MATH 314 - Advanced Calculus

McGill University - Winter 2025

Jack Borthwick

This version: November 10, 2025

The most up to date version can be found here.

Contents

Chapter -1

What is this course about?

The aim of this course is to polish up your knowledge of calculus and provide some exposure to the modern mathematical theory that was developed through the study of diverse physical phenomena such as gravitation, electromagnetism, movement of fluid bodies, etc.

A central notion we will become accustomed to is that of a *field of quantities*, let us first consider a few examples:

Examples

1. **From mechanics: the velocity vector $\vec{v}$ of a particle.**

   Classically, the movement of a particle P in space $\mathscr{E}$, is represented as a *curve* $t \in \mathbb{R} \mapsto P(t) \in \mathscr{E}$. This notation means that to every time t is mapped to a point $P(t)$ in space $\mathscr{E}$: the position of the particle at time t. Making the usual assumptions about $\mathscr{E}$ in classical mechanics, one may fix an arbitrary origin O and it is possible to locate the particle P at anytime t if we know the position vector $\overrightarrow{OP(t)}$. The velocity vector at time t_0 is then defined to by:

   $$\vec{v}(t) = \lim_{t \to t_0} \frac{\overrightarrow{OP}(t) - \overrightarrow{OP}(t_0)}{t - t_0} = \lim_{t \to t_0} \frac{\overrightarrow{P(t_0)P(t)}}{t - t_0}.$$

   Observe that it is independent on the choice of origin and only depends on the point $P(t_0)$, hence we can assign to every point $P(t)$ on the trajectory a vector $\vec{t}$: this is a vector field along a curve.

   Note that the vector is more than just 3 real numbers: I need to fix a frame to represent $\vec{v}$ as an element $\begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}$ of $\mathbb{R}^3$ and these numbers v_x, v_y, v_z depend on that choice.

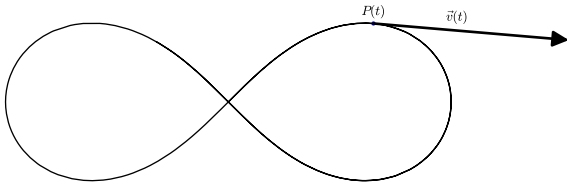
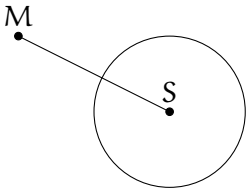


Figure 1: A curve in a plane



Furthermore, if one fixes a reference frame then the trajectory can be modelled as a function from: $\mathbb{R} \to \mathbb{R}^3$.

2. **Gravitation**

Assume now that you are considering the gravitational force felt by a small test particle with mass $m << M_S$ due to a nearby star S with mass M_S. It is subject to the force:

$$\overrightarrow{F} = -\frac{mGM_S}{||\overrightarrow{SM}||^3}\vec{SM}.$$

Save m, this expression only depends on S and the point of M in space so one may rewrite:

$$\overrightarrow{F} = m\overrightarrow{\mathscr{G}(M)},$$

where here:

$$\mathscr{G}(M) = -\frac{GM_S}{||\overrightarrow{SM}||^3}\overrightarrow{SM}.$$

This is a *vector field*, which we can refer to as the gravitational field of the star. It assigns to each point M in space a vector. If we choose a reference frame, then this will be assimilated to a function from: $\mathbb{R}^3 \to \mathbb{R}^3$.

3. **Quantum mechanics**

In quantum mechanics, a particle is described by its wave function, which is a map ψ that to each point of space $\mathscr{E}$ assigns a complex number, i.e.: $\psi : \mathscr{E} \to \mathbb{C}$. Once more, a choice of reference frame leads us to model ψ as a map $\mathbb{R}^3 \to \mathbb{C}$. This is known as a complex scalar field.

The wave function is satisfies a *partial differential equation* known as the Schrödinger equation:

$$-i\frac{\partial\psi}{\partial t} = \hat{H}\psi.$$

4. **Electromagnetism: Maxwell's equations**

   Classically, the electromagnetic field is described by two vector fields $\vec{E}, \vec{B}$. They satisfy a famous system of equations known as the Maxwell's equations. This case aims to give you the tools to understand and analyse them, they are often written on one of the following ways.

$$\begin{cases} \text{div}\ \vec{E} = \frac{\rho}{\varepsilon_0}, & \vec{\nabla} \cdot \vec{E} = \frac{\rho}{\varepsilon_0}, \\ \overrightarrow{\text{curl}}\vec{E} = -\frac{\partial \vec{B}}{\partial t}, & \vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}, \\ \text{div}\vec{B} = 0, & \vec{\nabla} \cdot \vec{E} = 0, \\ \overrightarrow{\text{curl}}\vec{B} = \mu_0(\vec{j} + \varepsilon_0 \frac{\partial \vec{E}}{\partial t}) & \vec{\nabla} \times \vec{B} = \mu_0(\vec{j} + \varepsilon_0 \frac{\partial \vec{E}}{\partial t}). \end{cases}$$

What these examples show is that it will be interesting to understand how to do calculus on functions between *vector spaces*. $\mathbb{R}, \mathbb{R}^2, \mathbb{R}^3, \mathbb{R}^n$ are the classical examples (and where we work most of the time) but it is conceptually interesting to develop the theory for an arbitrary vector space. Vaguely (and this is the extent in which we need to understand these things) this is some set V, composed of vectors, in which we have a way of adding elements together, i.e. a $+$ operation, as well as a way of scaling them using a *scalar* which for us is just a real number. i.e. we can make sense of $\lambda \cdot \vec{v}$, for $\lambda \in \mathbb{R}$ and $\vec{v} \in V$.

We will not have any issue in working with an abstract vector space as the algebraic operations are modelled on the usual algebraic operations available to us in vector spaces and *behave the same way*. You can look up the formal axiomatic definition in any textbook on Linear Algebra.

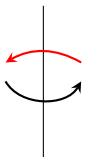
Chapter 0

An (augmented) review of some basic vector geometry.

0.1 Orientation

← Start Lecture 1

To get started, we will review some basic notions from linear algebra by studying the notion of *orientation*. This will enable us to work with the notions of *basis* and determinant of a matrix.



The idea of things have some defined orientation is relatively, say for instance, which way is up or down, left or right? When we look at something spin, we can ask which *way* is it spinning.

We can capture something of this idea with notions from linear algebra, let us first consider some examples:

Orienting a line

The simplest case is a 1 dimensional vector space or, in other words, a line. Consider the following figure:

Figure 1: Vectors on a line

Clearly the red $\vec{r}$ and blue $\vec{b}$ vectors are oriented in different ways, but how can I appreciate this algebraically ?

Well, since a line is 1 dimensional, then any non-zero vector generates it. In other words there is a constant $a \neq 0$: such that:

$$\vec{r} = a\vec{b}.$$

Intuitively the fact that $\vec{r}$ and $\vec{b}$ are pointing in different ways means that:

$$a < 0,$$

and had they been pointing the same way we would have:

$$a > 0.$$

So the *sign* of the constant between them can be used to ascertain if they are pointing the same way or not. We can orient the line by declaring that the positive direction is one of the vectors, and then determine the orientation of any other vector by comparing to this reference.

Orienting a plane

When we do trigonometry it is usual to define the positive sense of rotation to be anti-clockwise as in the left below: We could say that the sense of rotation is defined

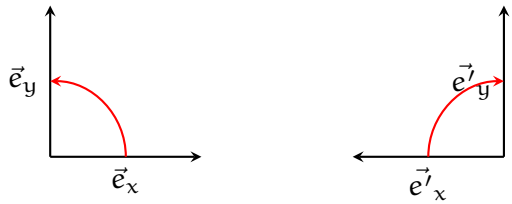


Figure 2: Orientations of a plane

by that needed to send $\vec{e}_x \to \vec{e}_y$. If we apply this principle to the diagram on the right then we find the *opposite* sense of rotation.

The couples $\mathscr{B} = (\vec{e}_x, \vec{e}_y)$ and $\mathscr{B}' = (\vec{e'}_x, \vec{e'}_y)$ are both *ordered* basis of the plane, and we can again compare them by expressing one in function of the other, writing down the change of basis matrix.

Here from the diagram we see that: $\vec{e'}_x = -\vec{e}_x$ and $\vec{e'}_y = \vec{e}_y$. So the change of basis matrix is given by:

$$P_{\mathscr{B}\to\mathscr{B}'} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

Remark 0.1.1. The change of basis matrix between two bases $\mathscr{B} = (\vec{e}_1, \ldots, \vec{e}_n)$ (old basis) and $\mathscr{B}' = (\vec{e}'_1, \ldots, \vec{e}'_n)$ is obtained by placing in the i-th column the coordinates of the vector e'_i expressed in the basis $\mathscr{B}'$. Explictly, assume that:

$$\vec{e}'_i = \sum_{k=1}^{n} P_{kj} \vec{e}_k,$$

then:

$$P_{\mathscr{B}\to\mathscr{B}'} = \begin{pmatrix} P_{11} & \cdots & P_{1n} \\ \vdots & \ddots & \vdots \\ P_{n1} & \cdots & P_{nn} \end{pmatrix}.$$

Observe that in this case it is $\det P_{\mathscr{B}\to\mathscr{B}'} < 0$. So the orientations of bases can be compared by considering the sign of the determinant of the change of basis matrix.

These examples motivate the following definition:

Definition 0.1

Let $\mathscr{B}$ and $\mathscr{B}'$ be two *ordered bases* of a vector space V. We say that they have *same orientation* if $\det P_{\mathscr{B}\to\mathscr{B}'} > 0$ and *opposite* orientation otherwise.

Example 0.1. If $\mathscr{B} = (\vec{e}_1, \vec{e}_2)$ and $\mathscr{B}' = (\vec{e}_2, \vec{e}_1)$ then:

$$P_{\mathscr{B}\to\mathscr{B}'} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

and $\det P_{\mathscr{B}\to\mathscr{B}'} < 0$ so they have *opposite* orientation. This example illustrates the importance of using *ordered* bases: changing the order of the basis vectors can change its orientation. ◇

In order to define an orientation for some space, we need to decide which bases are *positive* and which are *negative*. This is a convention that needs to be made.

One may check that the relation "has same orientation as" defines an equivalence relation on the set of all ordered basis of a vector space which has two equivalence classes. In other words, since a basis either has same orientation or not as another basis, we can split the set of all ordered basis up into two classes in which all the basis have same orientation. A choice of one of these classes is what we refer to as an *orientation*.

Definition 0.2: Orientation

Let V be a finite dimensional vector space. An **orientation** of V is a choice of one of the two equivalence classes. When such a choice has been made we refer to V as an *oriented vector space*.

