oriented adapted frame is $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{N})$. □

Definition 10.10. *Two local surfaces $\boldsymbol{x} : D \to \mathbb{E}^3$ and $\tilde{\boldsymbol{x}} : D \to \mathbb{E}^3$ are* **isometric** *if $I = \tilde{I}$.*

Example 10.11. The cylinder (5.8) is isometric to a rectangular domain in a plane.

Remark 10.12. Isometric surfaces have the same measure of distance and angles. Sometimes one says isometric surfaces are related by "bending".

Theorem 10.13 (Gauss's Theorema Egregium). *Isometric surfaces have the same Gauss curvature.*

Proof. The result follows because the following algorithm for computing the Gauss curvature of a local surface depends only on I.

- Find 1-forms θ_1, θ_2 such that

  $$\boxed{\mathrm{I} = \theta_1^2 + \theta_2^2 \ .}$$

  This step is not unique, but is always possible and as shown above corresponds to a choice of adapted frame. In practice there will often be an "obvious" choice.

- Solve the first structure equations for a surface

$$\boxed{d\theta_1 + \omega_{12} \wedge \theta_2 = 0 \quad \text{and} \quad d\theta_2 - \omega_{12} \wedge \theta_1 = 0}$$

  for the connection form ω_{12}. Given θ_1, θ_2 this step is unique.

- Calculate K from the Gauss equation

$$\boxed{d\omega_{12} = K\theta_1 \wedge \theta_2 \ .}$$

  A tedious calculation shows that K does not depend on the choice of θ_1, θ_2.

□

This is a remarkable theorem which shows the Gauss curvature of a surface depends only on its first fundamental form I and not the explicit map $\boldsymbol{x} : D \to \mathbb{E}^3$. More succintly,

> “ The Gauss curvature is an intrinsic invariant of a surface ”

Remark 10.14. The converse of this theorem is not true. There are examples of surfaces with the same Gauss curvature which are however not isometric.

Example 10.15. The sphere of radius a in polar coordinates: $\mathrm{I} = a^2 d\alpha^2 + a^2 \sin^2 \alpha d\phi^2$. Clearly we can choose

$$\theta_1 = a d\alpha, \qquad \theta_2 = a \sin \alpha d\phi.$$

Hence $d\theta_1 = 0$ and $d\theta_2 = a \cos \alpha d\alpha \wedge d\phi$. The first structure equations thus reduce to $\omega_{12} \wedge \theta_2 = 0$ and $\omega_{12} \wedge \theta_1 = \cos \alpha\, \theta_1 \wedge d\phi$. Expanding ω_{12} in the θ_1, θ_2 basis (or $d\alpha, d\phi$ basis) shows that

$$\omega_{12} = -\cos \alpha d\phi.$$

Thus

$$d\omega_{12} = \sin \alpha d\alpha \wedge d\phi = a^{-2} \theta_1 \wedge \theta_2$$

and hence the Gauss equation gives

$$K = \frac{1}{a^2}.$$

10.3 Riemannian geometry

The above results suggest we can study the intrinsic geometry of a surface by working with just a first fundamental form. This leads to the (enormous) subject of Riemannian geometry. The idea is that one can consider a domain D (or some more complicated 2-dimensional object like a surface) and specify a first fundamental form (often called a **metric**) on D *without it necessarily having arisen from a map* $\boldsymbol{x} : D \to \mathbb{E}^3$. The idea generalises to higher dimensions. We will content ourselves with an example.

Definition 10.16. Two-dimensional hyperbolic space *is the upper half plane*

$$H = \{(x, y) \mid y > 0\} \subset \mathbb{R}^2$$

equipped with the first fundamental form given by

$$I = \frac{(dx)^2 + (dy)^2}{y^2}.$$

Theorem 10.17. *Hyperbolic space as above has $K = -1$.*

Proof. We take

$$\theta_1 = \frac{1}{y}dx, \quad \theta_2 = \frac{1}{y}dy$$

so that $\mathrm{I} = \theta_1^2 + \theta_2^2$. Calculating, the first structure equations become

$$\frac{1}{y^2}dx \wedge dy + \omega_{12} \wedge \frac{1}{y}dy = 0, \qquad 0 - \omega_{12} \wedge \frac{1}{y}dx = 0$$

which clearly are solved by

$$\omega_{12} = -\frac{1}{y}dx.$$

To derive this one can always expand in a basis $\omega_{12} = \alpha dx + \beta dy$ and substitute into the first structure equations. In this case the equation on the left implies $\beta = -1/y$ and the one on the right $\alpha = 0$. Thus,

$$d\omega_{12} = -\frac{1}{y^2}dx \wedge dy = -\theta_1 \wedge \theta_2$$

and the Gauss equation immediately gives $K = -1$. □

10.4 Exercises

Exercise 10.18. Consider the sphere of radius a in Exercise 5.20. Define an adapted frame by taking $\boldsymbol{e}_1, \boldsymbol{e}_2$ to be the unit vectors in the direction of $\boldsymbol{x}_\alpha, \boldsymbol{x}_\phi$ respectively. (You should check these are orthogonal.) Compute the forms $\theta_1, \theta_2, \omega_{jk}$ and check all the structure equations hold. Compute the functions a, b, c. Write down the first and second fundamental forms, and check that the formulae for H, K in terms of a, b, c agree with your previous computations. Check finally that $d\omega_{12} = K\theta_1 \wedge \theta_2$.

Solution. You should get

$$\theta_1 = a d\alpha, \quad \theta_2 = a \sin \alpha d\phi.$$

The connection forms are

$$\omega_{12} = -\cos\alpha\, d\phi, \quad \omega_{13} = d\alpha, \quad \omega_{23} = \sin \alpha d\phi.$$

Expanding, we get

$$\omega_{13} = \frac{1}{a}\theta_1, \quad \omega_{23} = \frac{1}{a}\theta_2.$$

Now using the formulae for H, K in terms of the coefficients here (don't get confused by the name of one of these being "a" which we are using here for the radius of the sphere) we get

$$H = -1/a, \quad K = 1/a^2.$$

◆

Exercise 10.19. Given a local surface defined by $\boldsymbol{x} : D \to \mathbb{E}^3$, consider the local surface defined by the scaling $\hat{\boldsymbol{x}}(u, v) = \Omega\, \boldsymbol{x}(u, v)$ where $\Omega > 0$ is a constant. Write the fundamental forms and various curvatures of $\hat{\boldsymbol{x}}$ in terms of those of $\boldsymbol{x}$.

Solution. Clearly one has $d\hat{\boldsymbol{x}} = \Omega d\boldsymbol{x}$ and thus $\hat{\mathrm{I}} = \Omega^2\mathrm{I}$. Also, $\hat{\boldsymbol{N}} = \boldsymbol{N}$ and thus $\hat{\mathrm{II}} = \Omega\mathrm{II}$. It follows that $\det(\hat{\mathrm{II}} - \hat{\lambda}\hat{\mathrm{I}}) = \Omega^2 \det(\mathrm{II} - \hat{\lambda}\Omega\mathrm{I})$ and therefore the principal curvatures are $\hat{k}_1 = k_1/\Omega$ and $\hat{k}_2 = k_2/\Omega$. Thus, $\hat{K} = \Omega^{-2}K$ and $\hat{H} = \Omega^{-1}H$. ◆

Exercise 10.20. Consider the (half) cone

$$x^2 + y^2 = z^2, \quad z > 0$$

in $\mathbb{E}^3$. Parametrise the cone (cylindrical polar coordinates are a good start) and show that the cone is flat. Why is this to be expected?

Solution. Parametrise the cone by

$$\boldsymbol{x}(u, \phi) = \begin{pmatrix} u\cos\phi \\ u\sin\phi \\ u \end{pmatrix}.$$

The first fundamental form (by calculation, or better by staring at a picture) is

$$\mathrm{I} = 2(du)^2 + u^2(d\phi)^2.$$

Thus we can choose

$$\theta_1 = \sqrt{2}du, \quad \theta_2 = ud\phi.$$

Then

$$d\theta_1 = 0, \quad d\theta_2 = du \wedge d\phi.$$

The first structure equations are

$$\omega_{12} \wedge d\phi = 0, \quad du \wedge d\phi - \omega_{12} \wedge \sqrt{2}du = 0.$$

These are clearly solved by

$$\omega_{12} = -\frac{1}{\sqrt{2}}d\phi$$

and $d\omega_{12} = 0$ and so $K = 0$.

This is to be expected because if you snip the cone open, you can clearly lay it flat. ◆

Exercise 10.21. Consider the standard surface of revolution

$$\boldsymbol{x}(u, \phi) = \begin{pmatrix} f(u)\cos\phi \\ f(u)\sin\phi \\ u \end{pmatrix}.$$

Writing $f_u = df/du$ argue geometrically (i.e. from arclength) that

$$\mathrm{I} = (1 + {f_u}^2)(du)^2 + f^2(d\phi)^2.$$

Taking the obvious forms $\theta_1 = \sqrt{1 + {f_u}^2}du$ and $\theta_2 = f d\phi$, show that

$$\omega_{12} = \frac{-f_u}{\sqrt{1 + {f_u}^2}} d\phi$$

and hence obtain a formula for the Gauss curvature.

Solution. Use the first structure equations to get ω_{12}. Explicitly, choosing θ_i as above we have $d\theta_1 = 0$ and $d\theta_2 = f_u du \wedge d\phi$. The first structure equation for $d\theta_1$ then tells us that $\omega_{12} \wedge d\phi = 0$ which implies ω_{12} only has a $d\phi$ component. The other first structure equation gives the coefficient of $d\phi$ in ω_{12} as above. Differentiating, we get

$$d\omega_{12} = \frac{-f_{uu}}{(1 + {f_u}^2)^{3/2}} du \wedge d\phi$$

and comparing with $\theta_1 \wedge \theta_2$ we get

$$K = \frac{-f_{uu}}{f(1 + {f_u}^2)^2}$$

◆

Exercise 10.22. Consider the unit disc $x^2 + y^2 < 1$ in the plane, and give it the abstract metric or first fundamental form given by $\mathrm{I} = (dr)^2 + r^2(d\phi)^2$ where (r, ϕ) are the usual polar coordinates. Show that $K = 0$. Why is this to be expected?

Solution. Take

$$\theta_1 = dr, \quad \theta_2 = r d\theta$$

so that the structure equations are

$$\omega_{12} \wedge d\theta = 0, \quad dr \wedge d\theta - \omega_{12} \wedge dr = 0.$$

The solution is $\omega_{12} = -d\theta$ and $d\omega_{12} = 0$ and so $K = 0$. *Note that what we are doing here is just the structure equations for a map to Euclidean space $\mathbb{E}^2$. The fact that $d\omega_{12} = 0$ is just the second structure equations for such maps.* ◆

Exercise 10.23. Take the unit disc $x^2 + y^2 < 1$ in the plane, and give it the abstract metric or first fundamental form given by

$$\frac{4((dr)^2 + r^2(d\phi)^2)}{(1 - r^2)^2},$$

where r, ϕ are the usual polar coordinates. Write I as $\mathrm{I} = \theta_1^2 + \theta_2^2$ and hence compute ω_{12} and K. You should get $K = -1$, just as for hyperbolic space described in lectures.

In fact, although it looks different, this is just another realisation of essentially the same space. The map which relates the two pictures is a Möbius transformation (as in complex analysis) that takes the disc to the upper half-plane. It is given by

$$z = x + iy \mapsto \frac{z - i}{z + i} = re^{i\phi}.$$

Regarding that as a change of coordinates, you should be able to check that the metrics are the same. (Challenge!)

Solution. Clearly
$$\theta_1 = \frac{2dr}{1-r^2}, \quad \theta_2 = \frac{2rd\phi}{1-r^2}$$
is the obvious choice. Differentiating and writing down the structure equations and solving for ω_{12} you should get
$$\omega_{12} = -\frac{1+r^2}{1-r^2}d\phi.$$
Using the Gauss equation, you get $K = -1$. ◆

Section 11: Geodesics

In this section we introduce the notion of geodesics. As we will show these correspond to curves which extremise arclength among "nearby" curves. We start with the simpler case of curves in Euclidean space.

11.1 Geodesics in Euclidean space

Consider a curve in Euclidean space $\boldsymbol{x} : [a,b] \to \mathbb{E}^n, t \mapsto \boldsymbol{x}(t)$, with unit speed $|\boldsymbol{x}'| = 1$, joining two points $\boldsymbol{x}(a)$, $\boldsymbol{x}(b) \in \mathbb{E}^n$. Consider a 1-parameter family of **nearby** curves

$$\boldsymbol{x}_\epsilon(t) = \boldsymbol{x}(t) + \epsilon \boldsymbol{y}(t)$$

where $\boldsymbol{x}'(t) \cdot \boldsymbol{y}(t) = 0$ and $\boldsymbol{y}(a) = \boldsymbol{y}(b) = 0$ so that all the curves in the family join $\boldsymbol{x}(a)$ to $\boldsymbol{x}(b)$. We refer to $\boldsymbol{y}$ as a **connecting vector**.

Remark 11.1. It is always possible to connect a curve to a sufficiently nearby curve with a connecting vector $\boldsymbol{y}(t)$ satisfying $\boldsymbol{x}'(t) \cdot \boldsymbol{y}(t) = 0$. This is because, if $\boldsymbol{y}$ has a component along $\boldsymbol{x}'$ this could be removed by reparameterisation of the curve.

Definition 11.2. *We say a unit speed curve $\boldsymbol{x}(t)$ as above has* **stationary length** *if the length of nearby curves $s_\epsilon = \int_a^b |\boldsymbol{x}_\epsilon'| dt$ satisfies*

$$\left.\frac{ds_\epsilon}{d\epsilon}\right|_{\epsilon=0} = 0$$

for all connecting vectors $\boldsymbol{y}(t)$.

Proposition 11.3. *A unit speed curve $\boldsymbol{x}(t)$ in Euclidean space has stationary length if and only if it is the straight line joining the two points.*

Proof. The speed of the tangent to our family of curves $\boldsymbol{x}_\epsilon$ is given by

$$|\boldsymbol{x}_\epsilon'| = \sqrt{1 + 2\boldsymbol{x}' \cdot \boldsymbol{y}'\epsilon + \epsilon^2 |\boldsymbol{y}'|^2} = 1 + \boldsymbol{x}' \cdot \boldsymbol{y}'\epsilon + O(\epsilon^2).$$

Thus

$$\begin{aligned} \left.\frac{ds_\epsilon}{d\epsilon}\right|_{\epsilon=0} &= \left.\left(\frac{d}{d\epsilon}\int_a^b |\boldsymbol{x}_\epsilon'| dt\right)\right|_{\epsilon=0} = \int_a^b \boldsymbol{x}' \cdot \boldsymbol{y}' dt \\ &= \Big[\boldsymbol{x}' \cdot \boldsymbol{y}\Big]_a^b - \int_a^b \boldsymbol{x}'' \cdot \boldsymbol{y} dt = -\int_a^b \boldsymbol{x}'' \cdot \boldsymbol{y} dt \end{aligned}$$

where the boundary terms vanish because $\boldsymbol{y}(a) = \boldsymbol{y}(b) = 0$. The final expression vanishes for all such $\boldsymbol{y}$ if and only if $\boldsymbol{x}'' = 0$ (because $\boldsymbol{x}'' \cdot \boldsymbol{x}' = 0$ since $\boldsymbol{x}(t)$ is unit speed). Hence $\boldsymbol{x}'' = 0$ which integrates to give the straight line between the two points. □

Remark 11.4. The above is a "Calculus of Variations" argument. We have only shown that straight lines extremise the arclength. Of course in Euclidean space they in fact minimise arclength.

11.2 Geodesics on surfaces

Suppose that we have a unit speed curve $\boldsymbol{x}(c(t))$ lying in a surface connecting two given points. We wish to consider infinitesimally nearby curves also lying in the surface. We can do this by imposing the additional restriction that the connecting vector $\boldsymbol{y}(t)$ is everywhere tangent to the surface. Hence we say that a curve $\boldsymbol{x}(c(t))$ on a surface has **stationary length** if

$$\left.\frac{ds_\epsilon}{d\epsilon}\right|_{\epsilon=0} = 0$$

for all connecting vectors $\boldsymbol{y}(t)$ satisfying the conditions in the previous section and such that $\boldsymbol{y}(t) \cdot \boldsymbol{N} = 0$.

Definition 11.5. *A unit-speed curve $\boldsymbol{x}(c(t))$ lying in a surface is a* **geodesic** *if its acceleration is everywhere normal to the surface, that is,*

$$\frac{d^2}{dt^2}\boldsymbol{x}(c(t)) = A(t)\boldsymbol{N}(c(t))$$

where $\boldsymbol{N}$ is the unit normal to the surface and A is some function along the curve.

Remark 11.6. This definition states that for a geodesic the acceleration in the directions tangent to the surface vanishes thus generalising the concept of a straight line in a plane.

Proposition 11.7. *A curve lying in a surface has stationary length (among nearby curves on the surface joining the same endpoints) if and only if it is a geodesic.*

Proof. The proof is exactly as in Proposition 11.3 except that the final integral has to vanish for all $\boldsymbol{y}(t)$ satisfying the additional constraint $\boldsymbol{y} \cdot \boldsymbol{N} = 0$. This is so if and only if $\boldsymbol{x}''$ is everywhere normal to the surface. □

Example 11.8. For a sphere the geodesics are always part of the **great circles**. The great circles are closed curves on the sphere defined by the intersection of the sphere and any plane going through its centre. It is easy to see that the acceleration of a curve on a great circle is normal to the sphere and hence correspond to geodesics.