In practice, you only need to know *one* representative of the equivalence class; so in other words it is sufficient to choose 1 reference basis $\mathscr{B} = (\vec{e}_1, \ldots, \vec{e}_n)$ that you declare positive and subsequently any other basis $\mathscr{B}'$ is:

- *positive* or *positively oriented* if $\mathscr{B}'$ has same orientation as $\mathscr{B}$,
- *negative* or *negatively oriented* if $\mathscr{B}'$ has opposite orientation as $\mathscr{B}$

Example 0.2. The standard orientations of the plane and space are given by the following bases: In 3-space this is the mathematical statement corresponding to the

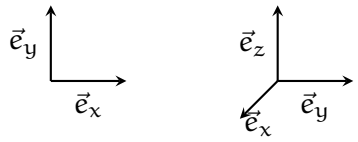


Figure 3: Standard orientations of planes and 3-space

right-hand rule. ◇

Example 0.3. Let $V = \mathbb{R}^3$ and $\mathscr{B} = (\vec{i}, \vec{j}, \vec{k})$ the standard basis. Orient V with $\mathscr{B}$ i.e., $(\vec{i}, \vec{j}, \vec{k})$ is positively oriented then:

- $\mathscr{B}' = (-\vec{i}, \vec{j}, \vec{k})$ is *negatively oriented*. $P_{\mathscr{B}\to\mathscr{B}'} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.
- $\mathscr{B}' = (\vec{j}, -\vec{i}, \vec{k})$ is *positively oriented*. $P_{\mathscr{B}\to\mathscr{B}'} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

End Lecture 1→ Check this using the right hand rule ! ◇

0.2 General Euclidean spaces

Start Lecture 2→

The standard n-dimensional Euclidean space $\mathbb{R}^n$ comes with a natural scalar product (also referred to as an inner product, or the dot product). Defined by:

$$\vec{x} \cdot \vec{y} = \vec{x} \bullet \vec{y} = \langle \vec{x}, \vec{y} \rangle = (\vec{x} \,|\vec{y}) = \sum_{i=1}^{n} x_i y_i \in \mathbb{R},$$

where: $\vec{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$. We have listed a number of different notations used in the literature for this product. They often come in handy, when there are many other different dots $\cdot$ indicating other types of notation.

As always it can be a conceptual aid to extract out the main algebraic properties of our inner product, and work only with those. This enables us to treat many cases

uniformly. In this instance, it is straightforward to check that the dot product on $E = \mathbb{R}^n$ has the following properties:

1. Linearity in the first variable:
$$\forall \vec{u}, \vec{v}, \vec{w} \in E, \quad \lambda \in \mathbb{R}, \quad (\lambda\vec{u} + \vec{w}) \bullet \vec{v} = \lambda(\vec{u} \bullet \vec{v}) + \vec{w} \bullet \vec{v}.$$
2. Symmetry:
$$\forall \vec{u}, \vec{v} \in E, \quad \vec{u} \bullet \vec{v} = \vec{v} \bullet \vec{u}$$
3. Positivity:
$$\forall \vec{u} \in E, \quad \vec{u} \bullet \vec{u} \geqslant 0$$
4. Definiteness:
$$\forall \vec{u} \in E, \qquad \vec{u} \bullet \vec{u} = 0 \quad \Rightarrow \quad \vec{u} = 0.$$

Observe that the properties of linearity in the first variable plus and symmetry imply that the inner product is also linear in the second variable so that the assignment is *bilinear*.

An n-dimensional Euclidean space, or inner product space, is *any* vector space E of dimension n on which we have given a rule, that allows us to combine two vectors $\vec{v}_1$ and $\vec{v}_2$ to obtain a real number $\mathbb{R}$, formalised as a mapping :

$$\begin{array}{ccc} E \times E & \longrightarrow & \mathbb{R} \\ (\vec{v}_1, \vec{v}_2) & \longmapsto & \vec{v}_1 \bullet \vec{v}_2 \end{array}$$

that obeys the algebraic rules above.

Example 0.4. Let $E = M_n(\mathbb{R})$ the set of square $n \times n$ matrices with real coefficients, this is a vector space. Define:

$$\langle A, B \rangle = \mathbf{tr}({}^tAB),$$

where ${}^tA$ is the transpose matrix of $A = (A_{ij})$ whose coefficients are $({}^tA)_{ij} = A_{ji}$, we also recall that if $A = (A_{ij})$ then its trace is defined to be:

$$\mathbf{tr}(A) = \sum_{i=1}^{n} A_{ii}.$$

This defines a scalar product on $M_n(\mathbb{R})$.

It is a straightforward exercise to check that the tr is a *linear map* from $M_n(\mathbb{R})$ into $\mathbb{R}$, and that the transpose ^{t} is a linear map from $M_n(\mathbb{R}) \to M_n(\mathbb{R})$ i.e.

$$\begin{cases} \mathbf{tr}(\lambda A + B) = \lambda \mathbf{tr}(A) + \mathbf{tr}(B), \\ {}^t(\lambda A + B) = \lambda {}^tA + {}^tB \end{cases} \qquad \lambda \in \mathbb{R}, \quad A, B \in M_n(\mathbb{R})$$

and deduce from it linearity in the first variable of A.

Symmetry follows from the facts:

$$\begin{cases} {}^t(AB) = {}^tB\,{}^tA, \\ \mathbf{tr}(A) = \mathbf{tr}({}^tA), \end{cases}$$

which can be verified by explicit computation. The coefficient at position (i,j) in AB is given by:

$$\sum_{k=1}^{n} A_{ik}B_{kj},$$

so the coefficient coefficient in position (i,j) of ${}^tAB$ is

$$\sum_{k=1}^{n} B_{ki}A_{jk} = \sum_{k=1}^{n} ({}^tB)_{ik}({}^tA)_{kj}.$$

Indeed for any matrices $A, B \in M_n(\mathbb{R})$ we have:

$$\langle A, B\rangle = \mathbf{tr}({}^tAB) = \mathbf{tr}({}^t({}^tAB)) = \mathbf{tr}({}^tB\,{}^t({}^tA)) = \mathbf{tr}({}^tBA) = \langle B, A\rangle.$$

$\diamond$

The final properties also follow from an explicit computation, let $A = (A_{ij}) \in M_n(\mathbb{R})$ then:

$$({}^tAA)_{ij} = \sum_{k=1} A_{ki}A_{kj},$$

and then:

$$\langle A, A\rangle = \sum_{i,k=1}^{n} A_{ki}^2 \geqslant 0.$$

If the sum vanishes then each term of the sum must vanish individually, but then this means that all of the coefficients of the matrix must vanish, i.e. $A = 0$.

Remark 0.2.1. The actual definition of an inner product does not rely in anyway on the dimension of the vector space E, and continues to be meaningful if E is infinite dimensional; these spaces are known as *pre-Hilbert spaces*.

When doing geometry 2 or 3 dimensions, you are accustomed to working with orthonormal bases. Their definition extends straightforwardly to higher dimensions i.e. these are bases $(\vec{e}_1, \ldots, \vec{e}_n)$ of E with the following property:

$$\vec{e}_i \cdot \vec{e}_j = \delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise,} \end{cases}$$

One of the fundamental results of the theory of Euclidean spaces, that we state without proof is:

Proposition 0.1: Orthonormal basis

Any Euclidean space E has an orthonormal basis.

Remark 0.2.2. The proof we omit here is actually algorithmic, i.e. given an arbitrary basis in E we can use it to construct an orthonormal basis. The process is known as the **Gramm-Schmidt orthonormalisation procedure**.

Example 0.5. Let $E = M_2(\mathbb{R})$ and $\langle A, B\rangle = \mathrm{tr}({}^tAB)$ as above, then the following matrices form an orthonormal basis of E.

$$\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}.$$

◇

The major theoretical consequence of Proposition 0.1 is that, *up to a choice of orthonormal basis*, **all** Euclidean spaces behave like $\mathbb{R}^n$ for some positive integer n!

To see this, take an arbitrary space E with a scalar product $\bullet$ (for example: $E = M_n(\mathbb{R})$ and $\langle A, B\rangle = \mathrm{tr}({}^tAB)$ as above.)

Let $(\vec{e}_1, \ldots, \vec{e}_n)$ be an orthonormal basis. By definition of a basis, all vectors $\vec{v}$,$\vec{u} \in E$ can be uniquely decomposed into a linear combination of the basis elements, i.e. there are unique real numbers[1] $v_1, \ldots, v_n$ and $u_1, \ldots, u_n$ such that:

$$v = \sum_{i=1}^{n} v_i \vec{e}_i, \qquad u = \sum_{i=1}^{n} u_i \vec{e}_i.$$

Let us calculate, $\vec{u} \bullet \vec{v}$ using only the rules we gave above:

$$\vec{u} \bullet \vec{v} = \left(\sum_{i=1}^{n} u_i \vec{e}_i\right) \bullet \left(\sum_{j=1}^{n} v_j \vec{e}_j\right),$$

by linearity in the first variable we have:

$$\vec{u} \bullet \vec{v} = \sum_{i=1}^{n} u_i \vec{e_i} \bullet \left(\sum_{j=1}^{n} v_j \vec{e}_j\right),$$

and then by linearity in the second variable:

$$\vec{u} \bullet \vec{v} = \sum_{i=1}^{n}\sum_{j=1}^{n} u_i v_j \vec{e_i} \bullet \vec{e}_j.$$

[1] These are known as the coordinates of $\vec{u}$ and $\vec{v}$ in the basis

Then since the basis is orthonormal, the only terms in the some that survive are when $i = j$, hence:

$$\vec{u} \bullet \vec{v} = \sum_{i=1}^{n} \sum_{j=1}^{n} u_i v_j \delta_{ij} = \sum_{i=1}^{n} u_i v_i.$$

Which is the standard scalar product in $\mathbb{R}^n$.

Remark 0.2.3. In essence, we have shown that from an algebraic point of view, there is only one Euclidean space, namely $\mathbb{R}^n$ with its standard dot product, which is why it is some times referred to as *the* Euclidean space.

The takeaway here should be that, from an algebraic point of view, there is no real difference between working with an arbitrary Euclidean space E or $\mathbb{R}^n$: it is exactly the same if we work in an orthonormal basis. Conceptually, however, it is interesting to allow E to be arbitrary. For instance, when working with $E = M_n(\mathbb{R})$, it is enlightening to keep representing matrices as matrices, instead of identifying them with column vectors in $\mathbb{R}^{n^2}$. This also has the advantage of not making it complicated to keep track of other algebraic structures that a space may have. For example, multiplication of matrices in $M_n(\mathbb{R})$.

0.3 The special case of 3 dimensional Euclidean space

0.3.1 Cross product and triple product

The case of 3 dimensions has been of interest for engineers and physicists, our perception of space is that there are 3 dimensions. Incidentally, it is well studied. Furthermore, there are some "accidents" in this low dimension that do not occur in higher dimensions, such as the possibility of defining the cross product.