Proposition 11.9. *A curve $\boldsymbol{x}(c(t))$ lying in a surface is a geodesic if and only if, in an adapted frame it obeys the* **geodesic equations**

$$\boxed{\begin{aligned} &\frac{d}{dt}(\theta_1(c')) + \omega_{12}(c')\theta_2(c') = 0 \\ &\frac{d}{dt}(\theta_2(c')) - \omega_{12}(c')\theta_1(c') = 0 \end{aligned}}$$

and the **energy equation**

$$\boxed{(\theta_1(c'))^2 + (\theta_2(c'))^2 = 1\ .}$$

Proof. The energy equation is simply the unit speed condition $\boldsymbol{x}' \cdot \boldsymbol{x}' = \mathrm{I}(c', c') = 1$. For a curve to have $\boldsymbol{x}''$ normal to the surface it is necessary and sufficient that in an adapted frame $\boldsymbol{e}_1 \cdot \boldsymbol{x}'' = 0$ and $\boldsymbol{e}_2 \cdot \boldsymbol{x}'' = 0$. Computing, we get

$$\begin{aligned} \boldsymbol{e}_1 \cdot \boldsymbol{x}'' &= \frac{d}{dt}(\boldsymbol{e}_1 \cdot \boldsymbol{x}') - \boldsymbol{e}_1' \cdot \boldsymbol{x}' = \frac{d}{dt}(\boldsymbol{e}_1 \cdot d\boldsymbol{x}(c')) - d\boldsymbol{e}_1(c') \cdot d\boldsymbol{x}(c') \\ &= \frac{d}{dt}(\theta_1(c')) - \omega_{21}(c')\boldsymbol{e}_2 \cdot d\boldsymbol{x}(c') = \frac{d}{dt}(\theta_1(c')) + \omega_{12}(c')\theta_2(c') \end{aligned}$$

where in the second line we used the definitions $\theta_i = \boldsymbol{e}_i \cdot d\boldsymbol{x}$ and $d\boldsymbol{e}_i = \sum_{j=1}^{3} \boldsymbol{e}_j \omega_{ji}$. The same calculation for $\boldsymbol{e}_2 \cdot \boldsymbol{x}''$ gives the second geodesic equation. □

Example 11.10. Let us solve for the geodesics of the cylinder of radius a. Recall $\theta_1 = dz$ and $\theta_2 = ad\phi$ and that $\omega_{12} = 0$. Hence the geodesic equations for a curve $c(t) = (z(t), \phi(t))$ are $z'' = 0$ and $\phi'' = 0$. This means $z = z(0) + z'(0)t$ and $\phi = \phi(0) + \phi'(0)t$, where $z'(0)$ and $\phi'(0)$ are constants constrained by the energy equation to satisfy $z'(0)^2 + a^2\phi'(0)^2 = 1$. The curves in the domain are straight lines. This is not surprising since the cylinder is flat. On the cylinder the curves are in general spirals.

Proposition 11.11. *The fact that the "energy"*

$$E = (\theta_1(c'))^2 + (\theta_2(c'))^2$$

is constant is a consequence of the geodesic equations. Setting it to one just means we get curves of unit speed, rather than some other constant speed.

Proof.

$$E' = 2\theta_1(c')\frac{d}{dt}(\theta_1(c')) + 2\theta_2(c')\frac{d}{dt}(\theta_2(c')) = 0$$

where the geodesic equations have been used in the last equality. □

Proposition 11.12. *Given a point $\boldsymbol{p}$ on a surface and a unit tangent vector $\boldsymbol{v}$ to the surface at $\boldsymbol{p}$, there exists a unique geodesic on the surface $t \mapsto \boldsymbol{x}(c(t))$ for $|t| < \epsilon$ (with ϵ sufficiently small), such that $\boldsymbol{x}(0) = \boldsymbol{p}$ and $\boldsymbol{x}'(0) = \boldsymbol{v}$.*

Proof. The geodesic equations are a pair of second order differential equations for $\boldsymbol{x}(c(t))$. The given initial conditions ensure a unique solution exists near $\boldsymbol{p}$ by the usual theorem for ODEs. □

Remark 11.13. Given two points on a surface which is a closed subset of $\mathbb{E}^3$, there exists a geodesic joining those points. The closed condition is to exclude pathologies like taking the (x, y)-plane and removing the origin, so there is no geodesic between $(1, 0)$ and $(-1, 0)$.

Notice that the geodesic equations depend only on the first fundamental form of a surface. Hence they are part of the intrinsic geometry of a surface and isometric surfaces have the same geodesics. This also means that we can define them for abstract surfaces like the hyperbolic plane.

Example 11.14. Take the hyperbolic plane H as in Definition 10.16, with

$$\mathrm{I} = \frac{(dx)^2 + (dy)^2}{y^2} \ .$$

We can write $\mathrm{I} = \theta_1^2 + \theta_2^2$ where $\theta_1 = \frac{1}{y}dx$ and $\theta_2 = \frac{1}{y}dy$ and we have computed $\omega_{12} = -\frac{1}{y}dx$. Thus the geodesic equations for a curve $c(t) = (x(t), y(t))$ with tangent vector $c' = x'\frac{\partial}{\partial x} + y'\frac{\partial}{\partial y}$ are therefore

$$\left(\frac{x'}{y}\right)' - \frac{x'y'}{y^2} = 0 \qquad \text{and} \qquad \left(\frac{y'}{y}\right)' + \frac{x'^2}{y^2} = 0 \ ,$$

and the energy equation is

$$\frac{x'^2}{y^2} + \frac{y'^2}{y^2} = 1.$$

To solve this system notice that the first geodesic equation is equivalent to

$$\frac{x'}{y^2} = A = \text{constant} \ .$$

If $A = 0$ then the equations are straight lines $x = $ constant, otherwise substitute into the energy equation to get

$$\left(\frac{dy}{dx}\right)^2 = \frac{1}{A^2y^2} - 1$$

and solve by separation of variables to get

$$(x + C)^2 + y^2 = \frac{1}{A^2} \ ,$$

which is a circle of radius $1/|A|$.

11.3 Exercises

Exercise 11.15. Show that for a ruled surface $\boldsymbol{x}(u, v) = \boldsymbol{z}(u) + v\boldsymbol{p}(u)$ the curves defined by $v \mapsto \boldsymbol{x}(u_0, v)$ for any constant u_0 are geodesics.

Solution. Such curves have zero acceleration and so in particular zero acceleration in directions tangent to the surfaces. Hence they must be geodesics. ◆

Exercise 11.16. Consider the unit sphere parameterised in polar coordinates (α, ϕ). Using the definition of a geodesic on a surface show that the equator $\alpha(t) = \pi/2, \phi(t) = t$ is a geodesic. Also show that the lines of longitude $\alpha(t) = t, \phi(t) = \phi_0$ where ϕ_0 is a constant are geodesics.

Exercise 11.17. Derive the geodesic equations and the energy equation for the sphere, with the usual forms θ_1, θ_2. Show that $\sin^2\alpha\,\phi' = A = $ constant for any solution. Using the energy equation show that $\sin^2\alpha \geq A^2$. Deduce that the only geodesics that join the poles of the sphere are lines of longitude.

Solution. From previously,

$$\mathrm{I} = r^2(d\alpha)^2 + r^2\sin^2\alpha(d\phi)^2$$

and

$$\theta_1 = r d\alpha, \quad \theta_2 = r\sin\alpha d\phi.$$

and

$$\omega_{12} = -\cos\alpha d\phi.$$

Thus, for a curve with $c(t) = (\alpha(t), \phi(t))$ we have

$$\theta_1(c') = r\alpha', \quad \theta_2(c') = r\sin\alpha\phi', \quad \omega_{12}(c') = -\cos\alpha\phi'.$$

The geodesic equations thus become

$$\alpha'' - \sin\alpha\cos\alpha(\phi')^2 = 0$$
$$\frac{d}{dt}(\sin\alpha\phi') + \cos\alpha\,\alpha'\phi' = 0.$$

The energy equation becomes

$$(\alpha')^2 + \sin^2\alpha(\phi')^2 = \frac{1}{r^2}.$$

From the second geodesic equation,

$$(\sin^2\alpha\,\phi')' = \sin\alpha(\sin\alpha\,\phi')' + \alpha'\cos\alpha(\sin\alpha\,\phi') = 0$$

and so

$$\sin^2\alpha\,\phi' = A = \text{constant}$$

for any solution. ◆

Exercise 11.18. In the "disc model" of hyperbolic space discussed in Exercise 10.23, the geodesics are lines through the centre and arcs of circles that cut the boundary orthogonally. Can you show that? (Remark: if you did Complex Analysis, it follows from properties of Möbius transformations without any further calculation!)

Exercise 11.19. Consider a local surface such that $\boldsymbol{x}_u \cdot \boldsymbol{x}_v = 0$. Choose the adapted frame such that

$$\boldsymbol{x}_u = p\boldsymbol{e}_1, \quad \boldsymbol{x}_v = q\boldsymbol{e}_2$$

for positive functions p, q on D. Calculate θ_1, θ_2 and hence find an expression for K in terms of p, q and their u and v derivatives.

One can think of the curves of constant v as being a family of curves parametrised by u. The vector $\boldsymbol{x}_v$ is then a "connecting vector" connecting you orthogonally to a nearby curve of constant v just as in our discussion of geodesics. Note that $q = |\boldsymbol{x}_v|$. Suppose now that we have chosen our parametrisation of the surface so that the curves of constant v are the images of geodesics. Thus, we have that for some function f that the curves

$$u(t) = f(t, v), \quad v = \text{constant}$$

are geodesics. Write down the geodesic and energy equations and show that $p_v = 0$. Deduce that $q' = q_u f' = q_u/p$. Differentiate again and deduce that

$$q'' + Kq = 0.$$

Since q is the "distance to an infinitesimally neighbouring geodesic", we see that positive K has the effect of pulling diverging geodesics back together and negative K pushes them apart.

Solution. Since $d\boldsymbol{x} = p\boldsymbol{e}_1 du + q\boldsymbol{e}_2 dv$ one gets $\theta_1 = pdu$ and $\theta_2 = qdv$. Hence $d\theta_1 = -p_v du \wedge dv$ and $d\theta_2 = q_u du \wedge dv$. The first structure equations then imply

$$\omega_{12} = \frac{p_v}{q} du - \frac{q_u}{p} dv$$

and the Gauss equation gives

$$K = -\frac{1}{pq}\left[\left(\frac{p_v}{q}\right)_v + \left(\frac{q_u}{p}\right)_u\right]$$

The curves $c(t) = (f(t, v_0), v_0)$ have tangent vectors $c' = f' \frac{\partial}{\partial u}$. Hence $\theta_1(c') = pf'$ and $\theta_2(c') = 0$ and $\omega_{12}(c') = \frac{p_v}{q} f'$. The first geodesic equation then says $(pf')' = 0$ and the second geodesic equations says $\frac{p_v}{q} f' p f' = 0$. The latter tells us that $p_v = 0$ as required. The energy equation is just $p^2 f'^2 = 1$ and so $f' = 1/p$ (choosing a sign). Hence $q' = q_u u' + q_v v' = q_u f' = q_u/p$. Finally, $q'' = (q_u/p)_u f' = (q_u/p)_u p^{-1} = -Kq$, where in the last equality we used the expression for K (and $p_v = 0$). ◆

Section 12: Integration over surfaces

12.1 Integration of 2-forms over surfaces

Earlier we saw that one can integrate 3-forms over domains in $\mathbb{R}^3$ and 1-forms over curves in $\mathbb{R}^3$. Thus we expect to be able to integrate 2-forms over surfaces. This is indeed the case. We have previously considered surfaces in Euclidean space, but here the Euclidean structure is unnecessary.

Definition 12.1. *Let $x : D \to \mathbb{R}^3$ define a local surface*

$$x(u, v) = (x_1(u, v), x_2(u, v), x_3(u, v)).$$

(Note we do not write the map defining the surface in bold here, to emphasise we are not going to use the Euclidean structure.) Let

$$\alpha = \alpha_{12} dx_1 \wedge dx_2 + \alpha_{13} dx_1 \wedge dx_3 + \alpha_{23} dx_2 \wedge dx_3$$

be a 2-form on $\mathbb{R}^3$. We define the **pull-back** *$x^*\alpha$ of α by the map x to be the 2-form on D given by*

$$x^*\alpha = \alpha_{12}(x(u, v))dx_1 \wedge dx_2 + \alpha_{13}(x(u, v))dx_1 \wedge dx_3 + \alpha_{23}(x(u, v))dx_2 \wedge dx_3$$

where here dx_k is the exterior derivative of $x_k : D \to \mathbb{R}$ so $dx_k = \frac{\partial x_k}{\partial u} du + \frac{\partial x_k}{\partial v} dv$. (Note carefully this distinction: dx_k stands for the basis 1-form in $\mathbb{R}^3$ in the definition of the form α, but it stands for "d" of the function x_k on D in the definition of the pull-back.)

Definition 12.2. *Let $x : D \to \mathbb{R}^3$ be a local surface and let α be a 2-form on $\mathbb{R}^3$. We define the* **integral** *of α over the local surface to be*

$$\boxed{\int_{x(D)} \alpha = \int_D x^*\alpha}$$

(assuming the integral on the right hand-side exists).

Definition 12.3. *If Σ is an oriented regular surface, the integral over Σ of a 2-form α is defined by writing Σ as a union of images of regular oriented local surfaces, and adding the integrals of α over these local surfaces. In many cases, e.g. the cylinder or sphere, one can find a single local surface that covers all except a curve – this is sufficient.*

Example 12.4. Consider the sphere of radius a in the usual polar coordinates $x(\alpha, \phi) = (a \sin\alpha \cos\phi, a \sin\alpha \sin\phi, a \cos\alpha)$ and the two form $\alpha = x_1 dx_2 \wedge dx_3$. The pull-back of α under the map x is

$$\begin{aligned} x^*\alpha &= a \sin\alpha \cos\phi d(a \sin\alpha \sin\phi) \wedge d(a \cos\alpha) \\ &= a \sin\alpha \cos\phi(a \cos\alpha \sin\phi d\alpha + a \sin\alpha \cos\phi d\phi) \wedge (-a \sin\alpha d\alpha) \\ &= -a^3 \sin^3\alpha \cos^2\phi d\phi \wedge d\alpha = a^3 \sin^3\alpha \cos^2\phi d\alpha \wedge d\phi \end{aligned}$$

Hence the integral of α over the local surface is

$$\int_D x^*\alpha = \int_{\alpha=0}^{\alpha=\pi} \int_{\phi=0}^{\phi=2\pi} a^3 \sin^3\alpha \cos^2\phi \; d\alpha d\phi = \frac{4\pi a^3}{3} \,.$$

Definition 12.5. *A 1-dimensional submanifold of* $\mathbb{R}^3$ *is a regular curve, a 2-dimensional submanifold of* $\mathbb{R}^3$ *is a regular surface, a 3-dimensional submanifold of* $\mathbb{R}^3$ *is a domain.*

A central result in the theory of integration on surfaces is Stokes' theorem, which we will state without proof. For this we need to discuss the notion of boundaries.

Definition 12.6. *Let* Σ *be an oriented surface in* $\mathbb{R}^3$ *with boundary an oriented curve* C*. The* **induced orientation** *on* C *is such that if* (u_1, u_2) *is an oriented coordinate chart on* Σ *such that* $u_1 \leq 0$ *and* $u_1 = 0$ *is the boundary, then* u_2 *is an oriented coordinate on* C*. Let* M *be a 3-dimensional submanifold of* $\mathbb{R}^3$ *bounded by an oriented surface* Σ*. The induced orientation on* Σ *is such that if* (u_1, u_2, u_3) *is an oriented coordinate chart on* M *such that* $u_1 \leq 0$ *and* $u_1 = 0$ *is the boundary, then* (u_2, u_3) *are oriented coordinates on* Σ*.*

Example 12.7. Take the inside of a simple (i.e. non-self-intersecting), closed curve C in the plane with its standard orientation $dx_1 \wedge dx_2$. The induced orientation on C is dx_2, and corresponds to traversing C anti-clockwise. This can be seen by drawing oriented coordinate lines for the plane near a point on the curve in such a way that inside the curve is $x_1 < 0$ and the curve itself is given by $x_1 = 0$. Then you can see x_2 increasing as you go round the curve anti-clockwise.

Theorem 12.8 (Stokes)**.** *Let* Σ *be a* k*-dimensional oriented closed and bounded submanifold in* $\mathbb{R}^3$ *with boundary* $\partial\Sigma$ *given the induced orientation and* $\alpha \in \Omega^{k-1}(\Sigma)$*. Then*

$$\boxed{\int_\Sigma d\alpha = \int_{\partial\Sigma} \alpha.}$$

Remark 12.9. The $k = 1$ case of Stokes' theorem reduces to our earlier result

$$\int_c df = f(c(a)) - f(c(b))$$

for an oriented curve $c : [a, b] \to \mathbb{R}^3$ and function f (∂c is the two endpoints $c(a)$ and $c(b)$).

Remark 12.10. The Stokes' and divergence of vector calculus are the $k = 2$ and $k = 3$ special cases respectively, see exercises (12.28) and (12.29).

Example 12.11. We may calculate $\int_\Sigma x_1 dx_2 \wedge dx_3$ for any oriented bounded surface Σ, by employing Stokes' Theorem. Since the exterior derivative of the integrand is $dx_1 \wedge dx_2 \wedge dx_3$ we have $\int_\Sigma x_1 dx_2 \wedge dx_3 = \int_M dx_1 \wedge dx_2 \wedge dx_3$ where M is the region bounded by Σ. Therefore, the integral is simply the volume of the region bounded by Σ.

Remark 12.12. Stokes' theorem also holds for k-dimensional submanifolds of $\mathbb{R}^n$.