In this section, E is an *oriented* (recall that this means that we have chosen a reference basis for the orientation) 3-dimensional Euclidean space. Then we can always choose a *positively* oriented orthonormal basis (p.o.n.b.): $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$.

It is then custom to define the cross product as the following determinant:

$$\vec{v} \times \vec{w} = \begin{vmatrix} v_1 & w_1 & \vec{e}_1 \\ v_2 & w_2 & \vec{e}_2 \\ v_3 & w_3 & \vec{e}_3 \end{vmatrix},$$

where:

$$\vec{v} = \sum_{i=1}^{3} v_i \vec{e}_i, \qquad \vec{w} = \sum_{i=1}^{3} w_i \vec{e}_i$$

Remark 0.3.1. You may have encountered this formula in transposed form, since $\det(A) = \det({}^tA)$ for any matrix A, you can use either definitions.

This definition is slightly unsatisfactory: it is not clear *a priori* that this definition does not depend on the choice of orthonormal basis. One of the goals of the remainder of this section is for us to check that it does **not**.

For this we introduce the:

Definition 0.3: Triple product

Let $\vec{u}, \vec{v}, \vec{w} \in E$, then the triple product of $\vec{u}, \vec{v}, \vec{w}$ is defined by:

$$[\vec{u}, \vec{v}, \vec{w}] = (\vec{u} \times \vec{v}) \bullet \vec{w}.$$

Remark 0.3.2. This can sometimes be called the *mixed* product as it combines both the scalar product with the cross product.

The triple product is an important mathematical object with a clear geometric interpretation we will uncover at the end of this section. For now, let us note that it has the following algebraic properties:

Proposition 0.2: Properties of the triple product

1. Viewed as map with three variables $E \times E \times E \to \mathbb{R}$ it is linear in the each variable:
$$\begin{cases} [\lambda\vec{u_1} + \vec{u_2}, \vec{v}, \vec{w}] = \lambda[\vec{u_1}, \vec{v}, \vec{w}] + [\vec{u}_2, \vec{v}, \vec{w}], \\ [\vec{u}, \lambda\vec{v}_1 + \vec{v}_2, \vec{w}] = \lambda[\vec{u}, \vec{v_1}, \vec{w}] + [\vec{u}, \vec{v}_2, \vec{w}], \\ [\vec{u_1} + \vec{u_2}, \vec{v}, \lambda\vec{w_1} + \vec{w_2}] = \lambda[\vec{u_1}, \vec{v}, \vec{w_1}] + [\vec{u}_2, \vec{v}, \vec{w_2}], \end{cases}$$
we say that it is trilinear.
2. It is *alternating*, which means that if any two of $\vec{u}, \vec{v}, \vec{w}$ are equal then:
$$[\vec{u}, \vec{v}, \vec{w}] = 0.$$

Observe these properties imply the following:

1. If $(\vec{u}, \vec{v}, \vec{w})$ is *a linearly dependent family* then $[\vec{u}, \vec{v}, \vec{w}] = 0$.

2. Swapping any two of the vectors $\vec{u}, \vec{v}, \vec{w}$, will flip the sign of the triple product, e.g.
$$[\vec{u}, \vec{v}, \vec{w}] = -[\vec{v}, \vec{u}, \vec{w}].$$

Indeed, observe that:

$$\begin{aligned} 0 = [\vec{u}+\vec{v},\vec{u}+\vec{v},\vec{w}] &= [\vec{u},\vec{u}+\vec{v},\vec{w}] + [\vec{v},\vec{u},\vec{w}] \\ &= \underbrace{[\vec{u},\vec{u},\vec{w}]}_{=0} + [\vec{u},\vec{v},\vec{w}] + [\vec{v},\vec{u},\vec{w}] + \underbrace{[\vec{v},\vec{v},\vec{w}]}_{=0}. \end{aligned}$$

Concretely, we can give the expression of $[\vec{u},\vec{v},\vec{w}]$ in terms of the coordinates of the vectors in the orthonormal basis $(\vec{e}_1,\vec{e}_2,\vec{e}_3)$.

Proposition 0.3: Coordinate expression

Assuming:

$$\vec{u} = \sum_{i=1}^{3} u_i \vec{e}_i, \quad \vec{v} = \sum_{i=1}^{3} v_i \vec{e}_i, \quad \vec{w} = \sum_{i=1}^{3} w_i \vec{e}_i$$

then:

$$[\vec{u},\vec{v},\vec{w}] = \begin{vmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{vmatrix}.$$

This means it is the determinant of the matrix whose columns are the coordinate column vectors $\mathbf{U},\mathbf{V},\mathbf{W}$ of the vectors $\vec{u},\vec{v},\vec{w}$ respectively in the basis $(\vec{e}_1,\vec{e}_2,\vec{e}_3)$.

We will now show that the triple product does not depend on the *positively oriented orthonormal basis* chosen to calculate it, using the properties of the determinant you are familiar with.

We will need the following fact about orthonormal bases:

Lemma 0.1

The change of basis matrix $\mathsf{P}_{\mathscr{B}\to\mathscr{B}'}$ between two *orthonormal* basis has determinant ± 1. In particular, the change of basis matrix between two positively oriented orthonormal bases is exactly 1.

Let $\mathscr{B} = (\vec{e}_1,\vec{e}_2,\vec{e}_3)$ and $\mathscr{B}' = (\vec{e}\,'_1,\vec{e}\,'_2,\vec{e}\,'_3)$ be two positively oriented orthonormal bases. To simplify notation we set:

$$\mathsf{P} \equiv \mathsf{P}_{\mathscr{B}\to\mathscr{B}'}.$$

Let $\vec{u},\vec{v},\vec{w}$ be 3 vectors and assume that $\mathbf{U},\mathbf{V},\mathbf{W}$ (resp. $\mathbf{U}',\mathbf{V}',\mathbf{W}'$) are the column matrices formed by their coordinates in the basis $\mathscr{B}$ (resp. $\mathscr{B}'$). Then these columns are related by:

$$\mathbf{U} = \mathsf{P}\mathbf{U}', \quad \mathbf{V} = \mathsf{P}\mathbf{V}', \quad \mathbf{W} = \mathsf{P}\mathbf{W}'.$$

Hence:

$$\begin{aligned}[\vec{u},\vec{v},\vec{w}] &= \det\begin{pmatrix}\mathbf{U} & \mathbf{V} & \mathbf{W}\end{pmatrix}\\ &= \det\begin{pmatrix}P\mathbf{U}' & P\mathbf{V}' & P\mathbf{W}'\end{pmatrix}\\ &= \det P\begin{pmatrix}\mathbf{U}' & \mathbf{V}' & \mathbf{W}'\end{pmatrix}\\ &= \underbrace{\det P}_{=1}\det\begin{pmatrix}\mathbf{U}' & \mathbf{V}' & \mathbf{W}'\end{pmatrix}\\ &= \det\begin{pmatrix}\mathbf{U}' & \mathbf{V}' & \mathbf{W}'\end{pmatrix}.\end{aligned}$$

The value of the triple product does not depend on the choice of *positively oriented orthonormal basis* chosen to calculate it.

Remark 0.3.3. As you may have noticed, this independence relies on a choice of orientation.

Since the cross product may be *defined* by the formula (in this case we define $[\vec{u},\vec{v},\vec{w}]$ by its coordinate expression in any positively oriented orthonormal basis.)

$$[\vec{u},\vec{v},\vec{w}] = (\vec{u}\times\vec{v})\bullet\vec{w}.$$

It follows that:

The value of the cross product does not depend on the choice of *positively oriented orthonormal basis* chosen to calculate it.

Conceptually, it is actually preferable to define the triple product first as a determinant and then define the cross product. You can try to show that you can recover the basic properties of the cross product from this definition using the properties of the determinant.

0.3.2 Some geometric properties of the cross and triple products

Whilst the cross product in many ways behaves like the usual notions of multiplication, there is one property it does not share: associativity, i.e. in general:

$$(\vec{u}\times\vec{v})\times\vec{w} \neq \vec{u}\times(\vec{v}\times\vec{w}).$$

Instead we have the following:

Proposition 0.4: Double cross product

Let $\vec{a},\vec{b},\vec{c}\in E$ then:

$$\vec{a}\times(\vec{b}\times\vec{c}) = (\vec{a}\bullet\vec{c})\vec{b} - (\vec{b}\bullet\vec{a})\vec{c}.$$

Remark 0.3.4. There is a mnemonic to remember this formula, in French it reads (with French pronunciation of course)

A, B C'est Assez (pronounced AC) Bien au BAC.

It can be justified as follows (this was skipped during the lectures):

- By definition of the cross product: $\vec{b} \times \vec{c}$ is orthogonal to the plane spanned by the vectors $(\vec{b}, \vec{c})$, but the vector $\vec{a} \times (\vec{b} \times \vec{c})$ is orthogonal to $\vec{b} \times \vec{c}$ so it must lie in the plane $(\vec{b}, \vec{c})$. Since it must also be orthogonal to $\vec{a}$ there is some non constant $k \in \mathbb{R}$ such that:

$$\vec{a} \times (\vec{b} \times \vec{c}) = k\left((\vec{a} \bullet \vec{c})\vec{b} - (\vec{b} \bullet \vec{a})\vec{c}\right).$$

- It is possible to show that *the constant* k *does not depend on the choice of vectors* $\vec{a}, \vec{b}, \vec{c}$! (Can you see why?) It then suffices to determine k for a specific choice of $\vec{a}, \vec{b}, \vec{c}$. If $(\vec{e}_1, \vec{e}_2, \vec{e_3})$ is a positive orthonormal basis then setting:

$$\vec{a} = \vec{e}_1, \vec{b} = \vec{e}_2, \vec{c} = \vec{e}_1$$

we find that:

$$k = 1.$$

We can use this formula to derive a link between $\vec{u} \times \vec{v}$ and $\vec{u} \bullet \vec{v}$, for this, recall that we define the length of a vector by: $\|\vec{u}\|^2 = \vec{u} \bullet \vec{u}$.