Corollary 12.13 (Green's theorem in the plane)**.** *Let* $D \subset \mathbb{R}^2$ *with boundary curve* C*. The induced orientation on* C *from the standard orientation on* D *is anti-clockwise. Then Stokes' theorem for the 1-form* $\alpha = P(x, y)dx + Q(x, y)dy$ *is*

$$\int_C (Pdx + Qdy) = \int_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx \wedge dy.$$

Example 12.14. Consider a simple (i.e. non self-intersecting) closed curve C in the plane, oriented anti-clockwise. The integral $\int_C xdy$ can be calculated using Stokes' Theorem (or Green's Theorem). Thus $\int_C xdy = \int_D dx \wedge dy$, where D is the region bounded by C. This integral thus gives the area enclosed by the curve C.

12.2 Integration of functions over surfaces

Given a local surface $\boldsymbol{x} : D \to \mathbb{E}^3$, its area is given by

$$A = \int_D |\boldsymbol{x}_u \times \boldsymbol{x}_v| \, dudv \; .$$

Intuitively, the quantity $|\boldsymbol{x}_u \times \boldsymbol{x}_v| dudv$ is the area of an infinitessimal parallelogram-shaped piece of surface obtained by incrementing u, v by du, dv and so the integral gives the sum of these, hence the total area. More generally, given a function $f : \mathbb{E}^3 \to \mathbb{R}$, its integral over the surface is

$$\int_{\boldsymbol{x}(D)} f \quad \equiv \quad \int_D f(\boldsymbol{x}(u,v)) |\boldsymbol{x}_u \times \boldsymbol{x}_v| \, dudv \; .$$

The reason we can integrate a function over a surface, rather than a 2-form over a surface, is that the given notion of distance in Euclidean space means we also have a notion of area.

Lemma 12.15. *For a local surface $|\boldsymbol{x}_u \times \boldsymbol{x}_v| = \sqrt{\det I}$. Hence, we obtain an alternate expression for the area*

$$A = \int_D \sqrt{\det I} \, dudv \; .$$

Thus the area depends only on I and hence is an intrinsic property of the surface.

Proof. Recall the vector identity $|\boldsymbol{x}_u \times \boldsymbol{x}_v|^2 = |\boldsymbol{x}_u|^2 |\boldsymbol{x}_v|^2 - (\boldsymbol{x}_u \cdot \boldsymbol{x}_v)^2 = EF - G^2$, where E, F, G are the components of the first fundamental form. $\square$

Lemma 12.16. *For a local surface in $\mathbb{E}^3$ with adapted frame,*

$$\theta_1 \wedge \theta_2 = |\boldsymbol{x}_u \times \boldsymbol{x}_v| du \wedge dv.$$

Proof. Consider $d\boldsymbol{x} \overset{\wedge}{\times} d\boldsymbol{x}$ where the notation means that the vector cross product is to be taken, but with the entries wedged to make a vector of 2-forms. On the one hand,

$$d\boldsymbol{x} = \boldsymbol{x}_u du + \boldsymbol{x}_v dv$$

and we get

$$d\boldsymbol{x} \overset{\wedge}{\times} d\boldsymbol{x} = 2(\boldsymbol{x}_u \times \boldsymbol{x}_v) du \wedge dv.$$

On the other hand,

$$d\boldsymbol{x} = \theta_1 \boldsymbol{e}_1 + \theta_2 \boldsymbol{e}_2$$

and we get

$$d\boldsymbol{x} \overset{\wedge}{\times} d\boldsymbol{x} = 2(\theta_1 \wedge \theta_2) \boldsymbol{e}_3 \; .$$

But since we are in an adapted frame $\boldsymbol{x}_u \times \boldsymbol{x}_v = |\boldsymbol{x}_u \times \boldsymbol{x}_v| \boldsymbol{e}_3$, which gives result. $\square$

Corollary 12.17. *Let $\boldsymbol{x} : D \to \mathbb{E}^3$ be a local surface and $f : \mathbb{E}^3 \to \mathbb{R}$ be a function. Then the integral of f over the surface is given by*

$$\int_{\boldsymbol{x}(D)} f = \int_D f(\boldsymbol{x}(u,v)) \theta_1 \wedge \theta_2$$

In particular,

$$\boxed{A = \int_D \theta_1 \wedge \theta_2}$$

gives the area of the local surface. The 2-form $\theta_1 \wedge \theta_2$ is called the **area form**.

Example 12.18. For a sphere of radius a, in spherical polar coordinates, we had $\theta_1 = d\alpha$ and $\theta_2 = a \sin \alpha d\phi$. Hence the area form is $\theta_1 \wedge \theta_2 = a^2 \sin \alpha \, d\alpha \wedge d\phi$. Thus, the area is

$$A = \int_D a^2 \sin \alpha \, d\alpha \wedge d\phi = a^2 \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} \sin \alpha \, d\alpha d\phi = 4\pi a^2$$

Remark 12.19. The **normal variation** of a local surface $\boldsymbol{x} : D \to \mathbb{E}^3$, is the family of local surfaces defined by

$$\boldsymbol{x}_\epsilon(u, v) = \boldsymbol{x}(u, v) + \epsilon f(u, v) \boldsymbol{N}(u, v)$$

where f is a smooth function on D which vanishes outside some closed bounded subset. It can be shown that a local surface has **stationary area**

$$\left. \frac{dA_\epsilon}{d\epsilon} \right|_{\epsilon=0} = 0$$

for all f, where A_ϵ is the area of the normal variation, if and only if it is a minimal surface $H = 0$. This (partly) explains the definition of a minimal surface.

12.3 Exercises

Exercise 12.20. Let $x(\alpha, \phi) = (\sin \alpha \cos \phi, \sin \alpha \sin \phi, \cos \alpha)$ be the usual polar coordinate map for the unit sphere. Calculate $x^*(dx_1 \wedge dx_2)$. Hence show the integral of $dx_1 \wedge dx_2$ over the sphere is zero.

Solution. We have

$$x_1(\alpha, \phi) = \sin \alpha \cos \phi$$

and so

$$dx_1 = \cos \alpha \cos \phi d\alpha - \sin \alpha \sin \phi d\phi.$$

Similarly,

$$x_2(\alpha, \phi) = \sin \alpha \sin \phi$$

and so

$$dx_2 = \cos \alpha \sin \phi d\alpha + \sin \alpha \cos \phi d\phi.$$

Thus

$$\begin{aligned} x^*(dx_1 \wedge dx_2) &= (\cos \alpha \cos \phi d\alpha - \sin \alpha \sin \phi d\phi) \wedge (\cos \alpha \sin \phi d\alpha + \sin \alpha \cos \phi d\phi) \\ &= \sin \alpha \cos \alpha d\alpha \wedge d\phi. \end{aligned}$$

Thus the required integral becomes

$$\int_{\phi=0}^{2\pi} \int_{\alpha=0}^{\pi} \sin \alpha \cos \alpha \, d\alpha d\phi = 0.$$

◆

Exercise 12.21. Integrate

$$x_1 dx_2 \wedge dx_3 + x_2 dx_3 \wedge dx_1 + x_3 dx_1 \wedge dx_2$$

over a region of the sphere of radius a. (Use polar coordinates again.)

Solution. From the previous solution,

$$x^*(x_3 dx_1 \wedge dx_2) = \sin\alpha \cos^2\alpha \, d\alpha \wedge d\phi.$$

We have $x_3 = \cos\alpha$ and so

$$dx_3 = -\sin\alpha d\alpha.$$

Thus,

$$x^*(x_1 dx_2 \wedge dx_3) = \sin^3\alpha \cos^2\phi d\alpha \wedge d\phi$$

and

$$x^*(x_2 dx_3 \wedge dx_1) = \sin^3\alpha \sin^2\phi d\alpha \wedge d\phi.$$

Adding the three terms, we get

$$\sin\alpha d\alpha \wedge d\phi.$$

Thus, the integral will become

$$\int\int \sin\alpha d\alpha d\phi$$

which you should recognise as the integral that gives the area of whatever region of the sphere you are integrating over. ◆

Exercise 12.22. Integrate the 3-form $dx_1 \wedge dx_2 \wedge dx_3$ over the region $B = \{x_1^2 + x_2^2 + x_3^2 \leq 1\} \subset \mathbb{R}^3$, by changing to spherical polar coordinates. By appealing to Cartesian coordinates interpret your answer.

Solution. From a previous exercise we have $dx_1 \wedge dx_2 \wedge dx_3 = r^2 \sin\alpha \, dr \wedge d\alpha \wedge d\phi$. The region B is polar coordinates is simply $r \leq 1$. Hence

$$\int_B dx_1 \wedge dx_2 \wedge dx_3 = \int_B r^2 \sin\alpha \, dr \wedge d\alpha \wedge d\phi = \int_{r=0}^{1} \int_{\alpha=0}^{\pi} \int_{\phi=0}^{2\pi} r^2 \sin\alpha dr d\alpha d\phi = \frac{4\pi}{3}$$

which is of course the volume of the region B. ◆

Exercise 12.23. Consider the polar coordinate map in the plane:

$$x(r,\theta) = (r\cos\theta, r\sin\theta).$$

Show that pulling back a form by this map is the same as changing coordinates from Cartesian to polar.

Solution. You should get that

$$x^* dx = \cos\theta dr - r\sin\theta d\theta,$$

etc., which is exactly the same computation as changing coordinates. ◆

Exercise 12.24. Let

$$H^n = \{x \in \mathbb{R}^n \mid x_1 \leq 0\}$$

be the lower half space (oriented by $dx_1 \wedge \cdots \wedge dx_n$). The **boundary** ∂H^n of H^n is

$$\partial H^n = \mathbb{R}^{n-1} = \{x \in \mathbb{R}^n \mid x_1 = 0\}$$

which we will orient by $dx_2 \wedge \cdots \wedge dx_n$. Let α be an $n-1$ form on H^n, vanishing identically outside of some closed and bounded subset of H^n. Prove the identify

$$\int_{H^n} d\alpha = \int_{\partial H^n} \alpha$$

Focus on the $n = 2$ case.

Solution. The restriction $\alpha|_{\partial H^n} = \alpha_{2\ldots n}(0, x_2, \ldots, x_n) dx_2 \wedge \cdots \wedge dx_n$ is an $(n-1)$-form on the $(n-1)$-dimensional space ∂H^n. Then

$$\int_{\partial H^n} \alpha = \int_{\mathbb{R}^{n-1}} \alpha_{2\ldots n}(0, x_2, \ldots, x_n) dx_2 \wedge \cdots \wedge dx_n.$$

follows from our definition of integration. (Note that it is just one of the n terms in the general $(n-1)$-form that contributes and that term is evaluated at the boundary $x_1 = 0$).

Let us take $n = 2$ for simplicity (the generalisation to $n > 2$ is straightforward). Since α vanishes outside a sufficiently large box, we can take our limits to be

$$-R \leq x_2 \leq R, \quad -R \leq x_1 \leq 0$$

for some large R, where all the coefficients of the forms vanish at all the limits of integration except $x_1 = 0$. We have $\alpha = \alpha_1 dx_1 + \alpha_2 dx_2$ and by the preceding Lemma

$$\int_{\partial H} \alpha = \int_{-R}^{R} \alpha_2(0, x_2) dx_2.$$

Now, since

$$d\alpha = \left(\frac{\partial \alpha_2}{\partial x_1} - \frac{\partial \alpha_1}{\partial x_2} \right) dx_1 \wedge dx_2$$

we have

$$\int_H d\alpha = \int_{x_2=-R}^{R} \int_{x_1=-R}^{0} \left(\frac{\partial \alpha_2}{\partial x_1} - \frac{\partial \alpha_1}{\partial x_2} \right) dx_1 dx_2 \; .$$

The first term, on doing the x_1-integral, gives $\int_{-R}^{R} \alpha_2(0, x_2) dx_2$. The second term (doing the x_2 integral first) gives zero.

◆

Exercise 12.25. Taking the inside of the unit sphere with the usual orientation, show that the induced orientation on the sphere corresponds to taking the outward pointing normal.

Solution. If you take coordinates near a point on the surface with $x_1 < 0$ inside, and $x_1 = 0$ on the sphere, then x_2, x_3 provide coordinates on the sphere. If x_1, x_2, x_3 are oriented in $\mathbb{R}^3$ then taking $u = x_2, v = x_3$ as parameters for the surface, $\boldsymbol{x}_u \times \boldsymbol{x}_v$ is outward pointing. ◆

Exercise 12.26. Verify Stokes's theorem for the inside of the unit sphere bounded by the unit sphere, where

$$\alpha = x_1 dx_2 \wedge dx_3.$$

Recall example (12.4).

Solution. From a previous example, this form pulls back in polar coordinates to be

$$\sin^3\alpha\cos^2\phi d\alpha\wedge d\phi.$$

Thus

$$\int_{\text{surface of sphere}}\alpha=\int_{\alpha=0}^{\pi}\int_{\phi=0}^{2\pi}\sin^3\alpha\cos^2\phi d\alpha d\phi=\frac{4\pi}{3}.$$

On the other hand, we readily compute

$$d\alpha=dx_1\wedge dx_2\wedge dx_3$$

and so

$$\int_{\text{inside sphere}}d\alpha=\text{Vol of sphere}=\frac{4\pi}{3}.$$

◆

Exercise 12.27. Verify Stokes's theorem for the upper hemisphere $x_3\geq 0$ bounded by the unit circle in the $x_3=0$ plane, where

$$\alpha=x_2dx_1.$$

Solution. Easiest here is to parametrise the hemisphere as a graph

$$x(u,v)=\begin{pmatrix}u\\ v\\ \sqrt{1-u^2-v^2}\end{pmatrix}$$

so that $u^2+v^2\leq 1$ parameterise a unit disk. Then $d\alpha=-dx_1\wedge dx_2$ pulls back to give $-du\wedge dv$ so that the integral is $-\pi$. Spherical polars should also work. (The region of integration should be $0\leq\phi\leq 2\pi$, $0\leq\alpha\leq\pi/2$.)

The integral of α over the curve can be performed as follows. Parametrise the curve as $(x_1,x_2)=(\cos\phi,\sin\phi)$. Then, $\alpha=x_2dx_1=-\sin^2\phi d\phi$. Integrating with respect to to ϕ from 0 to 2π gives $-\pi$ as well. ◆

Exercise 12.28. Let $S\subset\mathbb{R}^3$ be an oriented surface which is the boundary of a region $V\in\mathbb{R}^3$, so $\partial V=S$. Let $\beta=\beta_1dx_2\wedge dx_3+\beta_2dx_3\wedge dx_1+\beta_3dx_1\wedge dx_2$ be a two-form defined on a submanifold of $\mathbb{R}^3$ which contains V. Show in this case Stokes theorem $\int_V d\beta=\int_S\beta$ can be rewritten as the divergence theorem of vector calculus $\int_V \operatorname{div}\boldsymbol{v}\,dV=\int_S\boldsymbol{v}\cdot d\mathbf{S}$.

Solution. Define a one to one map between vectors in $\mathbb{E}^3$ and 2-forms on $\mathbb{R}^3$ by $\boldsymbol{v}=(\beta_1,\beta_2,\beta_3)$. One can check that $d\beta=\operatorname{div}\boldsymbol{v}\;dx_1\wedge dx_2\wedge dx_3$. The surface element normal to the surface may be written as $d\mathbf{S}=(\boldsymbol{x}_u\times\boldsymbol{x}_v)\,du\wedge dv=(1/2)d\boldsymbol{x}\overset{\wedge}{\times}d\boldsymbol{x}$ which means take the cross product and wedge the entries of the vector. Hence $\boldsymbol{v}\cdot d\mathbf{S}=\beta$. ◆

Exercise 12.29. Let S be an oriented surface with a boundary curve C with the induced orientation. Let α be a 1-form defined on S. Show that in this case Stokes theorem $\int_S d\alpha=\int_C\alpha$ is nothing but Stokes theorem of vector calculus $\int_S\operatorname{curl}\boldsymbol{v}\cdot d\mathbf{S}=\int_C\boldsymbol{v}\cdot d\mathbf{l}$.

Solution. As in the previous exercise we define a one to one correspondence between vectors in $\mathbb{E}^3$ and 2-forms on $\mathbb{R}^3$. We also have an obvious correspondence between vectors in $\mathbb{E}^3$ and 1-forms on $\mathbb{R}^3$ given by $\boldsymbol{v} = (\alpha_1, \alpha_2, \alpha_3)$ so $\boldsymbol{v} \cdot d\mathbf{l} = \sum_i \alpha_i dx_i = \alpha$. Computing the 2-form $d\alpha$ we see that $d\alpha = \operatorname{curl} \boldsymbol{v} \cdot d\mathbf{S}$. ◆

Exercise 12.30. Integrate the function $f(\boldsymbol{x}) = z$ over the unit sphere in $\mathbb{E}^3$.

Solution. For the unit sphere we have

$$\theta_1 = d\alpha, \quad \theta_2 = \sin\alpha d\phi.$$

Thus,

$$\theta_1 \wedge \theta_2 = \sin\alpha d\alpha \wedge d\phi.$$

Also, $z = \cos\alpha$ and so the integral becomes

$$\int_D \cos\alpha \sin\alpha d\alpha \wedge d\phi = \int_{\alpha=0}^{\pi} \int_{\phi=0}^{2\pi} \frac{1}{2} \sin 2\alpha \, d\alpha d\phi = 0.$$

(You should expect this by symmetry.) ◆

Exercise 12.31. What is the area of the region in hyperbolic space given by $a < x < b$ and $p < y < q$, where p, q are both positive? Show that the area becomes infinite as p tends to zero.

Solution. For hyperbolic space,

$$\theta_1 = \frac{1}{y} dx, \quad \theta_2 = \frac{1}{y} dy.$$

Thus,

$$\theta_1 \wedge \theta_2 = \frac{1}{y^2} dx \wedge dy.$$

The area is thus

$$\int_D \frac{1}{y^2} dx \wedge dy = \int_{x=a}^{b} \int_{y=p}^{q} \frac{1}{y^2} dx dy = (b-a)\left(\frac{1}{p} - \frac{1}{q}\right).$$

Clearly this becomes infinite as $p \to 0$. (You should expect this — remember the x-axis is infinitely far away from points in hyperbolic space.) ◆

Exercise 12.32. The surface of revolution with profile curve $f(u) = \cosh u$ is a **catenoid**. Show it is a minimal surface. Compute the area of this surface of revolution between $u = -1$ and $u = +1$. Show that it is less than the area of the cylinder whose boundary is the same two circles.

Solution. Direct calculation shows that $H = 0$. As shown in exercise (10.21), for the standard surface of revolution we have

$$\theta_1 = \sqrt{1 + f'^2} du, \quad \theta_2 = f d\phi$$

and so

$$\theta_1 \wedge \theta_2 = f\sqrt{1 + f'^2} du \wedge d\phi.$$

In this case, $f(u) = \cosh u$ and this becomes

$$\cosh^2 u du \wedge d\phi.$$

Thus the area is

$$\int_{u=-1}^{1}\int_{\phi=0}^{2\pi}\cosh^2 u du d\phi = 2\pi\int_{u=-1}^{1}\cosh^2 u du = \frac{\pi}{2}(e^2 - e^{-2} + 4) \approx 17.68$$

On the other hand, the cylinder with the same boundary has radius $\cosh 1$ and length 2, and so has area

$$4\pi\cosh 1 \approx 19.39$$

which is indeed greater than the corresponding catenoid. Intuitively it is clear the catenoid minimises the area of a surface with the above given boundary. This would be the shape a soap film makes. ◆

Exercise 12.33.☆ Let $A(\epsilon) = \int_D \sqrt{\det \mathrm{I}_\epsilon}\, dudv$ be the area of the local surface $\boldsymbol{x}_\epsilon : D \to \mathbb{E}^3$ defined by the normal variation of a local surface $\boldsymbol{x} : D \to \mathbb{E}^3$ given above. Show that

$$A'(0) = -2\int_D fH\sqrt{\det \mathrm{I}}\, dudv\ .$$

Hence deduce that $A'(0) = 0$ for all normal variations (i.e. all f) if and only if $H = 0$.

Solution. The area $A(\epsilon) = \int_D \sqrt{\det \mathrm{I}_\epsilon}\, dudv$ where $\mathrm{I}_\epsilon = d\boldsymbol{x}_\epsilon \cdot d\boldsymbol{x}_\epsilon$. Since $d\boldsymbol{x}_\epsilon = d\boldsymbol{x} + \epsilon(f d\boldsymbol{N} + \boldsymbol{N} df)$ we have

$$\mathrm{I}_\epsilon = \mathrm{I} + 2\epsilon f\, d\boldsymbol{x}\cdot d\boldsymbol{N} + O(\epsilon^2) = \mathrm{I} - 2\epsilon f \mathrm{II} + O(\epsilon^2).$$

Hence

$$\det \mathrm{I}_\epsilon = \det \mathrm{I} \det(1 - 2\epsilon f \mathrm{I}^{-1}\mathrm{II} + O(\epsilon^2)) = \det \mathrm{I}\, [1 - 2\epsilon f \mathrm{Tr}(\mathrm{I}^{-1}\mathrm{II}) + O(\epsilon^2)]$$

where we have used the general relation $\det(\mathrm{Id} + \epsilon M) = 1 + \epsilon \mathrm{Tr} M + O(\epsilon^2)$. Noting that $2H = \mathrm{Tr}(\mathrm{I}^{-1}\mathrm{II})$ we deduce that $\sqrt{\det \mathrm{I}_\epsilon} = \sqrt{\det \mathrm{I}}\,(1 - 2\epsilon f H + O(\epsilon^2))$. Finally, $A'(0) = \int_D \frac{\partial\sqrt{\det \mathrm{I}_\epsilon}}{\partial \epsilon}\Big|_{\epsilon=0} dudv = -2\int_D fH\sqrt{\det \mathrm{I}}\, dudv$. ◆

Section 13: Gauss-Bonnet theorem

In this section we will prove the Gauss-Bonnet theorem, a profound result which relates the local geometry of a surface to its topology.

13.1 Geodesic triangles

Definition 13.1. *A* **geodesic triangle** Δ *in a surface is three points each joined to the other two by geodesics. We assume our triangles are sufficiently small that they fit in the image of a single coordinate chart (i.e. a local surface) on which we have an adapted frame.*

Proposition 13.2. *Let* $t \mapsto \boldsymbol{x}(c(t))$, $t \in [a, b]$, *be a geodesic in a local surface and* $(\boldsymbol{e}_1, \boldsymbol{e}_2)$ *be an adapted moving frame for the surface. Let* $\theta(t)$ *be such that*

$$\boldsymbol{x}' = \cos(\theta(t))\boldsymbol{e}_1 + \sin(\theta(t))\boldsymbol{e}_2$$

along the geodesic (we can write this since it is a unit speed curve). Then

$$\theta'(t) = \omega_{12}(c'(t)).$$

Proof. Recall for a geodesic $\boldsymbol{e}_1 \cdot \boldsymbol{x}'' = 0$ and $\boldsymbol{e}_2 \cdot \boldsymbol{x}'' = 0$. Computing

$$\begin{aligned}
\boldsymbol{e}_2 \cdot \boldsymbol{x}'' &= \boldsymbol{e}_2 \cdot (\cos\theta\, \boldsymbol{e}_1 + \sin\theta\, \boldsymbol{e}_2)' \\
&= \boldsymbol{e}_2 \cdot (\cos\theta\, \boldsymbol{e}_1' - \theta' \sin\theta\, \boldsymbol{e}_1 + \sin\theta\, \boldsymbol{e}_2' + \theta' \cos\theta\, \boldsymbol{e}_2) \\
&= \boldsymbol{e}_2 \cdot (\cos\theta\, d\boldsymbol{e}_1(c') + \theta' \cos\theta\, \boldsymbol{e}_2) = \cos\theta\, (-\omega_{12}(c') + \theta').
\end{aligned}$$

Repeating we find a similar identity $\boldsymbol{e}_1 \cdot \boldsymbol{x}'' = \sin\theta\ (\omega_{12}(c') - \theta')$. Since $\cos\theta$ and $\sin\theta$ cannot both be zero, the result now follows. □

Theorem 13.3. *Let* Δ *be a geodesic triangle on a local surface with interior angles* α, β, γ. *Assume* Δ *is contractible to a point. Then*

$$\int_\Delta K\theta_1 \wedge \theta_2 = \alpha + \beta + \gamma - \pi.$$

Proof. Recall that $d\omega_{12} = K\theta_1 \wedge \theta_2$ (Gauss). Stokes' Theorem can be then applied

$$\int_\Delta K\theta_1 \wedge \theta_2 = \int_\Delta d\omega_{12} = \int_{\partial\Delta} \omega_{12}$$

where $\partial\Delta$ is the boundary of the triangle Δ, with the induced orientation anticlockwise. The fact that Δ is contractible to a point ensures that the only boundaries to Δ are the three sides of the triangle C_1, C_2, C_2 so that $\partial\Delta = C_1 \cup C_2 \cup C_3$. Along each side of the triangle we have

$$\int \omega_{12} = \int_a^b \omega_{12}(c'(t))dt = \int_a^b \theta'(t)dt = \theta(b) - \theta(a)$$

which gives the difference in the values of θ at the two ends of the side. By the definition of $\theta(t)$ above, this gives the total angle of rotation of the unit tangent vector with respect to the frame field $\boldsymbol{e}_1, \boldsymbol{e}_2$ along that side.

As we traverse the whole triangle back to a chosen starting point, the unit tangent rotates by $(\pi-\alpha), (\pi-\beta), (\pi-\gamma)$ at each of the three vertices and by $\int_{C_1}\omega_{12}, \int_{C_2}\omega_{12}, \int_{C_3}\omega_{12}$ along each side. The total rotation as one goes round must be 2π and thus one concludes that

$$\int_{\partial\Delta}\omega_{12} = 2\pi - ((\pi-\alpha)+(\pi-\beta)+(\pi-\gamma)) = \alpha+\beta+\gamma-\pi.$$

□

Remark 13.4. This result only depends on the intrinsic geometry of the surface and hence is also valid for abstract surfaces like hyperbolic space.

Corollary 13.5. *Let Δ be a geodesic triangle with angles α, β, γ and area $A(\Delta)$. We have:*

1. *on the plane, $\alpha+\beta+\gamma=\pi$;*
2. *on the unit sphere $\alpha+\beta+\gamma=\pi+A(\Delta)>\pi$;*
3. *on hyperbolic space $\alpha+\beta+\gamma=\pi-A(\Delta)<\pi$.*

Example 13.6. Consider a geodesic triangle on the sphere of radius R with two vertices on the equator and one at the North Pole. Suppose the angle at the North Pole is α. The triangle is $\alpha/(2\pi)$ of the Northern hemisphere, and so it has area $\frac{\alpha}{2\pi}2\pi R^2 = \alpha R^2$. For the sphere, $K = 1/R^2$ (see earlier) and so the integral of K over the triangle is α. On the other hand, the angles in the triangle are α plus two right-angles. Thus the sum of the angles minus π is also equal to α, as required.

13.2 The Gauss–Bonnet Theorem

In this section Σ is an oriented closed and bounded surface with no boundary.

Definition 13.7. *Dissect Σ into polygons (triangles, quadrilaterals etc.). Let V be the number of vertices, E the number of edges and F the number of faces (polygons) of the dissection. The* **Euler characteristic** *of the dissected surface is*

$$\chi(\Sigma) = V - E + F.$$

Proposition 13.8. *The Euler characteristic of Σ is independent of the particular dissection chosen. Furthermore, it is unchanged by "reasonable" deformations of the surface.*

Proof. We will not prove this, or define what "reasonable" means (roughly, these are deformations that do not "tear" the surface). But it is not hard to see that χ is unchanged if polygons are subdivided to obtained a more refined dissection. □

Example 13.9. The Euler characteristic of a sphere is 2. One way to see this is to project a tetrahedron onto the sphere. The tetrahedron has 4 faces, 6 edges and 4 vertices.

Definition 13.10. *The* **torus** *is the surface of revolution whose profile is a circle $(f(u)-a)^2+u^2=b^2$ with $b<a$. The Euler characteristic of the torus is* 0.

Theorem 13.11 (Gauss–Bonnet)**.** *Let Σ be an oriented closed and bounded surface with no boundary. Then*

$$\boxed{\chi(\Sigma) = \frac{1}{2\pi}\int_\Sigma K.}$$

Proof. Dissect the surface into geodesic triangles. Since each triangle has 3 edges, and each edge is shared by 2 triangles we have $E = 3F/2$ and hence $\chi = V - \frac{F}{2}$. Now, over each geodesic triangle we have

$$\int_\Delta K = \alpha + \beta + \gamma - \pi$$

and so

$$\int_\Sigma K = (\text{the sum of the interior angles of all the triangles}) - \pi F.$$

But, the sum of the interior angles that appear round each vertex is 2π and so the total sum of all the interior angles is $2\pi V$. Thus

$$\int_\Sigma K = 2\pi V - \pi F = 2\pi\chi.$$

□

Remark 13.12. What is so remarkable is that if you deform a surface, K changes, yet the total integral remains unchanged. This Theorem is the origin of many results in modern geometry where a local geometric invariant is integrated to give a topological invariant.

Example 13.13. For a sphere $K = 1/a^2$ where $a > 0$ is its radius. Then $\int_\Sigma K = K \times 4\pi a^2 = 4\pi$ (since K is a constant this integral is simply K times the area). Therefore $\chi = 2$ as required. (Note this is independent of the radius a).

Example 13.14. The flat torus is defined as a parallelogram in a plane whose opposite edges are identified, equipped with the flat metric of the plane. This geometry cannot arise as a local surface in $\mathbb{E}^3$ (it can in $\mathbb{E}^4$!). We may now verify the Gauss-Bonnet theorem. Since $K = 0$ we immediately get $\chi = 0$ as required.

Corollary 13.15. *Suppose Σ is any surface obtained by a smooth deformation of a sphere, so $\chi = 2$. Then there must be a point on Σ where $K > 0$.*

Proof. By the Gauss Bonnet Theorem $\int_\Sigma K = 4\pi > 0$. This implies there exists a point on Σ such that $K > 0$. □

Corollary 13.16. *Suppose Σ is any surface obtained by a smooth deformation of a torus, so $\chi = 0$. Then the Gauss curvature of Σ must change sign.*

Proof. We have $\int_\Sigma K = 0$, so K cannot be everywhere positive or everywhere negative. □

13.3 Exercises

Exercise 13.17. A bear emerges from her cave, walks a mile South, then a mile West, then a mile North and finds she is back at her cave. What colour is the bear?

Solution. The bear must be at the North Pole. Hence it must be white. ◆

Exercise 13.18. Take $a > 0$ and find the area of the region in hyperbolic space bounded by the three geodesics $(x \pm a/2)^2 + y^2 = a^2/4$ and $x^2 + y^2 = a^2$. (Recall $y > 0$.) If you regard the region as a sort of geodesic triangle with all three vertices at infinity, does the result relating its angles to its area make any sort of sense?

Solution. From earlier, the area form in the hyperbolic plane is

$$\frac{1}{y^2} dx \wedge dy.$$

Take the half of the triangle which has $x > 0$. The area S of this is

$$\begin{aligned} S &= \int_{x=0}^{a} \int_{y=\sqrt{ax-x^2}}^{\sqrt{a^2-x^2}} \frac{1}{y^2} dxdy \\ &= \int_{x=0}^{a} \left[-\frac{1}{y} \right]_{\sqrt{ax-x^2}}^{\sqrt{a^2-x^2}} dx \\ &= \int_{0}^{a} \left(1/\sqrt{ax-x^2} - 1/\sqrt{a^2-x^2} \right) dx \\ &= \pi/2 \end{aligned}$$

Thus the area of the whole triangle is π and thus the integral of K over it is $-\pi$. Thus the standard result holds if all the angles are taken to be zero. This would seem to be a sensible value if you draw a picture, but of course the Theorem we proved does not immediately apply to these sorts of infinite triangles. ◆

Exercise 13.19. Show that for a geodesic quadrilateral $\square$ we have

$$\int_{\square} K\theta_1 \wedge \theta_2 = \Sigma - 2\pi.$$

where Σ is the sum of the four interior angles of $\square$.

Solution. Either duplicate the proof for the triangle to arrive at the total rotation around the quadrilateral being

$$\int_{\square} \omega_{12} + (\pi - \alpha) + (\pi - \beta) + (\pi - \gamma) + (\pi - \delta)$$

or perhaps more simply cut the quadrilateral into two triangles and add. ◆

Exercise 13.20. Consider the surface of revolution of a function $f(u)$ between the points $z = a$ and $z = b$. Suppose that f has a local minimum at a and a local maximum at b. By considering a geodesic quadrilateral consisting of half of the lines of latitude at $z = a$ and $z = b$ and the lines of longitude joining their ends, show that the integral of K over the region of surface between $z = a$ and $z = b$ is zero.

Solution. All four angles are right-angles, and thus by the previous exercise the integral of K over its inside must be zero. This is half the integral over the whole surface between $z = a$ and $z = b$. ◆

Exercise 13.21. What is the percentage error in assuming that for a geodesic triangle of area 1 square metre on the earths surface, the sum of the angles is π? (The radius of the earth is about 6400km.)

Solution. The angle surplus is the integral of K over the inside. For a sphere of radius R we have $K = 1/R^2$. Thus for the earth we have

$$K = \frac{1}{(6.4 \times 10^6)^2} = 2.4 \times 10^{-14}\text{metres}^{-2}.$$

Integrating over a square metre, we get an angle deficit of 2.4×10^{-14} radians. The percentage error is thus 7.8×10^{-13} percent! ◆

Exercise 13.22. Compute the Euler characteristic of the dissections of the sphere defined by projecting each of the platonic solids (recall these are the tetrahedron, cube, octahedron, dodecahedron and isocahedron) on to the sphere.

Solution. The tetrahedron has 4 faces, 6 edges and 4 vertices and defines a triangulation of the sphere. The cube has 6 faces, 12 edges and 8 vertices and dissects the sphere into quadrilaterals. The octahedron has 8 faces, 12 edges and 6 vertices and triangulates the sphere. The dodecahedron has 12 faces, 30 edges and 20 vertices and dissects the sphere into pentagons. The isocahedron has 20 faces, 30 edges and 12 vertices and triangulates the sphere. In all cases $\chi = V - E + F = 2$. ◆

Exercise 13.23. Compute the Euler characteristic of the torus. It is easiest to do this by describing the torus abstractly as a parallelogram with opposite edges identified.

Solution. The defining parallelogram for the flat torus gives $F = 1$. The number of edges in $E = 2$ because opposite sides are identified. The number of vertices is $V = 1$ due to the identifications. Hence $\chi = V - E + F = 0$ as required for a torus. ◆

Exercise 13.24. Show that if there is a regular solid whose faces are pentagons and such that 3 meet at each vertex, then there must be 12 faces. Similarly, if the faces are triangles and 5 meet at each vertex, show there are 20 faces.

Solution. Let there be F pentagons. Each edge is common to two pentagons and so we have $5F/2$ edges. Each vertex is common to three pentagons and so we have $5F/3$ vertices. Using the Euler characteristic, we have

$$2 = F - \frac{5F}{2} + \frac{5F}{3} = \frac{F}{6}.$$

The other is completely analogous. ◆

Exercise 13.25. A football is made by sewing together pentagons and hexagons, with three pieces meeting at each vertex. How many pentagons are there? If one assumes also that one pentagon and two hexagons meet at each vertex, how many hexagons are there?