Now:

$$\begin{aligned} \|\vec{u} \times \vec{v}\|^2 &= (\vec{u} \times \vec{v}) \bullet (\vec{u} \times \vec{v}) \\ &= [\vec{u}, \vec{v}, \vec{u} \times \vec{v}] = [\vec{u} \times \vec{v}, \vec{u}, \vec{v}] \\ &= ((\vec{u} \times \vec{v}) \times \vec{u}) \bullet \vec{v}. \end{aligned}$$

Now using the formula above:

$$(\vec{u} \times \vec{v}) \times \vec{u} = -\vec{u} \times (\vec{u} \times \vec{v}) = -\left((\vec{u} \bullet \vec{v})\vec{u} - \|\vec{u}\|^2 \vec{v}\right),$$

thus, combining with the above we have:

$$\|\vec{u} \times \vec{v}\|^2 = \|\vec{u}\|^2 \|\vec{v}\|^2 - (\vec{u} \bullet \vec{v})^2. \tag{0.1}$$

Now the (geometric) angle $\theta \in [0, \pi]$ between two vectors $\vec{u}$, $\vec{v}$ in 3 space is defined by:

$$\vec{u} \bullet \vec{v} = \|\vec{u}\| \|\vec{v}\| \cos\theta,$$

it then immediately follows from the above that:

$$\|\vec{u} \times \vec{v}\| = \|u\| \|v\| \sin\theta.$$

Remark 0.3.5. Observe that since $\theta \in [0, \pi]$, $\sin\theta \geqslant 0$.

Now introduce the geometric angle φ between the vectors $\vec{u} \times \vec{v}$ and $\vec{w}$, then, by the fundamental relationship between the cross product and triple product:

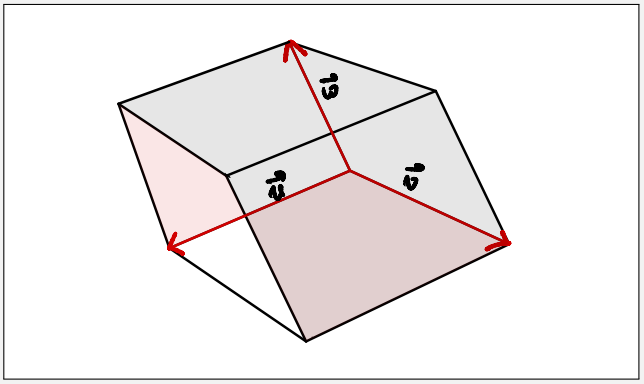
$$[\vec{u}, \vec{v}, \vec{w}] = (\vec{u} \times \vec{v}) \bullet \vec{w} = \|\vec{u} \times \vec{v}\| \|\vec{w}\| \cos\varphi = \|\vec{u}\| \|\vec{v}\| \|\vec{w}\| \sin\theta \cos\varphi.$$

A drawing (see optional reading) and some elementary geometric considerations will then convince you that:

Let $(\vec{u}, \vec{v}, \vec{w})$ be an ordered family of three vectors and consider $\mathcal{P}$ the parallelepiped supported by the vectors $(\vec{u}, \vec{v}, \vec{w})$ then:

$$[\vec{u}, \vec{v}, \vec{w}] = \begin{cases} \mathrm{Vol}(\mathcal{P}) & \text{if } (\vec{u}, \vec{v}, \vec{w}) \text{ is positively oriented,} \\ \mathrm{Vol}(\mathcal{P}) & \text{if } (\vec{u}, \vec{v}, \vec{w}) \text{ is negatively oriented.} \end{cases}$$

Here, Vol is the volume and it is understood that if $\mathcal{P}$ is degenerate, i.e. $(\vec{u}, \vec{v}, \vec{w})$ is linearly dependent, then the volume vanishes.



You should check that this is correct if $\vec{u}, \vec{v}, \vec{w}$ are orthogonal to one another.

Remark 0.3.6. This interpretation should be known to you !!! **This means its a potential exam question !!!!**

← End Lecture 2

0.4 Some complements (can be skipped)

In the above, we used the fact that $\mathscr{B} = (\vec{e}_1, \ldots, \vec{e}_n)$ and $\mathscr{B}' = (\vec{e}'_1, \ldots, \vec{e}'_n)$ are two *orthonormal* bases of a Euclidean space E, then the change of basis matrix

$$\mathsf{P} = \mathsf{P}_{\mathscr{B} \to \mathscr{B}'},$$

between them has determinant ± 1. This is because it most be an *orthogonal matrix*, i.e. it has the property that:

$${}^{t}\mathsf{P}\mathsf{P} = \mathrm{I}_n,$$

where I_n is the identity matrix.

Given vectors $\vec{u}, \vec{v} \in E$ we can decompose them on either orthonormal basis:

$$\vec{u} = \sum_{i=1}^{n} u_i \vec{e}_i = \sum_{i=1}^{n} u'_i \vec{e}'_i, \quad \vec{v} = \sum_{i=1}^{n} v_i \vec{e}_i = \sum_{i=1}^{n} v'_i \vec{e}'_i,$$

and according to the discussion at the end of Section 0.2, the following equalities hold:

$$\vec{u} \bullet \vec{v} = \sum_{i=1}^{n} u_i v_i = \sum_{i=1}^{n} u'_i v'_i.$$

Now, we have seen that the coordinates in the two bases are related by:

$$\mathbf{U} = P\mathbf{U}',$$

where $\mathbf{U} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$ and $\mathbf{U}' \begin{pmatrix} u'_1 \\ \vdots \\ u'_n \end{pmatrix}$. Writing this on components leads to:

$$u_i = \sum_{j=1}^{n} P_{ij} u_j.$$

Now we plugging this into the second equality above:

$$\sum_{i=1}^{n} \sum_{k=1}^{n} \sum_{l=1}^{n} P_{ik} u'_k P_{il} v'_l = \sum_{i=1}^{n} u'_i v'_i.$$

Rearranging the sums we conclude that:

$$\sum_{i=1}^{n} \sum_{k,l=1}^{n} \underbrace{\left(\sum_{i=1}^{n} P_{ik} P_{il} \right)}_{=(P^t P)_{kl}} u'_k v'_l = \sum_{i=1}^{n} u'_i v'_i.$$

This can be written in matrix form:

$${}^t\mathbf{U}' P^t P \mathbf{V}' = {}^t\mathbf{U}' \mathbf{V}'.$$

Since the vectors $\vec{u}$ and $\vec{v}$ where arbitrary it follows that this must hold for *any* column matrices $\mathbf{U}'$, $\mathbf{V}'$ and we conclude that:

$$P^t P = I_n.$$

Remark 0.4.1. An isometry of a Euclidean space E is a special kind of linear map $u : \mathsf{E} \to \mathsf{E}$ on an inner product space E that preserves the scalar product, i.e.:

$$\forall \vec{u}, \vec{v} \in \mathsf{E}, u(\vec{u}) \bullet u(\vec{v}) = \vec{u} \bullet \vec{v}.$$

They therefore are exactly the linear transformations that preserve angles and distances. Orthogonal matrices are *exactly* the matrices associated to isometries, expressed in orthonormal bases.

Chapter 1

Multivariable calculus

← Start Lecture 3

The examples of the introduction showed that in many of the physical situations we consider, after choosing a reference frame, we, in essence, want to study a function f mapping some $\mathbb{R}^n$ into $\mathbb{R}^m$, or even some abstract vector space V of dimension n into another such vector space V of dimension m. (In fact we won't specify the dimension unless we need to).

Adopting the latter point of view will enable us to emphasise on the essential role of the norm: $\|\vec{v}\|$ which, recall in $\mathbb{R}^n$ is usually defined by:

$$\|(x_1, \dots, x_n)\| = \left(\sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}},$$

and assigns a length to every vector in $\mathbb{R}^n$.

This plays a crucial role in how we define limits precisely, and how we can make sense of what it means for to vectors to be "close".

1.1 Basic topology of normed vector spaces

1.1.1 Norms and normed vector spaces

If it makes you more confortable, when first reading these notes you may mentally replace V and W par $\mathbb{R}^n$ and $\mathbb{R}^m$

The basic ingredients we will need in this section is a vector space V (you may take $V = \mathbb{R}^n$) and a way to assign a length to every vectors. This will be a function $V \to \mathbb{R}_+$ that we will denote by $\|\cdot\|$.

As usual, its definition is modelled on a selection of properties of the usual norm $\|\cdot\|$ in $\mathbb{R}^n$:

Definition 1.1: Norms

Let V be a vector space, a norm on V is a function $\|\cdot\| : V \to \mathbb{R}_+$ with the following properties:

1. $\vec{u} \in V, \|\vec{u}\| = 0 \Rightarrow \vec{u} = 0$ (separation),
2. If $\lambda \in \mathbb{R}, \vec{u} \in V$, $\|\lambda\vec{u}\| = |\lambda|\|\vec{u}\|$ (homogeneity),
3. If $\vec{u}, \vec{v} \in V$, $\|\vec{u} + \vec{v}\| \leqslant \|\vec{u}\| + \|\vec{v}\|$ (triangle inequality).

A *normed vector space* $(V, \|\cdot\|)$ is a vector space V on which we have specified a norm $\|\cdot\|$.

See Section 1.5 (p74) in Marsden and Tromba for the case of $\mathbb{R}^n$ with its standard norm, which is the fundamental example. Observe that on $\mathbb{R}$ the standard norm is just the absolute value function $|\cdot|$.

Example 1.1. A less obvious, but important, example is the following. Consider $V = M_n(\mathbb{R})$, and define a norm N on V by:

$$N(A) = \max_{X \in \mathbb{R}^n, X \neq 0} \frac{\|AX\|}{\|X\|}.$$

In this formula :

- AX is the result of the matrix multiplication of A with a non-vanishing column vector $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, AX is therefore a column vector in $\mathbb{R}^n$.
- $\|X\| = \left(\sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}}$ is the standard norm on $\mathbb{R}^n$.

We are therefore assigning the length of a matrix to be the maximum value $\|AX\|$ attains when $\|X\| = 1$.

Let us check it satisfies all the properties:

- Let $A \in M_n(\mathbb{R})$ and assume $N(A) = 0$. By definition, this means that the maximum is vanishing or in other words that for all $X \in \mathbb{R}^n$,

  $$\frac{\|AX\|}{\|X\|} = 0 \Rightarrow \|AX\| = 0,$$

  then by the properties of the usual norm on $\mathbb{R}^n$,

  $$\forall X \in \mathbb{R}^n, AX = 0.$$

  A is therefore the 0 matrix.