Solution. Let there be p pentagons and h hexagons. Thus $F = p+h$. We have $(5p+6h)/2$ edges and $(5p+6h)/3$ vertices. The Euler characteristic gives

$$2 = (p+h) - \frac{5p+6h}{2} + \frac{5p+6h}{3} = \frac{p}{6}.$$

Thus we have $p = 12$ but no immediate constraint on the number of hexagons.

Suppose now we have 2 hexagons and a pentagon meeting at each vertex. From above, putting $p = 12$ we have $20+2h$ vertices. If at each vertex one of the three meeting faces is a pentagon then we have $(20+2h)/5$ pentagons in all. So $(20+2h)/5 = 12$ and so $h = 20$. This is the standard football. It is aso the arrangement of the carbon atoms in the famous "bucky-ball" molecule $C_{60}H_{60}$. You can make one by taking an icosahedron and carefully planing down each of the 12 vertices to make a truncated icosahedron. ◆

Exercise 13.26. Given a local surface $\boldsymbol{x}$ as usual, one can regard the unit normal $\boldsymbol{e}_3$ itself as defining a parametrisation of a piece of the unit sphere. (This is called the "Gauss map".) Show that

$$d\boldsymbol{e}_3 \overset{\wedge}{\times} d\boldsymbol{e}_3 = 2\omega_{13} \wedge \omega_{23}\boldsymbol{e}_3$$

and hence deduce that the area form for the Gauss map is

$$K\theta_1 \wedge \theta_2$$

where K is the Gauss curvature of our original surface. Hence show that if the Gauss map of an oriented, compact surface without boundary S is a bijection, then S has Euler characteristic 2.

Solution. This is a fairly standard "forms and structure equations" computation. If the Gauss map is a bijection, then the integral of K over the original surface must equal the area of the unit sphere which is 4π. Thus the surface has Euler characteristic 2. ◆

Appendix

In differential geometry we often need to use a few theorems and constructions from advanced multi-variable calculus and linear algebra that unfortunately are frequently not discussed in analysis or algebra courses. We state them here for completeness. Note that this is not part of the examinable material for the course.

Note that the exercises given below are an integral part of the discussion. Ideally, you should do these exercises, but if you do not you should still read through their statements.

Section A: The inverse and implicit function theorems

A.1 The inverse function theorem: one dimensional case

Theorem A.1. *Let $f : (a,b) \subset \mathbb{R} \to \mathbb{R}$ be a smooth[4] real-valued function defined on an interval. If, for some $p \in (a,b)$, we have that $f'(p) \neq 0$, then there exists a subinterval $(c,d) \subset (a,b)$ containing p such that f restricted to (c,d) is invertible, and the inverse function is also smooth. Moreover, if $g : f(c,d) \to (c,d)$ is the inverse function, then*

$$g'(f(p)) = \frac{1}{f'(p)}.$$

Exercise A.2. Show that, if $f : (a,b) \to \mathbb{R}$ is a smooth function such that $f(p) \neq 0$ for all $p \in (a,b)$, then all of f has a smooth inverse function.

A.2 General case

Theorem A.3. *Let*

$$F = (f^1, \dots, f^n) : D \subset \mathbb{R}^n \to \mathbb{R}^n : (x^1, \dots, x^n) \mapsto (f^1(x^1, \dots, x^n), \dots, f^n(x^1, \dots, x^n))$$

be a smooth function defined on an open set[5] $D \subset \mathbb{R}^n$ to $\mathbb{R}^n$, with Jacobian matrix

$$J_F(p) = \left[\frac{\partial f^i}{\partial x^j}\Big|_p \right].$$

If, at any point $p \in D$, $J_F(p)$ is invertible as a matrix, then there exists an open set $\widetilde{D} \subset D \subset \mathbb{R}^n$ containing p such that $F|_{\widetilde{D}}$ is invertible. Moreover, if $G : E \subset \mathbb{R}^n \to \mathbb{R}^n$ is the inverse:

$$G \circ F\big|_{\widetilde{D}} = \mathrm{Id}_{\widetilde{D}} \ \text{ and } F\big|_{\widetilde{D}} \circ G = \mathrm{Id}_E$$

with E an open subset containing $F(p)$, then

$$J_G(F(p)) = \Big(J_F(p)\Big)^{-1}.$$

The Inverse Function Theorem guarantees that locally an inverse exists if the conditions are satisfied. It doesn't give you a formula for it though – and in general determining the inverse is very difficult, often impossible. We shall not give proofs of these theorems here; the most common technique to prove them is the *contracting mapping theorem*, which is discussed in the Honours Analysis course.

[4] In fact, it suffices for f to be differentiable with continuous derivative, in which case the same will be true of the inverse. In this course we are only concerned with smooth functions though.

[5] If you haven't heard of the notion of *open set* you can just read *open ball* here.

A.3 The implicit function theorem

Theorem A.4 (Implicit Function Theorem)**.** *Let*

$$F = (f^1, \dots, f^n) : D \subset \mathbb{R}^{n+m} \to \mathbb{R}^m : (x^1, \dots, x^n, y^1, \dots, y^m) \mapsto F(x^1, \dots, x^n, y^1, \dots, y^m)$$

be a smooth function defined on an open set $D \subset \mathbb{R}^{n+m}$, with Jacobian matrix

$$J_F = \begin{bmatrix} \frac{\partial f^1}{\partial x^1} & \cdots & \frac{\partial f^1}{\partial x^n} & \frac{\partial f^1}{\partial y^1} & \cdots & \frac{\partial f^1}{\partial y^m} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f^m}{\partial x^1} & \cdots & \frac{\partial f^m}{\partial x^n} & \frac{\partial f^m}{\partial y^1} & \cdots & \frac{\partial f^m}{\partial y^m} \end{bmatrix}.$$

If, at some point fixed $p \in D$, we have that the matrix

$$\begin{bmatrix} \left.\frac{\partial f^1}{\partial y^1}\right|_p & \cdots & \left.\frac{\partial f^1}{\partial y^m}\right|_p \\ \vdots & \ddots & \vdots \\ \left.\frac{\partial f^m}{\partial y^1}\right|_p & \cdots & \left.\frac{\partial f^m}{\partial y^m}\right|_p \end{bmatrix}$$

is invertible, then there exists a smooth function

$$\Phi : E \subset \mathbb{R}^n \to \mathbb{R}^m$$

(the implicit function*) defined on an open set $E \subset \mathbb{R}^n$, such that the following hold:*

1. *There exists an open set $\widetilde{D}$ with $p \in \widetilde{D} \subset D \subset \mathbb{R}^{n+m}$ of p, such that*

$$\widetilde{D} \cap \left(\mathbb{R}^n \subset \mathbb{R}^{n+m}\right) = E,$$

and such that for all $\widetilde{p} \in \widetilde{D}$ we have

$$F(\widetilde{p}) = F(p) \quad \Leftrightarrow \quad \widetilde{p} = \left(x^1, \dots, x^n, \Phi(x^1, \dots, x^n)\right) \quad \text{for some} \quad (x^1, \dots, x^n) \in E;$$

2. *for all $q \in E$, we have that the Jacobian matrix is given by*

$$J_\Phi(q) = -\begin{bmatrix} \frac{\partial f^1}{\partial y^1} & \cdots & \frac{\partial f^1}{\partial y^m} \\ \vdots & \ddots & \vdots \\ \frac{\partial f^m}{\partial y^1} & \cdots & \frac{\partial f^m}{\partial y^m} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial f^1}{\partial x^1} & \cdots & \frac{\partial f^1}{\partial x^n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f^m}{\partial x^1} & \cdots & \frac{\partial f^m}{\partial x^n} \end{bmatrix}.$$

A typical application of this theorem is in the case when $m = 1$, so we have a single real-valued function defined on $\mathbb{R}^{n+1}$. If, near some point $p \in \mathbb{R}^{n+1}$ the Jacobian matrix of this function is non-zero (i.e. not all partial derivatives vanish), the the level set will be a *smooth hypersurface.*

Section B: New vector spaces from old: introduction to multi-linear algebra

In differential geometry, it is often useful to think of vector spaces in a rather abstract way, without choosing a basis that identifies them with $\mathbb{R}^n$ or $\mathbb{C}^n$. Moreover, a number of constructions that *create new vector spaces out of old ones* are also used a lot, perhaps implicitly. Some of these may be familiar to you, but others will likely be new, so this appendix serves as a general introduction.

Throughout we will work with vectorspaces over a fixed base field K – for all practical purposes you can just think of this as the field of real numbers $\mathbb{R}$. For our purposes we will restrict to finite-dimensional vector spaces, even though a lot goes through without that assumption. Likewise, in the last section we will indicate how a lot of these constructions can be generalized to modules over more general commutative rings, not just vectorspaces over fields.

B.1 Notes on *canonical*

A main thread through the notes below is that we really need to think about linear algebra in a more abstract way, which often means not making reference to a basis. Related to that we shall use the term *canonical*, for instance when talking about an isomorphism between two vector spaces. Intuitively, this terms means *natural, not involving any choices.* It can often be made precise in terms of category theory, functoriality, and natural transformations, but will not use this language. This word is used a lot in (pure) mathematics, but the etymology of the term comes from religion – we could vaguely interpret it as *god-given.*

B.2 Dual vector spaces

B.2.1 Definition and basic properties

With each vector space V over K, we can associate a new vector space V^*, also over K, which consists just of the linear functionals on V:

$$V^* = \{\phi : V \to K \mid \phi \text{ is } K\text{-linear}\}.$$

If you choose a basis $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ for V, you automatically get a basis $\alpha^1, \ldots, \alpha^{\mathbf{n}}$ for V^*, where the new basis elements are defined by

$$\alpha^i(\mathbf{w}_j) = \delta^i_j.$$

Here δ^i_j is the Kronecker delta:

$$\delta^i_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

We refer to the new basis $\alpha^1, \ldots, \alpha^n$ for V^* as the *dual basis* to the basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ for V. Remark that you need to know all of the $\mathbf{v}_j$ to define each α^i.

Exercise B.1. Verify that the α^i indeed form a basis for V^*. As a corollary, deduce that, for a finite-dimensional vector space, $\dim V = \dim V^*$.

Exercise B.2. Show that, for a finite dimensional vector space, there exists a canonical isomorphism $V \cong (V^*)^*$ (note that this is not true generally for infinite dimensional spaces, there you only have a canonical injection $V \hookrightarrow (V^*)^*$).

Exercise B.3. Show that, given a linear map

$$f : V \to W$$

between two vector spaces, there exists a canonical linear map

$$f^* : W^* \to V^*$$

between the duals. Note that the direction of the arrow inverts – mathematicians say that taking the dual of a vector space is a *contravariant functor* to indicate this reversal, but don't worry about this if you haven't heard of the term functor before.

B.2.2 Dual spaces and bilinear forms

In linear algebra we often use bilinear forms

$$\langle ., . \rangle : V \times V \to K : (\mathbf{v}, \mathbf{w}) \mapsto \langle \mathbf{v}, \mathbf{w} \rangle,$$

e.g. if $K = \mathbb{R}$ then an Euclidean structure is a particular example of such a bilinear form (namely, it is a bilinear form that is also *symmetric*, *non-degenerate*, and *positive-definite*).

With every bilinear form $\langle ., . \rangle$ we can canonically associate a linear map $f_{\langle .,. \rangle}$ from V to V^*, as follows

$$f_{\langle .,. \rangle} : V \to V^* : \left(f_{\langle .,. \rangle}(\mathbf{v})\right)(\mathbf{w}) = \langle \mathbf{v}, \mathbf{w} \rangle$$

for all $\mathbf{v}, \mathbf{w} \in V$.

Exercise B.4. Show that $\langle ., . \rangle$ is *non-degenerate* if and only if $f_{\langle .,. \rangle}$ is an isomorphism.

Exercise B.5. Show that $\langle ., . \rangle$ is *symmetric* (i.e. $\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle$ for all $\mathbf{v}, \mathbf{w}$) if and only if

$$f_{\langle .,. \rangle} = f^*_{\langle .,. \rangle}$$

and likewise $\langle ., . \rangle$ is *anti-symmetric* (i.e. $\langle \mathbf{v}, \mathbf{w} \rangle = -\langle \mathbf{w}, \mathbf{v} \rangle$ for all $\mathbf{v}, \mathbf{w}$) if and only if

$$f_{\langle .,. \rangle} = -f^*_{\langle .,. \rangle}.$$

Vice versa, with every linear map $V \to V^*$ we can associate a bilinear form – these two things are just different ways of encoding the same information.

Remark B.6. At this point you may be very confused about the need to think of something like the dual vector space. Why do we bother with it at all? Why don't we simply identify every n-dimensional vector space with K^n?

If we do this (i.e. if we choose a basis), we also get (through the dual basis) an identification of V^* with K^n, and hence an isomorphism $V \cong V^*$. However, this isomorphism really depends on this choice – if we choose a different basis for V, we get a different isomorphism $V \cong V^*$. So we say that these isomorphisms are *not canonical*.

Another way to obtain an isomorphism $V \cong V^*$ is by choosing a non-degenerate bilinear form, but again: if we change the bilinear form, the isomorphism changes.

On the other hand, given any vector space V (without the need of Euclidean structure or anything like this), the dual space V^* always canonically exists.

And moreover, the isomorphism $V \cong (V^*)^*$ for a finite dimensional V is always naturally defined, no choice is needed.[6]

B.3 Direct sums and products

Another way to create new vector spaces is by taking the *direct sum* of two vector spaces V and W, denoted by $V \oplus W$. As a set, this is just the Carthesian product:

$$V \oplus W = V \times W = \{(\mathbf{v}, \mathbf{w}) \mid \mathbf{v} \in V, \mathbf{w} \in W\}.$$

We want to make this into a vector space, so we need the vector space operations – addition of vectors and scalar multiplication. This is done by defining

$$\lambda(\mathbf{v}, \mathbf{w}) = (\lambda \mathbf{v}, \lambda \mathbf{w}) \tag{13}$$

and

$$(\mathbf{v}_1, \mathbf{w}_1) + (\mathbf{v}_2, \mathbf{w}_2) = (\mathbf{v}_1 + \mathbf{v}_2, \mathbf{w}_1 + \mathbf{w}_2).$$

Exercise B.7. Check that this makes $V \oplus W$ into a vector space.

Often we will think of V and W as subspaces of $V \oplus W$, by identifying $\mathbf{v} \in V$ with $(\mathbf{v}, 0)$ and $\mathbf{w} \in W$ with $(0, \mathbf{w})$. So even though $\mathbf{v}$ and $\mathbf{w}$ a priori live in different vector spaces, we can abuse notation and write $\mathbf{v} + \mathbf{w}$, where it is understood that this refers to the element $(\mathbf{v}, \mathbf{w}) \in V \oplus W$.

Exercise B.8. Show that, if V and W are both finite-dimensional vector spaces, then

$$\dim(V \oplus W) = \dim(V) + \dim(W). \tag{14}$$

You can of course take the direct sum of more than two vector spaces as well, and because there are canonical isomorphisms

$$(U \oplus V) \oplus W \cong U \oplus (V \oplus W),$$

as one easily sees, there is no need to worry about the order in which we do this. Note that you do have to be a little careful if you want to take the direct sum of infinitely many vector spaces: as a set this is no longer the Carthesian product, but rather a subset:

$$\bigoplus_{i \in I} V_i = \left\{(\mathbf{v}_i)_i \;\middle|\; \mathbf{v}_i \in V_i \text{ and only finitely many } \mathbf{v}_i \neq \mathbf{0}\right\}.$$

If one drops the finiteness condition, one still gets a vectorspace, but this is referred to as the *direct product*. So for a finite number of vector spaces the direct sum and direct product are the same, but for infinitely many the former is contained in the latter.

[6] At this point you may be tempted to think that any procedure that needs a basis will not be canonical. But you have to be careful! Sometimes it is just convenient to define something using a basis, but it turns out that changing the basis doesn't make any difference. E.g. if we have a linear map $f : V \to V$, we get a square matrix representing this as soon as we choose a basis for V, and we can now take the determinant of this square matrix. But if we choose a different basis, and hence get a different matrix, the determinant doesn't change! So we associate the determinant to f itself, not just to the matrix representing it. Later we will see how the determinant of such a linear map f can in fact be defined without invoking any basis at all.

B.4 The tensor product

Like the direct sum, the tensor product is a way of creating a new vector space out of two old ones, but it is a bit harder to get our hands on. Often students find it difficult to understand at the beginning at the beginning, because unlike the direct sum, the tensor product is not quite as easy describe as a set. Nevertheless, once one gets used to them they are extremely useful, and one can recognise a lot of other constructions as special cases of the tensor product.

So unlike the direct sum, we are not going to begin by defining the tensor product as a set. Rather, we are going to begin by listing some of properties of this hypothetical tensor product, and only after that shall we give a construction that realises these properties.

It all starts from two vector spaces, V and W, over the same base field K (e.g. both are real vector spaces). The (hypothetical) new vector space that we are interested in, the tensor product of V and W, will be denoted as

$$V \otimes W.$$

It will be a new vector space, defined over the same field K. Sometimes, if we want to clarify which base field we are working over, $V \otimes W$ will be denoted by $V \otimes_K W$.

B.4.1 Similarities and differences between $V \oplus W$ and $V \otimes W$

The first fact of $V \otimes W$ that we want to discuss is something pretty similar to direct products: it says that we can create elements of $V \otimes W$ out of elements in V and W:

$$\text{If } \mathbf{v} \in V \text{ and } \mathbf{w} \in W \ \Rightarrow \mathbf{v} \otimes \mathbf{w} \in V \otimes W.$$

But these new elements $\mathbf{v} \otimes \mathbf{w}$ behave a little different from their analogues in the direct sum:

$$\lambda(\mathbf{v} \otimes \mathbf{w}) = (\lambda \mathbf{v}) \otimes \mathbf{w} = \mathbf{v} \otimes (\lambda \mathbf{w}).$$

Compare this with (13): it may seem to be just a little bit different, but this small difference will ensure that direct sums and tensor products are really quite different beasts.