- If $\lambda \in \mathbb{R}, A \in M_n(\mathbb{R})$,

$$N(\lambda A) = \max_{X \in \mathbb{R}^n, X \neq 0} \frac{\lambda \|AX\|}{\|X\|} = \max_{X \in \mathbb{R}^n, X \neq 0} |\lambda| \frac{\|AX\|}{\|X\|} = |\lambda| \max_{X \in \mathbb{R}^n, X \neq 0} \frac{\|AX\|}{\|X\|}.$$

- By the triangle inequality in $\mathbb{R}^n$, if $A, B \in M_n(\mathbb{R})$ and $X \in \mathbb{R}^n \setminus \{0\}$, then:

$$\|(A+B)X\| = \|AX + BX\| \leqslant \|AX\| + \|BX\|,$$

and so dividing by $\|X\|$ we arrive at:

$$\frac{\|(A+B)X\|}{\|X\|} \leqslant \frac{\|AX\|}{\|X\|} + \frac{\|BX\|}{\|X\|},$$

for every $X \in \mathbb{R}^n \setminus \{0\}$, furthermore for all $X \in \mathbb{R}^n \setminus \{0\}$:

$$\frac{\|AX\|}{\|X\|} \leqslant \max_{X \in \mathbb{R}^n, X \neq 0} \frac{\|AX\|}{\|X\|}$$

and similarly for the other term. Consequently, for any $X \in \mathbb{R}^n \setminus \{0\}$:

$$\frac{\|(A+B)X\|}{\|X\|} \leqslant N(A) + N(B),$$

but this is certainly true for the X at which this maximum is attained so:

$$N(A+B) \leqslant N(A) + N(B),$$

as desired.

This norm is especially interesting because it is "compatible" with matrix multiplication in the sense that it has the property:

$$N(AB) \leqslant N(A)N(B).$$

Which you can try to show using the definition of N. ◇

1.1.2 Open balls and open sets

In many ways, $\|\cdot\|$ plays a similar role for a vector space to that which the absolute value $|\cdot|$ plays on real numbers, and we can imitate the definitions from 1-variable calculus in this multidimensional setting. Accordingly, if $\vec{v}$ and $\vec{w}$ are two vectors in a normed vector space $(V, \|\cdot\|)$, then the distance between is defined to be:

$$\|\vec{v} - \vec{w}\|.$$

Starting from this we will be able to define the notion of limits and continuous functions. However, the topology of a line is considerably simpler than in a higher dimensional case, (we basically only have one direction in which to explore). To help us

with our understanding of theses space in higher dimensions it is useful to introduce the following visual language.

Definition 1.2: Open ball

Let $(V, \|\cdot\|)$ be a normed vector space, $\vec{u} \in V$ and $r > 0$. The **open ball** of centre $\vec{u}$ and radius $r > 0$ is the set:

$$B(\vec{u}, r) = \{\vec{v} \in V, \|\vec{v} - \vec{u}\| < r\}.$$

Example 1.2. $V = \mathbb{R}^2$, $\|(x, y)\| = \sqrt{x^2 + y^2}$.

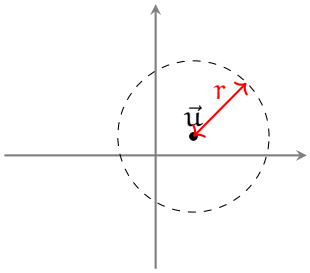


Figure 1.1: An open ball $B(\vec{u}, r)$ in the Euclidean space $\mathbb{R}^2$ is the "interior" of disk. The points of the boundary circle are **not** in the open ball.

Remark 1.1.1. In the practice exercises, you will see some other possible norms for $\mathbb{R}^2$.

◇

The intuitive idea of an open ball is that it enables us to explore locally vector space in all directions around a point, they are the basis for more general sets of this type.

Definition 1.3: Open sets

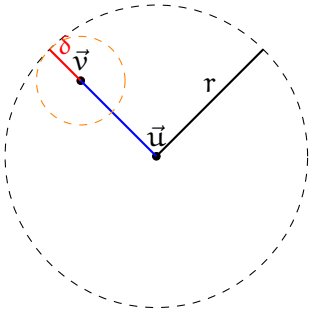
A subset $U \subset V$ of a normed vector space $(V, \|\cdot\|)$ is said to be *open* if:

$$\forall \vec{u} \in U, \quad \exists \delta > 0, \quad B(\vec{u}, \delta) \subset U.$$

i.e. For every point $\vec{u}$ in the set U, there exists a "small" open ball centered at $\vec{u}$ and entirely contained in the set U.

Open sets are sets in which one can "zoom in" around any point with a bit of room.

Example 1.3. Open balls are open. (phew!) The idea of the proof is sketched in the following diagram. Intuitively we just need to choose the size of the ball around $\vec{v}$ to



be smaller than the distance to the boundary.

The above diagram suggests that we could choose the radius of the ball to be $r-\|\vec{u}-\vec{v}\|$ but to be safer we shall instead choose:

$$\delta = \frac{1}{2}(r - \|\vec{u} - \vec{v}\|).$$

We just need to find *a* ball that works, not necessarily the best ball.

Let us show formally that this choice works. Let $\vec{w} \in B(\vec{v}, \delta)$ then we must show that $\vec{w} \in B(\vec{u}, r)$ or in other words:

$$\|\vec{w} - \vec{u}\| < r.$$

We estimate this quantity using the *triangle inequality*:

$$\begin{aligned}\|\vec{w} - \vec{u}\| = \|\vec{w} - \vec{v} + \vec{v} - \vec{u}\| &\leqslant \underbrace{\|\vec{w} - \vec{v}\|}_{<\delta} + \underbrace{\|\vec{v} - \vec{u}\|}, \\ &\leqslant \frac{1}{2}(r - \|\vec{v} - \vec{u}\|) + \|\vec{v} - \vec{u}\| = \frac{1}{2}(r + \underbrace{\|\vec{v} - \vec{u}\|}_{<r}) \\ &< r.\end{aligned}$$

Which concludes the proof, note that since we only used the triangle inequality, this proves it for *any* open ball in *any* normed vector space and not just the Euclidean disc. This is why it was "better" to include the factor $\frac{1}{2}$ in the choice of δ, our estimations with the triangle inequality do not use any particular feature of the Euclidean ball.

◇

1.2 Limits of functions

1.2.1 The definition

We are now ready to formalise the notion of limit of a function. In this section, $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ are two normed vector spaces and $A \subset V$, we will consider a function:

$$f : A \subset V \to W.$$

To simplify notation, we will now abandon the arrow notation for elements of V.

We want to make precise the notion that $f(a), a \in A$ approaches a value $l \in W$ when x_0 approaches some $x_0 \in V$ but "close" to A. This will be achieved using open balls:

Definition 1.4: Limits

We write that $\lim_{a \to x_0} f(a) = l$ if when given an arbitrary open ball $B_W(l, \varepsilon) \subset W$, one can always find an open ball $B_V(x_0, \delta) \subset V$ that is mapped entirely into $B_W(l, \varepsilon)$ by the function f. Formally:

$$\forall \varepsilon > 0, \exists \delta > 0, f(B_V(x_0, \delta) \cap A) \subset B_W(l, \varepsilon).$$

In some sense this definition means I can approximate l with arbitrary precision by all $f(a)$ as long as a lies in some small enough ball $B(x_0, \delta)$. Another way of saying this is that f should send points of A as close as we want to the limit l, provided that these points lie in a small enough ball around x_0.

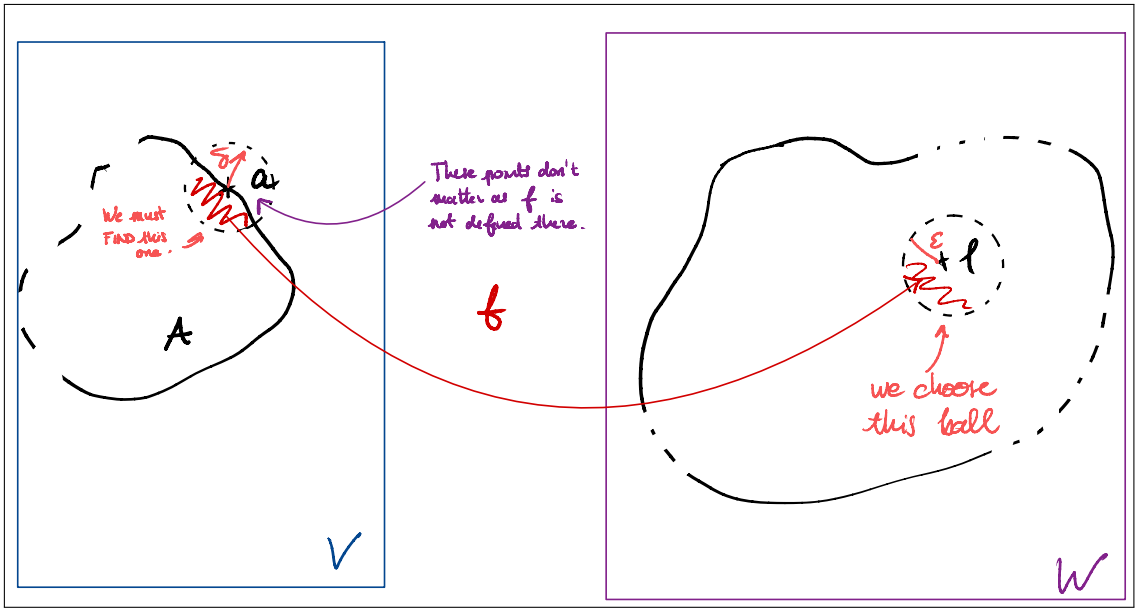


Figure 1.2: Diagram illustrating the intuitive idea of the definition of limits.

End Lecture 3→
Start Lecture 4→

Observe that, the limit cannot exist if the intersection $B(x_0, \delta) \cap A$ is empty for some $\delta > 0$. This is intuively reasonable as such a point cannot be approached from A. See, for instance the situation in Figure 1.3.

Hence, the limit can only make sense if this never occurs, the following formal definition captures this idea:

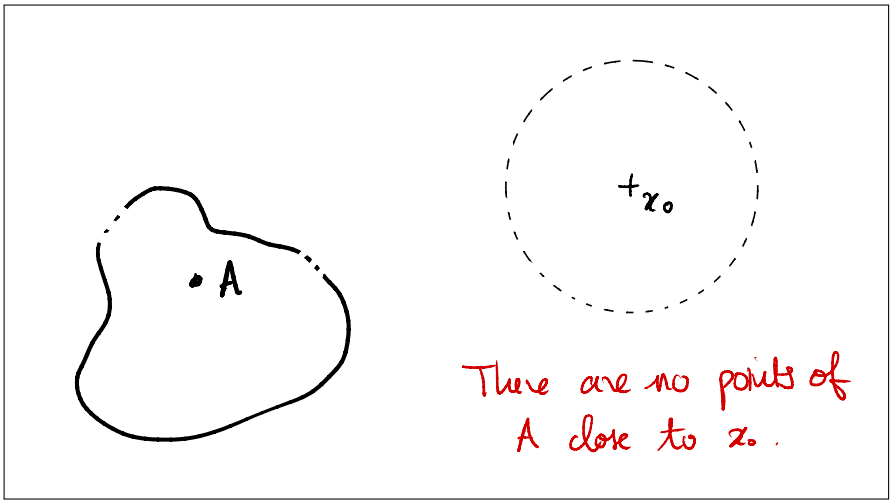


Figure 1.3: This point cannot be "reached" from the set A.

Definition 1.5: Adherent points

Let $A \subset V$ be an arbitrary subset of a normed vector space $(V, \|\cdot\|)$ and a point $x_0 \in V$ (*not necessarily in* A), then x_0 is said to be **adherent** to A if and only if, every open ball $B(x_0, r)$ meets the set A in at least one point, formally:

$$\forall r > 0, B(x_0, r) \cap A \neq \emptyset.$$

Example 1.4.

- Points $a \in A$ are clearly adherent to A.
- Boundary points, defined below.

◇

Intuitively, a point will be "on the boundary" if it is simultanenously "close" enough to A but also to exterior of A $V \setminus A$, in other words if it is adherent to both A and to $V \setminus A$.

Definition 1.6: Boundary points

Let A be a subset of a normed vector space $(V, \|\cdot\|)$ and $x_0 \in V$ (once more *not necessarily* a point A). x_0 is a **boundary point** if it is adherent to both A and $V \setminus A$, in other words, if every open ball $B(x_0, r)$ contains a point of A and a point that is *not* in A.

Remark 1.2.1. It is important to note that boundary sets do not need to be points of A! For instance, open sets, by definition, cannot contain any of their boundary points. However, it is clear that the open ball has a boundary...

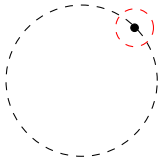


Figure 1.4: A boundary point on the open unit ball

1.2.2 Simple example of limits one can show using the definition

1. The obvious limit: $\lim_{x \to x_0} x = x_0$ is true, where we interpret x as the identity map $f : V \to V$, $f(x) = x$.
2. The triangle inequality can be used to show that the norm as a function $\|\cdot\| : (V, \|\cdot\|) \to (\mathbb{R}, |\cdot|)$ satisfies the limit:
$$\lim_{x \to x_0} \|x\| = \|x_0\|.$$
3. The limit coincides with the standard one in the one dimensional case, $V = W = \mathbb{R}$ and $\|\cdot\|_V = \|\cdot\|_W = |\cdot|$.

1.2.3 Properties of limits

The usual properties of limits carry over without modification to the new case, we refer the reader to Marsden & Tromba for a complete list (p115 Thm 3). Most of these properties can be obtained as a special case of the following one that we will investigate in more detail:

Theorem 1.1: Composition of limits

Let $(V_1, \|\cdot\|_{V_1})$, $(V_2, \|\cdot\|_{V_2})$, $(V_3, \|\cdot\|_{V_3})$ be three normed vector spaces and two functions:
$$f : A \subset V_1 \to V_2, \qquad g : B \subset V_2 \to V_3.$$
Assume that:
- $f(A) \subset B$ (so that the composition makes sense)
- x_0 is adherent to A and $\lim_{a \to x_0} f(a) = l \in V_2$,
- $\lim_{b \to l} g(b) = L \in V_3$.

Then:
$$\lim_{a \to x_0} (g \circ f)(a) = g(f(a)) = L.$$

Proof. Let $\varepsilon > 0$, then using $\lim_{b \to l} g(b) = L \in V_3$, there is $\delta_1 > 0$ such that
$$g(B_{V_2}(l, \delta_1) \cap B) \subset B_{V_3}(L, \varepsilon).$$

Moreover, using: $\lim_{a \to x_0} f(a) = l \in V_2$, there is $\delta_2 > 0$ such that:

$$f(B_{V_1}(x_0, \delta_2) \cap A) \subset B_{V_2}(l, \delta_1).$$

Since $f(A) \subset B$ it follows that $f(B_{V_1}(x_0), \delta_2) \subset B_{V_2}(l, \delta_1) \cap B$. Putting everything together we see then that:

$$g(f(B_{V_1}(x_0), \delta_2)) \subset g(B_{V_2}(l, \delta_1) \cap B) \subset B_{V_3}(L, \varepsilon),$$

which concludes the proof. □

The main idea of the proof is depicted in Figure 1.5:

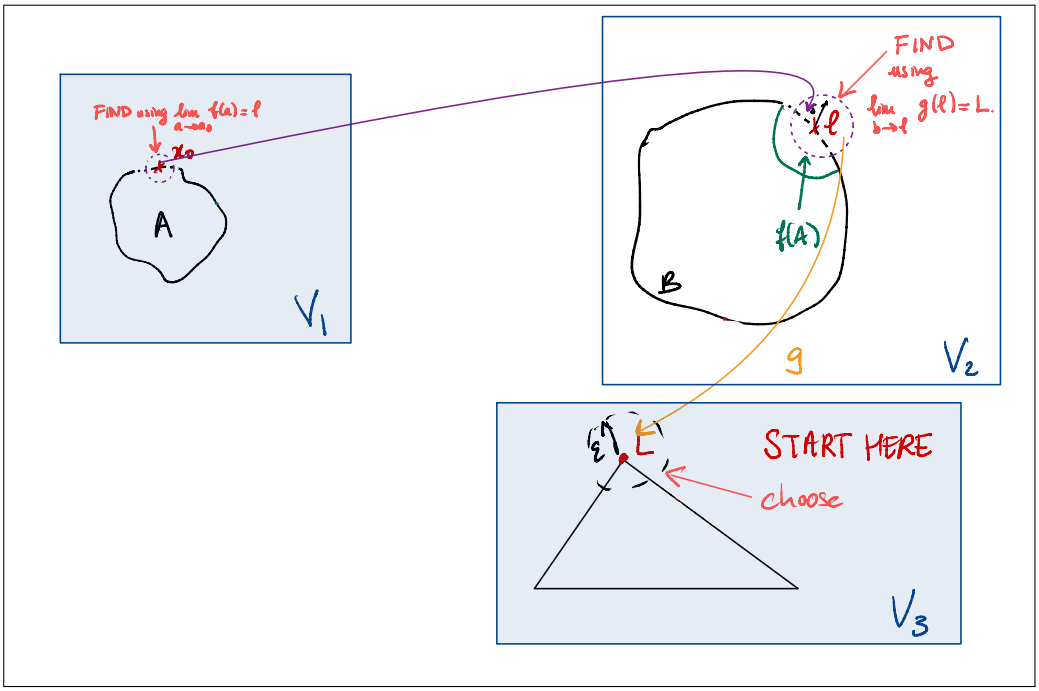


Figure 1.5: Composition of limits

1.2.4 For your information (can be skipped, not mentioned in class)

There is a slight subtlety with the notion of limit when $a \to x_0 \in A$. One can observe that in this case, according to our definition either:

$$\lim_{a \to x_0} f(a) = f(x_0) \quad \text{or} \quad \text{The limit does not exist.}$$

This can be shown as follows, suppose that $\lim_{x \to x_0} f(x) = l$, then taking $\varepsilon = \frac{1}{n}$ for $n \geqslant 1$, then one can find $\delta_n > 0$ such that:

$$f(B_V(x_0, \delta_n)) \subset B(l, \frac{1}{n}).$$

Since $x_0 \in B_V(x_0, \delta_n)$ for all $n \geqslant 1$, we conclude that *for all* $n \geqslant 1$:

$$\|f(x_0) - l\|_V < \frac{1}{n},$$

but taking the limit $n \to \infty$ we conclude that:

$$0 \leqslant \|f(x_0) - l\|_V \leqslant 0$$

so that:

$$\|f(x_0) - l\|_V = 0 \Rightarrow f(x_0) = l.$$

This is *not* a problem for most applications because we *want* $\lim_{x \to x_0} f(x) = f(x_0)$.

Some authors (like Marsden & Tromba) choose to systematically *exclude* the point x_0 in the definition of the limit. This reason for this is because there are cases in which the value of $f(x_0)$ really pollutes our understanding of the behaviour of f near x_0 even though f actually has a clear limiting behaviour.

Consider as an example a function f defined as follows:

$$\begin{array}{rccl} f : \mathbb{R} & \longrightarrow & \mathbb{R} & \\ x & \longmapsto & \begin{cases} x^2 & \text{if } x \neq 2 \\ 25 & \text{if } x = 2 \end{cases} \end{array}$$

In this example, f clearly *behaves* like $x \mapsto x^2$ near the point $x = 2$, despite the fact the function assumes a weird value there. It is clear that there is some sense in which the limit should be 4, but it will not exist according to our definition.

However, there is a price to pay for choosing to systematically exclude the point x_0 when computing the limit: the limit composition theorem as stated above is *false* with the alternative definition, although only a small modification makes it correct again: g needs to be assumed continuous.

It is a question of taste to which definition one might use, we just have to be consistent. For us, during this course, we will not encounter places where the difference actually matters. Furthermore, the alternative definition chosen by like Marsden & Tromba is the same as our definition of we restrict the domain of f to $A \setminus \{x_0\}$, this is sometimes written:

$$\lim_{\substack{x \to x_0 \\ x \neq x_0}} f(x).$$

For the above example, this limit will exist, and will be equal to 4 reconciling with the intuition.

Remark 1.2.2. More generally, we can define the notion of limit *in the direction* of some subset $B \subset A$, this would be written:

$$\lim_{x \in B} f(x),$$

and is defined to be the limit of the restriction $f|_B$ of the function f to the subset B. The above case corresponds to $B = A \setminus \{x_0\}$.

1.3 Continuous functions

We will now generalise the notion of continuous functions to the setting of normed vector spaces.

Throughout this section, $(V, \|\cdot\|_V)$ *and* $(W, \|\cdot\|_W)$ *are two normed vector spaces.*

The standard intuitive description at a precalculus level is that a function is continuous if we can draw it without lifting the pen off the paper. Of course, this intuition will not carry over to higher dimensions. Continuous functions play a role in topology that is similar to the role that linear transformations play in linear algebra: they enable to compare topologies of different spaces. We will not develop this point of view further in this course and will simply state the following practical local definition:

Definition 1.7: Continuous functions

Let $A \subset V$, $f : A \longrightarrow W$ be a function, $x_0 \in A$:

- f is said to be *continuous* at x_0 if:

$$\lim_{x \to x_0} f(x) = f(x_0).$$

- f is said to be *continuous* (on A) if f is continuous at every point of A.

This definition generalises the usual one (so all of the functions you know to be continuous are still continuous) and will be used to construct new ones in higher dimensions. For us, continuity will have two main applications. First, determining limits: if a function is continuous at x_0 to find the limit at x_0 one simply needs to evaluate it to find the limit. This is only useful, however, if we have ways to show that functions are continuous without applying the definition of the limit. This is the role of the usual rules of continuous functions which will carry over without modification: algebraic operations (when they make sense) are continuous and we will find that sums, products... of continuous functions will be continuous.