Another difference is that we will have the following rule (a kind of distributivity):

$$\mathbf{v}_1, \mathbf{v}_2 \in V, \mathbf{w} \in W \ \Rightarrow (\mathbf{v}_1 + \mathbf{v}_2) \otimes \mathbf{w} = (\mathbf{v}_1 \otimes w) + (\mathbf{v}_2 \otimes \mathbf{w}),$$

and similarly

$$\mathbf{v} \in V, \mathbf{w}_1, \mathbf{w}_2 \in W \ \Rightarrow \mathbf{v} \otimes (\mathbf{w}_1 + \mathbf{w}_2) = (\mathbf{v} \otimes w_1) + (\mathbf{v} \otimes \mathbf{w}_2).$$

It will be very important to note that these properties can be interpreted as saying that there is a morphism

$$V \times W \to V \otimes W : (\mathbf{v}, \mathbf{w}) \mapsto \mathbf{v} \otimes \mathbf{w}$$

that is *bilinear*, i.e. linear in each entry. Another way of phrasing bilinearity is to say that given any $\mathbf{w} \in W$ the map

$$. \otimes \mathbf{w} : V \to V \otimes W : \mathbf{v} \mapsto \mathbf{v} \otimes \mathbf{w}$$

is linear, and similarly for the corresponding map $\mathbf{v} \otimes . : W \to V \otimes W$ given any $\mathbf{v} \in V$. This morphism will turn out to be crucial in thinking about the tensor product.

B.4.2 Pure tensors and bases for tensor product spaces

Remark that though we can create elements in both $V \oplus W$ and $V \otimes W$ out of an element $\mathbf{v}$ in V and an element $\mathbf{w}$ in W, there is a crucial difference: *every* element in $V \oplus W$ can be written in this way as $\mathbf{v} \oplus \mathbf{w}$ for some $\mathbf{v} \in V$ and $\mathbf{w} \in W$. But not every element in $V \otimes W$ can be written as $\mathbf{v} \otimes \mathbf{w}$. What is true however is that every element of $V \otimes W$ can be written as a finite sum

$$\sum_i \mathbf{v}_i \otimes \mathbf{w}_i.$$

Indeed, we have the following:

Proposition B.9. *If we have a basis $(\mathbf{v}_i)_i$ for V and a basis $(\mathbf{w}_j)_j$ for W, then*

$$(\mathbf{v}_i \otimes \mathbf{w}_j)_{i,j}$$

forms a basis for $V \otimes W$.

Tensors (that is, elements of the tensor product $V \otimes W$) of the form $\mathbf{v} \otimes \mathbf{w}$ are called *pure* tensors, and, if we want to define a linear map from a tensor product $V \otimes W$ to some other vector space, it is sufficient, because of Proposition B.9, to define it on the pure tensors. We will use this fact throughout.

As a corollary of Proposition B.9 we see that, if V and W are both finite-dimensional, then so is $V \otimes W$, and we have

$$\dim(V \otimes W) = \dim(V)\dim(W), \tag{15}$$

which contrasts with (14).

B.4.3 Tensoring with K

Note that we can think of K as a one-dimensional vector space over itself. Tensoring by K, by (15), shouldn't change the dimension, and indeed we have canonical isomorphisms

$$K \otimes_K V \cong V \cong V \otimes_K K,$$

given by identifying $\lambda \otimes \mathbf{v}$ and $\lambda\mathbf{v}$ etc.

B.4.4 Linear maps as tensors

To show that this recovers some known spaces, think about the following: we can write down a map from $W \otimes V^*$ to $\mathrm{Hom}(V, W)$, the space of all linear maps from V to W, as follows: with every element of the form $\mathbf{w} \otimes \alpha$ in $W \otimes V^*$ we can associate a linear map $\Phi_{\mathbf{w}\otimes\alpha}$ from V to W, as follows:

$$\Phi_{\mathbf{w}\otimes\alpha}(\mathbf{v}) = (\alpha(\mathbf{v}))\mathbf{w}.$$

As we said, not every element in $W \otimes V^*$ is of the form $\mathbf{w} \otimes \alpha$, but it will be finite sum of such elements, and we can take the sum of the corresponding linear maps from V to W. So we end up with a linear map from $W \otimes V^*$ to $\mathrm{Hom}(V, W)$. This map is an isomorphism:

$$\mathrm{Hom}(V, W) \cong W \otimes V^*.$$

To make this a bit more concrete: suppose that $(\mathbf{v}_i)_i$ is a basis for V, with dual basis α^i, and $(\mathbf{w}_j)_j$ is a basis for W. Then for any pair of indices (i, j) the linear map $\Phi_{\mathbf{w}_j\otimes\alpha_i}$ corresponding to $\mathbf{w}_j \otimes \alpha_i \in W \otimes V^*$ has a matrix representation (with respect to the bases $(\mathbf{v}_i)_i$ and $(\mathbf{w}_j)_j$) all of whose entries are 0, except for the one in the (i, j)-th position, which is 1.

B.4.5 Universal property of tensor products

Now so far we have just stated some properties of this hypothetical new vector space $V \otimes W$, but we haven't actually constructed it yet. We will do that shortly, but before doing so that we will divert into abstraction a bit. It turns out that the tensor product satisfies something called a *universal property.*

Theorem B.10 (Uniqueness of tensor products)**.** *Suppose that, given two vector spaces V and W, we have a new vector space Z and a bilinear map*

$$\Phi : V \times W \to Z$$

that satisfies the following (known as the universal property of tensor products*):*

> *For every other bilinear map and K-vector space U*
>
> $$\Psi : V \times W \to U$$
>
> *there exists a unique linear map $\Xi : Z \to U$ that makes the following diagram commute:*
>
> $$\begin{tikzcd} V \times W \arrow[r, "\Phi"] \arrow[dr, "\Psi"'] & Z \arrow[d, "\Xi"] \\ & U \end{tikzcd} \tag{16}$$

Then Z is unique up to unique isomorphism, i.e. for any other such $\widetilde{Z}$ there exist a canonical isomorphism $Z \to \widetilde{Z}$.

The statement that this diagram (16) commutes just means that $\Psi = \Xi \circ \Phi$.

Proof. Suppose that we have another such vector space $\widetilde{Z}$ and linear map

$$\widetilde{\Phi} : V \times W \to \widetilde{Z}.$$

Then we can apply the universal property in both directions, so we get unique maps Ξ_1 and Ξ_2 such that the following diagrams commute:

$$\begin{tikzcd} V \times W \arrow[r, "\Phi"] \arrow[dr, "\widetilde{\Phi}"'] & Z \arrow[d, "\Xi_1"] \\ & \widetilde{Z} \end{tikzcd}$$

and

$$\begin{tikzcd} V \times W \arrow[r, "\widetilde{\Phi}"] \arrow[dr, "\Phi"'] & \widetilde{Z} \arrow[d, "\Xi_2"] \\ & Z. \end{tikzcd}$$

So now we can compose them, to get maps

$$\Xi_2 \circ \Xi_1 : Z \to Z,$$

and
$$\Xi_1 \circ \Xi_2 : \widetilde{Z} \to \widetilde{Z}.$$
However, since by construction both of those maps make the diagrams

$$\begin{tikzcd} V \times W \arrow[r, "\Phi"] \arrow[dr, "\Phi"'] & Z \arrow[d, "\Xi_1 \circ \Xi_2"] \\ & \widetilde{Z} \end{tikzcd}$$

and

$$\begin{tikzcd} V \times W \arrow[r, "\Phi"] \arrow[dr, "\widetilde{\Phi}"'] & \widetilde{Z} \arrow[d, "\Xi_2 \circ \Xi_1"] \\ & \widetilde{Z} \end{tikzcd}$$

commute, and the same is true if we replace them by the identity, from the uniqueness part of the universal property, they have to be the identity. So Ξ_1 and Ξ_2 are each other inverses, and hence they are both isomorphisms. Hence Z is unique up to unique isomorphism. □

Now at first sight this might be a very confusing statement, and a hard-nosed mathematician could be forgiven for suspecting that the author has spent too much time drinking wine on the left bank in Paris reading post-modern nonsense.

Nevertheless, this property is quite useful to guide our thinking about tensor products. Indeed, it tells us that if such a Z exist (and the tensor product is indeed of this nature), then it is essentially defined by being minimal whilst having the desired properties. A proper, high-brow, way to say this mathematically is that a tensor product $V \otimes W$ is an *initial object* in the category of bilinear maps $\Psi : V \times W \to Z$.

This also means that we should not give too much importance to an actual construction of the tensor product. Such a construction only serves to prove that the tensor product exists, but once we know it exists we can forget the construction, and focus entirely on the defining properties. One can give several different constructions of the tensor product, but by Theorem B.10 they are all automatically canonically isomorphic.

Remark B.11. Note that one can similarly give a universal property for many other constructions in mathematics – for example the direct sum of two vector spaces. When a construction is easy enough to do, and clear enough to understand the main properties directly from it, one often does not bother too much with the universal property. For tensor products the construction is in general messy enough that the universal property is actually helpful in clarifying what the tensor product is.

Exercise B.12. Show, using the universal property, that if a tensor product $V \otimes W$ exists, there automatically is a canonical isomorphism
$$V \otimes W \to W \otimes V.$$

Exercise B.13. Show that, given linear maps $f : V_1 \to V_2$ and $g : W_1 \to W_2$, there exists a canonical map
$$f \otimes g : V_1 \otimes W_1 \to V_2 \otimes W_2.$$

B.4.6 Construction

There are a number of ways we can construct a tensor product (that is, a vector space Z and a bilinear map $\Phi : V \times W \to Z$ that satisfies the universal property). For more general notions of tensor products, for modules over rings (see Section B.8), these can be quite complicated. For vector spaces however, the situation is not too bad.

Indeed, we can simply take $V \otimes W$ to be

$$V \otimes W = \left\{ \phi : V^* \times W^* \to K \;\middle|\; \phi \text{ is bi-linear} \right\}. \tag{17}$$

This set becomes a K-vector space (check this!) by defining the sum $\phi_1 + \phi_2$ of two elements as

$$(\phi_1 + \phi_2)(\mu, \nu) = \phi_1(\mu, \nu) + \phi_2(\mu, \nu)$$

and similarly the scalar product $\lambda\alpha$ as

$$(\lambda\alpha)(\mu, \nu = \lambda(\alpha(\mu, \nu)).$$

Of course, it is not enough to describe what $V \otimes W$ is as a set, we also need to give a bilinear map

$$\Phi : V \times W \to V \otimes W : (\mathbf{v}, \mathbf{w}) \mapsto \mathbf{v} \otimes \mathbf{w}.$$

So with every $\mathbf{v} \in V$ and $\mathbf{w} \in W$ we need to associate a bilinear map

$$\Phi(\mathbf{v}, \mathbf{w}) = \mathbf{v} \otimes \mathbf{w} : V^* \times W^* \to K.$$

It is obvious though which one we should take:

$$(\mathbf{v} \otimes \mathbf{w})(\mu, \nu) = (\mu(\mathbf{v}))(\nu(\mathbf{w})). \tag{18}$$

We can now show

Theorem B.14 (Existence of tensor products)**.** *The space $V \otimes W$ and the linear map Φ defined in (17) and (18) satisfy the Universal Property of tensor products.*

Proof. Suppose we have a vector space U, and a bilinear map $\Psi : V \times W \to U$. Then we want to find a map $\Phi : V \otimes W \to U$, such that the diagram

$$\begin{tikzcd} V \times W \arrow[r, "\Phi"] \arrow[dr, "\Psi"'] & V \otimes W \arrow[d, "\Xi"] \\ & U \end{tikzcd} \tag{19}$$

commutes, and moreover this map has to be unique.

Now we can begin by remarking that $V \otimes W$ is indeed (as was claimed in Proposition B.9) generated as a vectorspace by elements of the form $\mathbf{v} \otimes \mathbf{w}$. Indeed, if we just pick a basis $(\mathbf{v}_i)_i$ of V and $(\mathbf{w}_j)_J$ of W, with dual bases $(\alpha^k)_k$ and $(\beta^l)_l$ for V^* and W^* respectively, then we can express any $\mu \in V^*$ as (using the Einstein summation convention)

$$\mu = \mu(\mathbf{v}_i)\alpha^i$$

and any $\nu \in W^*$ as

$$\nu = \nu(\mathbf{w}_j)\beta^j.$$

Therefore, for any element $\phi \in V \otimes W$, we have by bilinearity

$$\phi(\mu, \nu) = \mu(\mathbf{v}_i)\nu(\mathbf{w}_j)\phi(\alpha^i, \beta^j).$$

But by (18) this is just the same thing as saying

$$\phi = \phi(\alpha^i, \beta^j)\mathbf{v}_i \otimes \mathbf{w}_j,$$

and hence indeed $V \otimes W$ is generated by elements of the form $\mathbf{v} \otimes \mathbf{w}$. So in order to determine what the linear map

$$\Xi : V \otimes W \to U$$

should be, it suffices to determine what $\Xi(\mathbf{v} \otimes \mathbf{w})$ should be, for any choice of $\mathbf{v}$ and $\mathbf{w}$. But it is now clear that one (and only one) choice will make the diagram (19) commute, namely

$$\Xi(\mathbf{v} \otimes \mathbf{w}) = \Psi(\mathbf{v}, \mathbf{w}).$$

It is finally a straightforward verification that this map Ξ is indeed a linear map from $V \otimes W$ to U. $\square$

This construction may not feel all that enlightening. But remember: the main purpose the construction serves is just to prove that tensor products indeed exist. That is all you need to remember!

B.5 Exterior and Symmetric products

B.5.1 Exterior products (a.k.a. wedge products)

There are two closely related constructions that we also should introduce: the *exterior product* and *symmetric product*. The ideas behind the exterior product will already be familiar from the discussion of p-forms: given a vector space V over a field K of dimension n, there exist $n+1$ K-vectorspace (of which the first two are just K and V itself):

$$\bigwedge^0 V = K, \ \bigwedge^1 V = V, \bigwedge^2 V, \ldots, \bigwedge^n V.$$

For each k-tuple $\mathbf{v}_1, \ldots, \mathbf{v}_k$, there is an element

$$\mathbf{v}_1 \wedge \ldots \wedge \mathbf{v}_k \in \bigwedge^k V.$$

This assignment

$$(\mathbf{v}_1, \ldots, \mathbf{v}_k) \mapsto \mathbf{v}_1 \wedge \ldots \wedge \mathbf{v}_k$$

is multi-linear (i.e. linear in each entry), and moreover it is anti-symmetric: if $\sigma \in \Sigma_k$ is a permutation, then

$$\mathbf{v}_{\sigma(1)} \wedge \ldots \wedge \mathbf{v}_{\sigma(k)} = (-1)^{\mathrm{sign}(\sigma)}\mathbf{v}_1 \wedge \ldots \wedge \mathbf{v}_k.$$

This anti-symmetry immediately shows why we can only go up to $\bigwedge^n V$: if we take k to be any larger than the dimension of V, $\bigwedge^k V$ has to be $\{\mathbf{0}\}$.

Note that because the map $V^k \to \bigwedge^k V$ is multilinear, by the universal property for the tensor product, it has to factor uniquely through the k-fold tensor power

$$V^{\otimes k} = \underbrace{V \otimes \cdots \otimes V}_{k \text{ times}}.$$

So we get a canonical, unique map

$$P_\wedge : V^{\otimes k} \to \bigwedge^k V. \tag{20}$$

In general, $\bigwedge^k V$ can be constructed as a quotient space (factor module) of $V^{\otimes k}$ (the notion of factor modules is discussed in Honours Algebra). But, if the characteristic of the field K is zero (e.g. if K is $\mathbb{Q}$ or $\mathbb{R}$ or $\mathbb{C}$), then we can also consider it as a *subspace* of $V^{\otimes k}$. Indeed, remember from exercise B.12 that for any two K-vector spaces V and W we get a canonical isomorphism $V \otimes W \to W \otimes V$. Now, we can also use this in the case when $V = W$. This isomorphism $V \otimes V \to V \otimes V$ is not the identity, indeed it will always send $\mathbf{v}_1 \otimes \mathbf{v}_2 \mapsto \mathbf{v}_2 \otimes \mathbf{v}_1$. Now, one can easily see that, for any permutation $\sigma \in \Sigma_k$, we similarly get an isomorphism

$$\Upsilon_\sigma : V^{\otimes k} \to V^{\otimes k} : \mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_k \mapsto \mathbf{v}_{\sigma(1)} \otimes \cdots \otimes \mathbf{v}_{\sigma(k)}.$$

With this in hand, we can define $\bigwedge^k V$ to be the subspace of $V \otimes k$ consisting of those elements $\mathbf{t} \in V^{\otimes k}$ such that

$$\Upsilon_\sigma(\mathbf{t}) = (-1)^{\mathrm{sign}(\sigma)}\mathbf{t}.$$

Moreover, the map $P_\wedge$ from (20) can also be made explicit:

$$P_\wedge : V^{\otimes k} \to \bigwedge^k V : \mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_k \mapsto \frac{1}{k!} \sum_{\sigma \in \Sigma_k} \mathbf{v}_{\sigma(1)} \otimes \cdots \otimes \mathbf{v}_{\sigma(k)}.$$

Because of the factor $\frac{1}{k!}$ that is present this doesn't work anymore in positive characteristic.

Exercise B.15. Verify that V is well-defined.

Exercise B.16. Given a basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ for V, write down a basis for $\bigwedge^k V$. Show that

$$\dim\left(\bigwedge^k V\right) = \binom{\dim V}{k}.$$

Exercise B.17. Formulate a universal property satisfied by the exterior product.