In order to gain some experience with the definition of the limit and how it can be used we will explore how to deduce some of these statements from the following immediate consequence of the composition of limits theorem:

Theorem 1.2

Compositions of continuous functions are continuous.

Let us now show explicitly that:

In a normed vector space the algebraic operations in V are continuous when considered as maps:

$$\begin{array}{rcccc} +: V\times V & \to & V & \qquad \cdot : \mathbb{R}\times V & \to & V \\ (v_1,v_2) & \mapsto & v_1+v_2 \ , & (\lambda, v) & \mapsto & \lambda\cdot v \end{array}$$

and the spaces $V\times V$ are equipped with the norms:

$$\begin{cases} \|(v_1,v_2)\|_{V\times V} = \|v_1\|_V + \|v_2\|_V, \\ \|(\lambda,v)\|_{\mathbb{R}\times V} = |\lambda| + \|v\|_V. \end{cases}$$

The keys to this fact are in a nutshell the triangle inequality and the homogeneity property of norms.

Let us begin with addition: we must choose a *fixed* but arbitrary pair $(v_1,v_2)\in V\times V$ and show that:

$$\lim_{(x,y)\to(v_1,v_2)} x+y = v_1+v_2.$$

For this we fix $\varepsilon > 0$ and we need to show that we can find $\delta > 0$ such that if $(x,y)\in B_{V\times V}((v_1,v_2),\delta)$ then: $x+y\in B_V(v_1+v_2,\varepsilon)$.

To find an appropriate $\delta>0$ we begin by trying to estimate $\|x+y-(v_1+v_2)\|_V$ with $\|(x,y)-(v_1,v_2)\|_{V\times V}$, now by the triangle inequality:

$$\|x+y-(v_1+v_2)\|_V \leqslant \|x-v_1\|_V+\|y-v_2\|_V = \|(x-v_1,y-v_2)\|_{V\times V} = \|(x,y)-(v_1,v_2)\|_{V\times V}.$$

Hence, to make: $\|x+y-(v_1+v_2)\|_V$ smaller than ε it suffices to make $\|(x,y)-(v_1,v_2)\|_{V\times V}$ smaller than $\delta=\varepsilon$. This shows that $\lim_{(x,y)\to(v_1,v_2)} x+y = v_1+v_2$.

The scalar multiplication is very similar. Let us fix $(\lambda_0,v_0)\in\mathbb{R}\times V$ and show that:

$$\lim_{(\lambda,v)\to(\lambda_0,v_0)} \lambda\cdot v = \lambda_0\cdot v_0.$$

Once more let us fix $\varepsilon>0$ and try to estimate $\|\lambda v-\lambda_0 v_0\|_V$ in terms of $\|(\lambda,v)-(\lambda,v_0)\|_{\mathbb{R}\times V}$. The main tool as always for these things is the triangle inequality:

$$\begin{aligned} \|\lambda v-\lambda_0 v_0\|_V &= \|\lambda v-\lambda_0 v+\lambda_0 v-\lambda_0 v_0\|_V \\ &\leqslant |\lambda-\lambda_0|\,\|v\|_V + |\lambda_0|\|v-v_0\|_V. \end{aligned}$$

For the last inequality we have also used the homogeneity property.

Now we are almost done but we must estimate $\|v\|_V$: the choice δ cannot depend on it: it is an arbitrary element in the ball of radius δ. For instance, observe that we can always choose $\delta<1$. (There may be a more possibly optimal larger value

of δ that works, but we only need one that works not the best.) Then, one can see by definition of the norm that if $(\lambda, v) \in B_{\mathbb{R}\times V}((\lambda_0, v), \delta)$ then:

$$\|v - v_0\|_V \leqslant \|v - v_0\|_V + |\lambda - \lambda_0| < \delta < 1.$$

Using the triangle inequality, we have:

$$\|v\|_V \leqslant \|v - v_0\|_V + \|v_0\|_V < 1 + \|v_0\|_V,$$

so in this case:

$$\|\lambda v - \lambda_0 v_0\|_V \leqslant |\lambda - \lambda_0|(1 + \|v_0\|_V) + |\lambda_0|\|v - v_0\|_V.$$

Now, if we set $\delta = \min(1, \frac{\varepsilon}{2|\lambda_0|}, \frac{\varepsilon}{2(1+\|v_0\|_V)})$ then if $(\lambda, v) \in B_{\mathbb{R}\times V}((\lambda_0, v), \delta)$ we have:

$$\|\lambda v - \ \lambda_0 v_0\|_V < \varepsilon.$$

Which proves the desired limit and continuity of scalar multiplication with vectors.

Remark 1.3.1. In the special case where $V = \mathbb{R}$, we have just established that **multiplication of real numbers is continuous.**

← End Lecture 4
← Start Lecture 5

1.3.1 Product spaces

The spaces $V \times V = \{(v_1, v_2), v_1 \in V, v_2 \ \in V\}$ and $\mathbb{R} \times V$, as well as $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$, that arose above are examples of product spaces. As we did above, for any normed vector spaces it is possible to make $V \times W$ a normed vector space with the norm:

$$\|(v, w)\|_{V\times W} = \|v\|_V + \|w\|_W.$$

The following basic facts about product spaces are used all (often implicitly) on numerous occasions. A product space is equipped with two natural projection maps that project onto each of its factors:

$$\begin{array}{rcl} \pi_V : V \times W & \to & V \\ (v, w) & \mapsto & v \end{array}, \qquad \begin{array}{rcl} \pi_W : V \times W & \to & W \\ (v, w) & \mapsto & w. \end{array}$$

It is a straightforward application of the definition to show that:

Proposition 1.1

The projection maps π_V and π_W are continuous.

The fundamental theorem about continuous maps in relation to product spaces is the following characterisation:

Theorem 1.3: Continuity of functions with values in a product space

Let $(Z, \|\cdot\|_Z)$ be a third normed vector space, then a function:

$$\begin{array}{rcl} f : A \subset Z & \to & V \times W \\ z & \mapsto & (h(z), g(z)) \end{array} .$$

Then f is continuous if and only if the maps $h = \Pi_V \circ f$ and $g = \Pi_W \circ f$ are continuous.

We leave the proof as an exercise to the interested reader.

Remark 1.3.2. We can generalise everything by induction to a finite number of factors $V_1 \times V_2 \times \cdots \times V_n$ without any modification.

Example 1.5. In $\mathbb{R}^n$, the coordinate functions $(x_1, \ldots, x_n) \in \mathbb{R}^n \mapsto x_i \in \mathbb{R}$ are continuous. This, for instance, justifies the statements about functions that are polynomials in the coordinates $(x_1, \ldots, x_n)$ are continuous. ◇

Using this we can justify theoretically the well known fact:

Proposition 1.2

Let $g : A \subset \mathbb{R}^n \to \mathbb{R}^m$ be defined by: $g(a) = (g_1(a), \ldots, g_m(a)), a \in A$ then g is continuous iff $g_1, \ldots, g_m$ are continuous.

Armed with these facts, the reader should now be able to justify, by composition of continuous functions Theorems 3 and 4 in Marsden & Tromba.

1.3.2 Some particularities of finite dimensional spaces

The theory we have developed above, save when explicitly mentioned in the examples, does not require V or W to be finite dimensional. Finite dimensional spaces are however special: we get some extra things for free.

First, the attentive reader may have observed that the definitions of limits and continuous functions *depend* on the norm you use on either space, and may be concerned that I tricked you, as the norm for instance used on $\mathbb{R}^2$, when viewed above as the product space $\mathbb{R} \times \mathbb{R}$ is *not* the usual Euclidean norm.

Luckily for me, there is a deep (magic) theorem:

Theorem 1.4: Equivalence of norms in finite dimensions

On a finite dimensional space all norms are *equivalent*.
More precisely, given any two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on a *finite dimensional* vector space V, there are constants $m, M > 0$ such that:

$$\forall v \in V, \quad m\|v\|_1 \leqslant \|v\|_2 < M\|v\|_1.$$

What this means in practice is that the notions of limit, continuous functions and open sets we have introduced **do not depend on the choice of norms** in finite dimensions. By this I mean, for instance, that if $\lim_{x \to x_0} f(x) = l$ for some norm then it is true for any other norm.

In conclusion, *in finite dimensions*, we can use any norm we want, and usually one that makes whatever we want to show as simple as possible.

Remark 1.3.3. For product spaces $V \times W$, even in infinite dimensions, the norms:

$$\|(v,w)\|_1 = \|v\|_V + \|w\|_W, \quad , \|(v,w)\|_2 = \sqrt{\|v\|_V^2 + \|w\|_W^2},$$

are always equivalent.

Finite dimensional have an important subclass of continuous functions:

Theorem 1.5

Let $u : V \to W$ be a *linear transformation* between a finite dimensional normed space V and an arbitrary normed space W. We recall that a map is linear u when it has the property:

$$u(\lambda v_1 + v_2) = \lambda u(v_1) + u(v_2).$$

Then u is continuous, furthermore there is a constant $M > 0$ such that:

$$\|u(x)\|_W \leqslant M\|x\|_V.$$

This theorem is proved in a Practice exercise.

The takeaway is, that in *finite dimension* linear maps are continuous, so for instance:

Example 1.6. Let $V = M_n(\mathbb{R})$, $W = M_n(\mathbb{R})$ (resp. $W = \mathbb{R}$) then the maps $A \mapsto {}^tA$ and $A \mapsto \mathrm{tr}(A)$ are linear and therefore continuous since V has finite dimension. $\diamond$

1.3.3 Continuous functions and open sets

We briefly mention the following fact (for a proof see the exercises) about continuity and open sets.

Theorem 1.6

Let $f : V \to W$ be a map between normed vector spaces, then f is continuous if and only if for every open set $U \subset W$, the preimage:

$$f^{-1}(U) = \{x \in V, f(x) \in U\},$$

is open.

Proof. See Practice Exercises. □

This conveys the deeper conceptual meaning of continuity: open sets in W are mapped into from open sets in V.

For us, it will be useful for recognising that some sets are open.

Example 1.7.

- $\mathbb{R}^n \setminus \{0\}$ is open since it is the preimage of the union of open intervals $(-\infty, 0) \cup (0, +\infty)$ under the continuous map: $x \mapsto \|x\|$.
- The set of invertible matrices, denoted by $GL_n(\mathbb{R})$, is open in $M_n(\mathbb{R})$, as the preimage of $(-\infty, 0) \cup (0, +\infty)$ under the map: $A \in M_n(\mathbb{R}) \mapsto \det A \in \mathbb{R}$ (which is continuous as it is a polynomial in the coefficients of A).