Exercise B.18. Show that the same universal property is satisfied by the space of all *alternating* multilinear maps

$$\underbrace{V^* \times \cdots \times V^*}_{k \text{ times}} \to K.$$

This links the approach taken in the appendix with Definition 4.1.

Exercise B.19. Given a linear map $f : V \to W$, show that there exists a canonical induced map

$$\bigwedge^k f : \bigwedge^k V \to \bigwedge^k W.$$

In particular, if $V = W$, this gives a map

$$\bigwedge^{\dim(V)} f : \bigwedge^{\dim(V)} V \to \bigwedge^{\dim(V)} V.$$

Now, since $\dim\left(\bigwedge^{\dim(V)} V\right) = 1$, and a linear map from a 1-dimensional vector space to itself is always given by multiplication by a scalar constant, this associates a scalar to f, independent of any choice of basis. Show that this scalar is the determinant, $\det(f)$:

$$\left(\bigwedge^{\dim(v)} f\right)(\boldsymbol{x}) = \det(f)\boldsymbol{x}.$$

Note: this can in fact be used as an entirely intrinsic way of defining what the determinant is, without ever invoking a basis.

B.6 Symmetric products

The story of exterior products also has a symmetric brother, the symmetric product. We have symmetric powers

$$\operatorname{Sym}^0 V = K, \operatorname{Sym}^1 V = V, \ldots, \operatorname{Sym}^r V, \ldots$$

Unlike the exterior products, this sequence doesn't end.

In fact, it turns out that the symmetric product is something already very familiar, as $\operatorname{Sym}^d V$ can best be understood as the space of homogeneous degree d polynomials on the space V^*.

To stay consistent with the product notation for polynomials, we do not use any symbol to denote the symmetric product. That is to say, we just write

$$\text{if } \mathbf{v}, \mathbf{w} \in V \Rightarrow \mathbf{vw} = \mathbf{wv} \in \operatorname{Sym}^2 V.$$

The key point here is that this assignment

$$(\mathbf{v}_1, \ldots, \mathbf{v}_k) \in V^k \mapsto \mathbf{v}_1 \ldots \mathbf{v}_k \in \operatorname{Sym}^k V$$

is symmetric:

$$\mathbf{v}_1 \ldots \mathbf{v}_k = \mathbf{v}_{\sigma(1)} \ldots \mathbf{v}_{\sigma(k)}$$

for each permutation σ.

Exercise B.20. Write down a universal property satisfied by the symmetric product.

Exercise B.21. Write down $\operatorname{Sym}^k V$ as a subspace of $V^{\otimes k}$, as well as a projection operator $P_{\text{Sym}} : V^{\otimes k} \to \operatorname{Sym}^k V$.

Exercise B.22. Prove that

$$\dim\left(\operatorname{Sym}^k V\right) = \binom{\dim V + d}{\dim V}.$$

B.7 Algebras from vector spaces

Given a vector space V, it is often useful to put all tensor powers together, to get an space

$$T(V) = \bigoplus_{k\in\mathbb{N}} V^{\otimes k}.$$

Note that this sum starts from 0, as we used the convention $0 \in \mathbb{N}$.

This space $T(V)$ is an infinite dimensional vector space (unless if $V = \{\mathbf{0}\}$, in which case also $T(V) = \{\mathbf{0}\}$), but more importantly the tensor product operation

$$\otimes : V^{\otimes k_1} \times V^{\otimes k_2} \to V^{\otimes(k_1+k_2)}$$

makes this into an $\mathbb{K}$-algebra (that is to say, a K-vector space that is also a ring – these notions are explained in the course *Honours Algebra*. We refer to it as the tensor algebra.

The same is true for the exterior and symmetric product: we can create the exterior algebra

$$\bigwedge^{\bullet} V = \bigoplus_{k\in\mathbb{N}} \bigwedge^{k} V$$

and the symmetric algebra

$$\operatorname{Sym}^{\bullet} V = \bigoplus_{k\in\mathbb{N}} \operatorname{Sym}^{k} V.$$

Note that $\bigwedge^{\bullet} V$ is always a finite-dimensional vector space if V is finite-dimensional, unlike $T(V)$ and $\operatorname{Sym}^{\bullet}$ V.

Remark B.23. If the characteristic of K is 0, we can think of each $\bigwedge^k(V)$ and $\operatorname{Sym}^k(V)$ as a subspace of $V^{\otimes k}$. This makes $\bigwedge^{\bullet} V$ and $\operatorname{Sym}^{\bullet} V$ into subspaces of $T(V)$, but these subspaces are not subalgebras! Indeed, the subspaces $\bigwedge^{\bullet} V$ and $\operatorname{Sym}^{\bullet} V$ are not preserved under the multiplication $\otimes$ on $T(V)$. Rather, what we do have (if we use the two projection operators $P_{\bigwedge}$ and P_{Sym}) is

$$\text{if } \alpha, \beta \in \bigwedge^{\bullet} V \subset T(V) \Rightarrow \alpha \wedge \beta = P_{\bigwedge}(\alpha \otimes \beta)$$

and similarly

$$\text{if } \alpha, \beta \in \operatorname{Sym}^{\bullet} V \subset T(V) \Rightarrow \alpha\beta = P_{\mathrm{Sym}}(\alpha \otimes \beta).$$

B.8 Generalizations

The operations outlined above one can apply to create new vector spaces out of old ones can be generalized to the much broader setting of *modules M over commutative rings R* (these notions are discussed in the course *Honours Algebra*). So one can take tensor products of modules, take wedge products, etc.

There are some caveats however. Firstly, though one can still take the dual module

$$M^* = \operatorname{Hom}(M, R),$$

this notion is not quite as useful as it was for vector spaces. One reason for this is that for vector spaces one essentially doesn't loose information. Indeed, for finite-dimensional vector spaces, since one canonically has $V \cong (V^*)^*$ one can in fact recover the original vector space from its dual. This is not always the case for modules: e.g. think of the abelian group $\mathbb{Z}/p\mathbb{Z}$ as a $\mathbb{Z}$-module. Then one has that

$$\mathrm{Hom}(\mathbb{Z}/p\mathbb{Z}, \mathbb{Z}) = \{0\}.$$

This is due to *torsion* – something which is not present for vector spaces (note that this refers to the algebraic notion of torsion, which is unrelated to the geometric notion of torsion for curves discussed in the course).

Similarly funny things can happen for tensor products. Because of formula (15), one could naively think that taking the tensor product of a vector space with any non-trivial vector space creates a *bigger* space in some sense. But for general modules, things like the following can happen: think of both $\mathbb{Q}$ and $\mathbb{Z}/p\mathbb{Z}$ as $\mathbb{Z}$-modules. Then we have

$$\mathbb{Q} \otimes_{\mathbb{Z}} \mathbb{Z}/p\mathbb{Z} = \{0\}.$$

Indeed, for every pure tensor $\lambda \otimes \mu \in \mathbb{Q} \otimes_{\mathbb{Z}} \mathbb{Z}/p\mathbb{Z}$ we have

$$\lambda\otimes\mu = \left(\left(p\frac{1}{p}\right)\lambda\right)\otimes\mu = \left(p\left(\frac{1}{p}\lambda\right)\right)\otimes\mu = p\left(\left(\frac{1}{p}\lambda\right)\otimes\mu\right) = \left(\frac{1}{p}\lambda\right)\otimes(p\mu) = \left(\frac{1}{p}\lambda\right)\otimes 0 = 0.$$

But apart from things like these, the constructions behave largely the same as for vector spaces (though the construction of the tensor product is necessarily a bit harder).

B.9 Use in differential geometry

In this course, the main application we make of these constructions is in the notion of p-form. We can now think about this in different ways. The first, most direct, way is to consider the co-tangent space T_p^*D at any point $p \in D$, then take the wedge-powers of this space $\bigwedge^k T_p^*D$, and then define a p-form to be the assignment of an element in $\bigwedge^k T_p^*D$ for each $p \in D$.

But we can also argue slightly differently. We could begin by considering $C^\infty(D)$, the set of smooth real-valued functions on D. Pointwise addition and multiplication makes this into a commutative ring (in fact a real algebra, as it is also a real vector space). Moreover, if we look at the set of all (smooth vector) fields, then we see that this set is actually a module over $C^\infty(D)$. Indeed, we can add vector fields (simply add them point wise), and we can multiply a vector field by a scalar to get a new vector field.

Also the set of 1-forms is a module over $C^\infty(D)$, in the same way. So a second way to think about $\omega^p(D)$ is to say that it is the p-fold wedge product of $\Omega^1(D)$ as a $C^\infty(D)$-module:

$$\Omega^p(D) = \bigwedge^p \Omega^1(D).$$

Both of these viewpoints give the same end result.

B.10 Excercises

Exercise B.24. One way to put tensor products to use is in the construction of *extension of scalars*. Suppose we have a K-vector space V, and a field extension $K \subset \widetilde{K}$. E.g. $K = \mathbb{R}$, V is a real vector space, and $\widetilde{K} = \mathbb{C}$. There are situations where we really would like to be able to multiply elements in V by elements in $\widetilde{K}$, but a priori there is no way to do this. However, consider the new vector space

$$V_{\widetilde{K}} := V \otimes_K \widetilde{K}.$$

At first sight this is just a K-vectorspace, but it actually canonically becomes a $\widetilde{K}$-vectorspace, by defining the scalar multiplication

$$\widetilde{\lambda}\left(\sum_i \mathbf{v}_i \otimes \gamma_i\right) = \sum_i \mathbf{v}_i \otimes \widetilde{\lambda}\gamma_i.$$

Moreover, we have an injection of K-vectorspaces:

$$V \hookrightarrow V_{\widetilde{K}} : \mathbf{v} \mapsto \mathbf{v} \otimes 1.$$

Note that this is true because the new scalar multiplication in $V_{\widetilde{K}}$ extends the old one! Show that any (K-)basis for V automatically becomes a $\widetilde{K}$-basis for $V_{\widetilde{K}}$. Deduce that

$$\dim_K V = \dim_{\widetilde{K}} V_{\widetilde{K}}.$$

Exercise B.25. If V is a complex vector space, we can restricts scalar multiplication to real scalars. In this way we end up with a real vector space, whose underlying additive group is the same as V. To clarify the distinction we will denote this as $V_{\mathbb{R}}$.

We can also create a new complex vector space, with the same structure of additive group, but where the new scalar multiplication is given by $\lambda \tilde{\cdot} \mathbf{v} = \overline{\lambda}\mathbf{v}$. We will denote this vector space as $\overline{V}$. Remark that $V_{\mathbb{R}} = \overline{V}_{\mathbb{R}}$.

We can now do a funny thing: we can first forget the complex structure of V, to only remember the real vector space $V_{\mathbb{R}}$, and then take the tensor product over $\mathbb{R}$ with $\mathbb{C}$ to end up with a new complex vector space, $(V_{\mathbb{R}}) \otimes_{\mathbb{R}} \mathbb{C}$.

Show now that there exists a canonical isomorphism of complex vector spaces

$$(V_{\mathbb{R}}) \otimes_{\mathbb{R}} \mathbb{C} \cong V \oplus \overline{V}. \tag{21}$$

Remark B.26. It is easy to get confused in this setting. Remark that every vector $\mathbf{v} \in (V_{\mathbb{R}}) \otimes_{\mathbb{R}} \mathbb{C}$ can be written as $\mathbf{v} = \mathbf{v}_1 \otimes 1 + \mathbf{v}_2 \otimes i$ for some $\mathbf{v}_1, \mathbf{v}_2 \in V$. This gives another decomposition

$$((V_{\mathbb{R}}) \otimes_{\mathbb{R}} \mathbb{C})_{\mathbb{R}} = V_{\mathbb{R}} \oplus iV_{\mathbb{R}}. \tag{22}$$

This is a different decomposition from (21)! The isomorphism (22) is only an isomorphism of real vector spaces, the decomposition (21) is an isomorphism of complex vector spaces.

Another way to think about the decomposition (21) is as follows: if we denote the multiplication by $\sqrt{-1}$ (coming from the 'original' complex structure on V) by I, then we can think of I as a real-linear map

$$I : V_{\mathbb{R}} \to V_{\mathbb{R}}$$

that satisfies $I^2 = -1$. Now any (real-)linear map between real vector spaces

$$f : W_1 \to W_2$$

automatically gives rise to a complex linear map between their complexifications

$$f_{\mathbb{C}} : (W_1)_{\mathbb{C}} \to (W_2)_{\mathbb{C}}.$$

Therefore, we can consider the map

$$I_{\mathbb{C}} : (V_{\mathbb{R}}) \otimes_{\mathbb{R}} \mathbb{C} \to (V_{\mathbb{R}}) \otimes_{\mathbb{R}} \mathbb{C}.$$

Now, this map still satisfies $(I_{\mathbb{C}})^2 = -1$, and so therefore if we think about its eigenvalues, these can only be i and $-i$ (in the 'new' complex structure on $(V_{\mathbb{R}}) \otimes_{\mathbb{R}} \mathbb{C}$). In fact, it turns out that the map $I_{\mathbb{C}}$ is diagonalizable, and both the i and the $-i$ eigenspaces are complex subspaces of $(V_{\mathbb{R}}) \otimes_{\mathbb{R}} \mathbb{C}$ of half the latter's dimension. The i-eigenspace is isomorphic to V (equipped with the 'old' complex structure), and the $-i$-eigenspace is isomorphic to $\overline{V}$.

Section C: Differential forms vs. 3-dimensional vector calculus

C.1 Cross-products and orientations

C.1.1 When does the cross-product exist?

In basic linear algebra the *cross-product* is introduced: it takes two vectors $\mathbf{v}$ and $\mathbf{w}$ in a three-dimensional real vector space V, and returns a new vector $\mathbf{v} \times \mathbf{w}$ in the same vector space.

This operation, crucially, only happens in three dimensions. Actually, we need a bit more: the vector space V needs to be Euclidean. Indeed, we want $\mathbf{v} \times \mathbf{w}$ to be orthogonal to both $\mathbf{v}$ and $\mathbf{w}$, and to have norm $|\mathbf{v}|\mathbf{w}|\,|\sin(\theta)|$, where θ is the angle between $\mathbf{v}$ and $\mathbf{w}$.

Upon close inspection, it is not even enough for V to be three-dimensional and Euclidean. We need an extra piece of structure on top of V: the choice of an orientation. Indeed, both $\mathbf{v} \times \mathbf{w}$ and $-\mathbf{v} \times \mathbf{w}$ have the property that they are orthogonal to $\mathbf{v}$ and $\mathbf{w}$, and they have the same norm. A choice of an orientation of V picks out $\mathbf{v} \times \mathbf{w}$.

C.1.2 Orientations of real vector spaces

There are a number of different ways to define what an orientation is, but probably the most basic starts from the following: let V be a real n-dimensional vector space. Let $\mathcal{A}$ be the set of all ordered bases for V. Then we can write $\mathcal{A}$ as the disjoint union of two subsets:

$$\mathcal{A} = \mathcal{A}_1 \coprod \mathcal{A}_2,$$

such that the change of basis matrix between any two bases in $\mathcal{A}_1$, or between any two bases in $\mathcal{A}_2$ has positive determinant, and such that the change of basis matrix between a basis in $\mathcal{A}_1$ and a basis in $\mathcal{A}_2$ (or vice versa) has negative determinant.

We have that, for any ordered basis $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ in $\mathcal{A}_1$, the basis $(\mathbf{v}_1, \dots, -\mathbf{v}_i, \dots, \mathbf{v}_n)$ is in $\mathcal{A}_2$ and vice versa. Similarly, for any basis $(\mathbf{v}_1, \dots, \mathbf{v}_i, \mathbf{v}_{i+1}, \dots, \mathbf{v}_n)$ in $\mathcal{A}_1$, the basis $(\mathbf{v}_1, \dots, \mathbf{v}_{i+1}, \mathbf{v}_i, \dots, \mathbf{v}_n)$ is in $\mathcal{A}_2$.

Remark C.1. It is important to remark however that a priori we can't single out $\mathcal{A}_1$ or $\mathcal{A}_2$, the only thing we can do is to divide the whole set $\mathcal{A}$ in two pieces.

Definition C.2. *An orientation of an n-dimensional real vector space is the choice of one of the two pieces $\mathcal{A}_1$ or $\mathcal{A}_2$ of $\mathcal{A}$. Bases in that subset will be referred to* positively oriented *bases, bases in the other as* negatively oriented *bases.*

One way to choose an orientation is to pick out a single oriented basis: we can then say that all other bases with positive determinant change of basis matrix relative to this basis are the positive ones.

Remark C.3. Note also that this is often done implicitly for $V = \mathbb{R}^n$: the standard oriented basis here picks out an orientation.

Remark C.4. An orientation for V immediately determines an orientation for V^*, and vice versa: an ordered basis for V^* is said to be positively-oriented if it is dual to a positively-oriented ordered basis for V.

Going back to the cross-product, we now see that the cross-product can exactly be defined in a 3-dimensional oriented euclidean vector space: the vector $\mathbf{v} \times \mathbf{w}$ is the unique vector with length $|\mathbf{v}|\ |\mathbf{w}|\ |\sin(\theta)|$ orthogonal to both $\mathbf{v}$ and $\mathbf{w}$ that makes the triple

$$(\mathbf{v}, \mathbf{w}, \mathbf{v} \times \mathbf{w})$$

into a positively oriented basis (if $\mathbf{v}$ and $\mathbf{w}$ are linearly independent, otherwise $\mathbf{v} \times \mathbf{w}$ is $\mathbf{0}$).

There are a number of other ways to define orientations though. One of them uses the *top exterior product* $\bigwedge^n V$. Recall that this is always a 1-dimensional vector space, and that if $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ is an ordered basis for V, then

$$\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_n$$

will be a basis for $\bigwedge^n V$.

Proposition C.5. *An orientation for V is equivalent to the choice of a* direction *for $\bigwedge^n V$, i.e. the choice of which non-zero elements in $\bigwedge^n V$ are positive and which are negative.*

Now if V is Euclidean, then so is $\bigwedge^n V$. This means that the choice of an orientation on a Euclidean real vector space induces an isomorphism

$$\bigwedge^n V \cong \mathbb{R}, \tag{23}$$

by identifying the single positive vector of norm 1 in $\bigwedge^n V$ with $1 \in \mathbb{R}$.

C.2 Exterior products as a generalization of the cross product

We now want to argue that the exterior product gives a generalization of the cross-product to vector spaces that are not necessarily 3-dimensional, or Euclidean, or oriented.

To do this let's recall the following: if a vector space V is Euclidean, the inner product provides an isomorphism

$$\Phi : V \cong V^* \tag{24}$$

defined by $\Phi(\mathbf{v})(\mathbf{w}) = \langle \mathbf{v}, \mathbf{w} \rangle$. If V is moreover Euclidean, and we identify $\bigwedge^n V^*$ with $\mathbb{R}$, we get a morphism

$$\Psi : \bigwedge^{n-1} V^* \to (V^*)^* \cong V : \Psi(\alpha)(\beta) = \alpha \wedge \beta \in \bigwedge^n V^* \cong \mathbb{R}. \tag{25}$$

Proposition C.6. *Let V be a real 3-dimensional oriented vector space. Then for any vectors $\mathbf{v}$, $\mathbf{w} \in V$ we have*

$$\mathbf{v} \times \mathbf{w} = \underbrace{\Psi\Big(\overbrace{\Phi(\mathbf{v}) \wedge \Phi(\mathbf{w})}^{\in \bigwedge^2 V^*}\Big)}_{\in V}.$$

C.3 Vector calculus and differential forms

Now, you may not be too impressed yet by the previous discussion. Indeed, it seems like a bit of a convoluted way to write the cross-product. Why bother with all that if you are just interested in three dimensions?

However, things become much more impressive when we start thinking about *vector calculus* in three dimensions.

C.3.1 Recap of vector calculus

So we will now think about vector fields on $\mathbb{R}^3$, i.e. maps $\mathbb{R}^3\mathbb{R}^3$. Everthing we will say can be generalised to oriented Riemannian 3-manifolds, this is discussed in courses like Manifolds and Geometry of General Relativity.

In vector calculus we have several differential operators: the gradient, which turns a scalar function into a vector field. In orthogonal coordinates on $\mathbb{R}^3$ this can be written, for a smooth real-valued function $f : \mathbb{R}^3 \to \mathbb{R}$, as

$$\mathrm{grad}(f) = \nabla f = \frac{\partial f}{\partial x^1}\frac{\partial}{\partial x^1} + \frac{\partial f}{\partial x^2}\frac{\partial}{\partial x^2} + \frac{\partial f}{\partial x^3}\frac{\partial}{\partial x^3}.$$

We also have the divergence, which turns a vector field into a scalar function. It can be written, for a smooth vectorfield $\mathbf{E} = E^1\frac{\partial}{\partial x^1} + E^2\frac{\partial}{\partial x^2} + E^3\frac{\partial}{\partial x^3}$, as

$$\mathrm{div}(\mathbf{E}) = \nabla.\mathbf{E} = \frac{\partial E^1}{\partial x^1} + \frac{\partial E^2}{\partial x^3} + \frac{\partial E^3}{\partial x^3}.$$

Note that both of these operations only need a Euclidean structure: they work in any dimension, and don't need an orientation.

The last differential operator is the curl, which takes a vector field, and returns a vector field. This one is particular to dimension 3, and needs an orientation. It is defined in (oriented orthonormal) coordinates by

$$\mathrm{curl}(\mathbf{E}) = \nabla \times \mathbf{E} = \left(\frac{\partial E^3}{\partial x^2} - \frac{\partial E^2}{\partial x^3}\right)\frac{\partial}{\partial x^1} + \left(\frac{\partial E^1}{\partial x^3} - \frac{\partial E^3}{\partial x^1}\right)\frac{\partial}{\partial x^2} + \left(\frac{\partial E^1}{\partial x^2} - \frac{\partial E^2}{\partial x^1}\right)\frac{\partial}{\partial x^3}.$$

Just as an indication of the relevance of these differential operators, let us write down the Maxwell equations in their standard (vector calculus) form:

$$\boxed{\begin{aligned} \nabla.\mathbf{E} &= \frac{\rho}{\epsilon_0}, \\ \nabla.\mathbf{B} &= 0, \\ \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t}, \\ \nabla \times \mathbf{B} &= \mu_0\mathbf{J} + \mu_0\epsilon_0\frac{\partial \mathbf{E}}{\partial t}. \end{aligned}}$$

These are the basic equations that describe all of electro-magnetism (here $\mathbf{E}$ is the electric field, $\mathbf{B}$ is the magnetic field, $\mathbf{J}$ is the electric current density (all three of these are vector fields on $\mathbb{E}^3$), and ρ is the electric charge density (given by a scalar function). ϵ_0 and μ_0 are constants of nature.

These differential operators satisfy several identities, two important of which are

$$\begin{aligned} \mathrm{curl}(\mathrm{grad}(f)) &= \nabla \times \nabla(f) = \mathbf{0}, \\ \mathrm{div}(\mathrm{curl}(\mathbf{E})) &= \nabla.(\nabla \times \mathbf{E}) = 0. \end{aligned}$$

for any function $f \in C^\infty(\mathbb{R}$ and vectorfield $\mathbf{E} \in \mathrm{Vec}(\mathbb{R}^3$.

C.3.2 Translation into differential forms

Our claim is now that one should really understand this in terms of differential forms, and that these identities are just incarnations of the much more general equation

$$d^2 = 0.$$

Indeed, if we denote by $\mathrm{Vec}(\mathbb{R}^3)$ the space of all (smooth) vectorfields on $\mathbb{R}^3$, we can write down the following commutative diagram

$$\begin{tikzcd}
C^\infty(\mathbb{R}^3) \arrow[r, "\cong"] \arrow[d, "\nabla"'] & \Omega^0(\mathbb{R}^3) \arrow[d, "d"] \\
\mathrm{Vec}(\mathbb{R}^3) \arrow[r, "\cong"] \arrow[d, "\nabla\times"'] & \Omega^1(\mathbb{R}^3) \arrow[d, "d"] \\
\mathrm{Vec}(\mathbb{R}^3) \arrow[r, "\cong"] \arrow[d, "\nabla."'] & \Omega^2(\mathbb{R}^3) \arrow[d, "d"] \\
C^\infty(\mathbb{R}^3) \arrow[r, "\cong"] & \Omega^3(\mathbb{R}^3)
\end{tikzcd}$$

where all horizontal maps are isomorphisms: $C^\infty(\mathbb{R}^3) \cong \Omega^0(\mathbb{R}^3)$ is just the definition of 0-forms, $\mathrm{Vec}(\mathbb{R}^3 \cong \Omega^1(\mathbb{R}^3)$ just uses the euclidean structure, as in (24), $\mathrm{Vec}(\mathbb{R}^3) \cong \Omega^2(\mathbb{R}^3)$ uses (25), using the euclidean structure and the orientation, and finally $C^\infty(\mathbb{R}^3) \cong \Omega^3(\mathbb{R}^3)$ relies on the identification $\bigwedge^3 T_p^*(\mathbb{R}^3 \cong \mathbb{R}$ as in (23), which also uses the orientation and the euclidean structure.

Note that the left column can be generalized to arbitrary oriented Riemannian 3-manifolds, but the right column always exists for any manifold – in any dimension, without the need of an orientation or a Riemannian structure.

C.3.3 Integral theorems – one theorem to rule them all

C.4 Moving into the complex world

Remark C.7. As the section below aims to show the Cauchy Residue Theorem (a result from complex analysis) is a special case of Stokes' theorem, it is only meant to be understood by people having familiarity with both, e.g. after having taking the Honours Complex Variables course.

C.4.1 Complexifying tangent and co-tangent spaces

Let us now slightly change perspective, and start thinking about the complex numbers. If D is an open set in $\mathbb{C}^n$, we can identify the latter with $\mathbb{R}^{2n}$, and think of D as a subset of $\mathbb{R}^{2n}$. If we used complex coordinates $(z^1, \dots, z^n)$ before, with $z^j = x^j + iy^j$ (here $i^2 = -1$), we can use the $(2n)$ real coordinates $(x^1, y^1, x^2, y^2, \dots, x^n, y^n)$ on D.

Now, because of the complex coordinates, we actually have that each T_pD is a complex vector space. Complex multiplication is determined by

$$I\left(\frac{\partial}{\partial x^j}\Big|_p\right) = \frac{\partial}{\partial y^j}\Big|_p. \tag{26}$$

Also the cotangent spaces are complex, with

$$I\left(dx^j|_p\right) = dy^j|_p. \tag{27}$$

Here I denotes multplication by $\sqrt{-1}$, i.e. $I^2 = -1$ (it will become clear below why we use a different notation rather than i for this). It is very important to remark that, as a set, T_pD still just consists of maps (derivations at p) from the vector space of smooth *real*-valued functions to $\mathbb{R}$.

Because of the complex context, we really want to change this: we want to be able to take exterior derivatives of *complex* valued functions on D. This can be done, but it requires a little attention (note that in many places this explanation is skipped, or kept implicit).

We have to begin by taking the tensor product (over the reals - this is important) of each T_pD with $\mathbb{C}$, and similarly we tensor T_p^*D with $\mathbb{C}$. But we have to be careful! On these new space there are now two different complex structures, one from (26) and (27), and the other one from the tensor product with $\mathbb{C}$. Just to be clear, we shall indicate complex multiplication from the latter (see Exercise B.24) with little i. So, to be precise, the little i is the complex structure obtained by taking the tensor products

$$(T_pD)_{\mathbb{R}} \otimes_{\mathbb{R}} \mathbb{C} \quad \text{and} \quad (T_pD)_{\mathbb{R}} \otimes_{\mathbb{R}} \mathbb{C}.$$

It is given by

$$(a+ib)\big(\mathbf{v} \otimes (c+id)\big) = \mathbf{v} \otimes \big((a+ib)(c+id)\big).$$

So now we can really work with complex-valued functions – we can take their derivatives, as well as apply d to them – and this is now just complex linear. In particular, we now have

$$d(z^j) = dx^j + idx^j, \quad \text{and} \quad d\overline{z}^j = dx^j - idy^j.$$

One easily verifies that these complex 1-forms form a basis for each

$$\left(T_p^*D\right)_{\mathbb{C}} = \left(T_p^*D\right)_{\mathbb{R}} \otimes_{\mathbb{R}} \mathbb{C}.$$

It is very useful to also use their dual basis for $(T_pD)_{\mathbb{C}} = (T_pD)_{\mathbb{R}} \otimes_{\mathbb{R}} \mathbb{C}$, and this is given by the 'complex tangent vectors'

$$\left.\frac{\partial}{\partial z^j}\right|_p = \frac{1}{2}\left(\left.\frac{\partial}{\partial x^j}\right|_p - i\left.\frac{\partial}{\partial y^j}\right|_p\right) \quad \text{and} \quad \left.\frac{\partial}{\partial \overline{z}^j}\right|_p = \frac{1}{2}\left(\left.\frac{\partial}{\partial x^j}\right|_p + i\left.\frac{\partial}{\partial y^j}\right|_p\right). \tag{28}$$

Note the signs and the factor $\frac{1}{2}$ here! It is a bit hard to give a clean geometric interpretation of these complex tangent vectors – you should think of them as partial derivatives in a 'complexified direction'. In particular, even when applied to a smooth real function on D they can return complex numbers. These 'complex directional derivatives' $\frac{\partial}{\partial z^j}$ and $\frac{\partial}{\partial \overline{z}^j}$ are sometimes referred to as *Wirtinger derivatives.*

Now, it is important to go back to the results of exercise B.25. It was shown there that for every complex vector space V, we have a canonical isomorphism of complex vector spaces

$$V_{\mathbb{R}} \otimes_{\mathbb{R}} \mathbb{C} \cong V \oplus \overline{V}.$$

We want to use this isomorphism at each point $p \in D$ now, which gives us

$$(T_pD)_{\mathbb{C}} \cong T_pD \oplus \overline{T_pD}$$

and

$$(T_pD)_{\mathbb{C}} \cong T_pD \oplus \overline{T_pD}.$$

It is easy to be confused here: the complex multiplication on the left hand sides is given by i, and on the right hand sides it is given by I in the first summand, and $-I$ in the second summand. To clarify things, in this decomposition the summand of $(T_pD)_{\mathbb{C}}$ corresponding to T_pD is often denoted by $T_p^{1,0}D$, and the summand corresponding to $\overline{T_pD}$ by $T_p^{0,1}D$.

Now, luckily, this decomposition interacts really well with our notation: it turns out that

$$\frac{\partial}{\partial z^j}\Big|_p \in T_p^{1,0}D \quad \text{and} \frac{\partial}{\partial \overline{z}^j}\Big|_p \in T_p^{0,1}D.$$

Similarly we have

$$dz^j|_p \in T_p^{*,1,0}D \quad \text{and} \quad d\overline{z}^j \in T_p^{*,0,1}D.$$

Also globally we have these decompositions, and so we can write the space of complexified one-forms as a direct sum:

$$\Omega^1(D)_{\mathbb{C}} = \Omega^1(D) \otimes_{\mathbb{R}} \mathbb{C} = \Omega^{1,0}(D) \oplus \Omega^{0,1}(D),$$

where $\alpha \in \Omega^{1,0}(D)$ if and only if $\alpha|_p \in T_p^{*,1,0}D$ for every $p \in D$.

Moreover, for any smooth, complex valued function f on D we can take the exterior derivative $df \in \Omega^1(D)_{\mathbb{C}}$, and at each point p break it up into two pieces, which we shall denote by $\partial f \in \Omega^{1,0}(D)$ and $\overline{\partial} f \in \Omega^{0,1}(D)$. In short, we write

$$d = \partial + \overline{\partial},$$

and we have

$$\partial f = \frac{\partial f}{\partial z^j} dz^j \quad \text{and} \quad \overline{\partial} f = \frac{\partial f}{\partial \overline{z}^j} d\overline{z}^j.$$

Now, a quick look at (28) will show that this operator $\overline{\partial}$ actually shows up in the Cauchy-Riemann equations, that test if a smooth complex-valued function on D is holomorphic. Indeed, we have that

$$f \text{ is holomorphic} \quad \Leftrightarrow \quad \overline{\partial} f = 0.$$

Remark C.8. In complex geometry, when working with complex tensors and differential forms, one often slightly extends the notational conventions of the Ricci Calculus by using *barred* indices. For example, one would write (in complex dimension three)

$$\begin{aligned} g_{k,\bar{l}} dz^k \wedge d\overline{z}^l = g_{1\bar{1}} dz^1 \wedge d\overline{z}^1 + g_{1\bar{2}} dz^1 \wedge d\overline{z}^2 + \dots \\ + g_{2\bar{1}} dz^2 \wedge d\overline{z}^1 + g_{2\bar{2}} dz^2 \wedge d\overline{z}^2 + \dots + g_{3\bar{3}} dz^3 \wedge d\overline{z}^3. \end{aligned}$$

C.4.2 Stokes' theorem and the Cauchy Integral and Residue Theorems

With all of this set-up, we can see that the famous Cauchy Residue Theorem just reduces to Stokes theorem in two real dimensions. The key thing to observe is the following:

Lemma C.9. *If f is a holomorphic function of one complex variable z, then $f dz$ is a closed form (i.e. its exterior derivative is zero):*

$$d\big(f dz\big) = 0.$$

Proof. Indeed, we have

$$d\big(f dz\big) = (\partial + \overline{\partial})\big(f dz\big) = \frac{\partial f}{\partial z} dz \wedge dz + \overline{\partial}\big(f\big)\, dz = 0 + 0 = 0.$$

□

Finally, to state the Cauchy Residue Theorem, recall that for a function f that is meromorphic at a point z_0, so that we can write

$$f(z) = \sum_{i=-N}^{\infty} a_i (z - z_0)^i$$

near z_0, we define the residue of f at z_0 to be

$$\mathrm{Res}(f, z_0) = a_{-1}.$$

Theorem C.10 (Cauchy Residue Theorem)**.** *Let γ be a closed regular curve in $\mathbb{C}$ with no self-intersection, which bounds an open subset $U \subset \mathbb{C}$. Assume γ is parametrized in a counter-clockwise sense. Let f be a function that is meromorphic on an open neighborhood of $\overline{U}$, with poles at a_k, $k = 1, \ldots, m$. Then*

$$\oint_\gamma f(z)\, dz = 2\pi i \sum_k \mathrm{Res}(f, a_k).$$

Here $\oint_\gamma f(z)\, dz$ is a complex line integral. By writting both $f(z)$ and dz as the sum of a real and imaginary part, one can write this as a complex number whose real and imaginary parts are ordinary line integrals.

Sketch of proof. Let μ_k be a little closed circle around each a_k, oriented in the counter clockwise sense, close enough so that the meromorphic expansion of $f(z)$ near a_k holds on μ_k. Then a direct calculation shows that

$$\oint_{\mu_k} f(z)\, dz = 2\pi i \,\mathrm{Res}(f, a_k).$$

We can now apply Stokes' theorem to the subset of $\overline{U}$ obtained by cutting out the disks bounded by the μ_k – let us call this M (a submanifold with boundary). Since f is holomorphic on M, Stokes' theorem and Lemma C.9 give

$$0 = \int_M d(\, f(z)\, dz\,) = -\oint_\gamma f(z)\, dz + \sum_k \oint_{\mu_k} f(z)\, dz = -\oint_\gamma f(z)\, dz + 2\pi i \sum_k \mathrm{Res}(f, a_k).$$

□