◇

1.4 Optional reading: Banach spaces

Save when we discussed equivalence of norms and the continuity of *all* linear maps, there has never been any need for us to assume that we are working with finite dimensional vector spaces. Normed vector spaces of infinite dimension have numerous applications, notably in the study of (partial) differential equations. In these cases we considered normed vector spaces of functions, for instance, the set of all continuous real-valued functions on $[0, 1]$, $\mathcal{C}^0([0, 1], \mathbb{R})$, can be considered a normed vector space with the norm:

$$\|f\|_\infty = \max_{x \in [0,1]} |f(x)|.$$

Whilst up to now, switching to infinite dimensional vector spaces seems pretty harmless, for some of the later theorems to carry over, one must make a further topological assumptions that is automatically satisfied in the finite dimensional case. Indeed, sometimes in mathematics we need to prove the existence of an object, solution to a given problem, but can only do so by an iterative limiting procedure: we construct a sequence that will "converge" to the solution.

The problem is according to the definition of the limit, to prove that something converges to a limit you must find the limit in advance. This is where the notion of *completeness* steps in. One way of expressing the idea is that if a sequence of points has the property that as n increases the terms of the sequence are getting closer and closer to one another, then it seems reasonable to assume that there should be a limit, a normed vector space is said to be complete if this is the case. To formalise this we use the notion of *Cauchy sequence*.

Definition 1.8

Let $(V, \|\cdot\|)$ be a normed vector space and $(v_n)_{n\in\mathbb{N}}$ a sequence of points of V. Then $(v_n)_{n\in\mathbb{N}}$ is said to be a Cauchy sequence if for any open ball $B(0,\varepsilon)$, one can find an integer $N \in \mathbb{N}$, such that for all $m, n \geqslant N$:

$$v_n - v_m \in B(0,\varepsilon).$$

Cauchy sequences are precisely the sequences who terms are arbitrarily close to one another for N large enough.

Example 1.8.

If a sequence $(v_n)_{n\in\mathbb{N}}$ converges to l, then $(v_n)_{n\in\mathbb{N}}$ is a Cauchy sequence. ◇

Definition 1.9

A normed vector space $(V, \|\cdot\|)$ is a said to be **complete** if every Cauchy sequence converges, in this case, V is said to be a **Banach space**.

Example 1.9.

- Finite dimensional normed vector spaces are Banach spaces.
- If $(E_1, \|\cdot\|_{E_1})$ and $(E_2, \|\cdot\|_{E_2})$ are Banach spaces then $(E_1 \times E_2, \|\cdot\|_{E_1\times E_2})$ is a Banach space.

◇

1.5 Some other topological notions (not covered in Winter 2025)

Open sets are not the only sets of interest when studying the topology of a given space. Other useful topological notions have also been identified, often in connection with generalisations of standard theorems from analysis of functions with one real variable.

The simplest remarkable sets are complementary subsets to the open sets, these are called closed:

Definition 1.10: Closed sets

F is closed iff $X \setminus F$ is open.

Whereas open sets contain none of their boundary points, closed sets contain all of them ! In fact, closed sets are characterised by the fact they contain all adherent points.

Another example is that of connected sets: the intermediate value theorem for continuous functions, i.e. if $f : [a,b] \subset \mathbb{R} \to \mathbb{R}$ is a continuous function and $y \in [f(a), f(b)]$, then there is $c \in [a,b], y = f(c)$, relies on the fact that the interval $[a,b]$ is "in one piece": we say that it is *connected*. This theorem can be generalised to the statement: *the image under a continuous function of a connected set is connected.*

Definition 1.11: Connected subsets

Let $(E, \|\cdot\|_E)$ be a normed vector space and A a subset then A is **connected** if and only if it is *not* the disjoint union of two sets of the form $U \cap A \neq \emptyset$, $W \cap A \neq \emptyset$, where U and W are open sets of E. In other words, if $(U \cap A) \cap (W \cap A) = \emptyset$:

$$A = (U \cap A) \cup (W \cap A) \Rightarrow U \cap A = \emptyset \text{ or } W \cap A = \emptyset.$$

This is the correct topological property for the intermediate value theorem to hold.

Proof. Let us assume that $f(A) = (U \cap f(A)) \cup (W \cap f(A))$ with $U \cap f(A) \cap W \cap f(A) = \emptyset$, then:

$$A = f^{-1}(U) \cap A \cup f^{-1}(W) \cap A,$$

and $f^{-1}(U)$ and $f^{-1}(W)$ are open in E by continuity. Furthermore, $f^{-1}(U) \cap A$ and $f^{-1}(W) \cap A$ must be disjoint by assumption, therefore since A is connected one of $f^{-1}(U) \cap A$ or $f^{-1}(W) \cap A$ is empty, meaning that one of $U \cap f(A)$ or $W \cap f(A)$ is empty. □

Remark 1.5.1. Observe that connectedness is an intrinsic property.

Example 1.10. Intervals are precisely the connected subsets of $\mathbb{R}$. ◇

1.6 Differentiation

1.6.1 Definition and examples

Linear transformations are the first types of functions we encounter when we begin studying linear algebra. Let us recall once more that these are functions: $u : V \to W$,

with the special property:

$$\forall x, y \in V, \lambda \in \mathbb{R}, u(\lambda x + y) = \lambda u(x) + u(y).$$

They are a direct generalisation of the functions:

$$x \in \mathbb{R} \mapsto ax \in \mathbb{R},$$

where $a \in \mathbb{R}$, whose graphs are the *lines*[1] $y = ax$. Indeed, such a function clearly has the linearity property: $a \cdot (\lambda x + y) = \lambda ax + ay$, and conversely, since $x = 1 \cdot x$ for every $x \in \mathbb{R}$, a linear map $u : \mathbb{R} \to \mathbb{R}$ satisfies:

$$u(x) = u(x \cdot 1) = x \underbrace{u(1)}_{\equiv a}, x \in \mathbb{R}.$$

When $V = \mathbb{R}^n$ and $W = \mathbb{R}^m$, then linear maps are entirely determined by a matrix A defined as follows: Any $x = (x_1, \ldots, x_n)$ can be decomposed as follows

$$(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i (0, \ldots, 0, \underbrace{1}_{\text{ith position}}, 0, \ldots, 0).$$

The vectors $e_i = (0, \ldots, 0, \underbrace{1}_{\text{ith position}}, 0, \ldots, 0)$ form a basis of $\mathbb{R}^n$ referred to as the standard or canonical basis of $\mathbb{R}^n$. Introducing this notation, the above equation can be written:

$$x = (x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i e_i.$$

Now let $u : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map, and $x \in \mathbb{R}^n$ be an arbitrary vector, then we can write:

$$u(x) = u(\sum_{i=1}^{n} x_i e_i) = \sum_{i=1}^{n} x_i u(e_i).$$

This shows that, generalising what we did above for linear functions in the case $n = m = 1$, that u is entirely determined by the values $u(e_i) \in \mathbb{R}^m$ it takes on the canonical basis vectors.

Since $u(e_i)$ can be decomposed onto the canonical basis of $\mathbb{R}^m$ we may find constants $A_{ij} \in \mathbb{R}$ such that

$$u(e_i) = \sum_{j=1}^{m} A_{ji} e_j.$$

[1] I guess this is where the term linear originates from

We call the matrix with coefficients (A_{ij}) is called the matrix of the linear map u (expressed in the canonical basis). Observe that it is defined so that:

$$u(x) = \sum_{j=1}^{m} \left(\sum_{i=1}^{n} A_{ji} x_i \right) e_j$$

In brackets, you might recognise the expression for the jth component of the column vector obtained by the matrix multiplication:

$$AX, \quad \text{where } X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

So if we identify instead $\mathbb{R}^n$ with the set of column matrices, we can represent a linear transformation as a map:

$$X \mapsto AX.$$

In conclusion linear maps between $\mathbb{R}^n$ and $\mathbb{R}^m$ can be thought of as multiplication maps like in the case $n = m = 1$ where the real constant a is replaced by a matrix A.

Remark 1.6.1. If V, W are arbitrary abstract, but finite dimensional vector spaces, say $\dim V = n$, $\dim W = m$, then linear maps u can still be represented as matrix multiplication, but the identification requires us to choose a bases $(v_1, \ldots, v_n)$, $(w_1, \ldots, w_n)$ of V, W. The only difference with $\mathbb{R}^n$ is that, in general, there is no standard choice of basis thus we need to specify which bases we choose.

Once the choice is made, the coefficients A_{ij} of the matrix forming what is known as the *matrix of* u *in the bases* $(v_1, \ldots, v_n)$, $(w_1, \ldots, w_n)$, are defined by the equation:

$$u(v_i) = \sum_{j=1}^{n} A_{ji} w_j,$$

and the identification of V with column matrices is:

$$x = \sum_{i=1}^{n} x_i v_i \mapsto X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Amongst all the functions from a vector space V to a vector space W, linear maps are amongst some of the most well understood. Furthermore, we have seen that in finite dimensions, they are all examples of continuous functions. Unfortunately, many functions that we want to understand are not linear...

The key idea of differential calculus, in modern terms, is **first order** or **linear** approximation.

Let us first relate this idea to the usual definition of the derivative of a function in 1 dimensions. Let $f : (a, b) \to \mathbb{R}$ be a function and $t_0 \in (a, b)$ the derivative of f at t_0, $f'(t_0)$ is the real number defined by:

$$f'(t_0) = \lim_{t \to t_0} \frac{f(t) - f(t_0)}{t - t_0} = \lim_{h \to 0} \frac{f(t_0 + h) - f(t_0)}{h}.$$

Remark 1.6.2. To go from the first quotient to the second note that it is just an application of the composition of limits by setting $t = t_0 + h$.

One might observe that this definition continues to make sense if f takes its values in some normed vector space V; in this case, we recall that it is referred to as the velocity vector.

Nevertheless, if we try to replace (a, b) by some set in $\mathbb{R}^n$ (or an abstract vector space), we will need to take $h \in \mathbb{R}^n$, but, division by h does not make sense. However, we may rewrite the definition of $f'(t_0)$ as follows:

$$\lim_{h \to 0} \frac{|f(t_0 + h) - f(t_0) - hf'(t_0)|}{|h|} = 0.$$

Let us set:

$$\varepsilon(h) = \frac{f(t_0 + h) - f(t_0) - hf'(t_0)}{h},$$

then we may write:

$$f(t_0 + h) = f(t_0) + hf'(t_0) + h\varepsilon(h),$$

where $\varepsilon(h) \to 0$ when $h \to 0$.

Now, the map $h \to f'(t_0)h$ is the archetype linear map, and therefore, we can interpret this formula as a "good" linear approximation of f near t_0: if h is small, the value of $f(t_0 + h)$ can be approximated the value obtained by adding $f'(t_0)h$ to $f(t_0)$. In this form, the definition can be generalised to functions defined on normed vector spaces.

Definition 1.12: Differentiability

Let $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ be normed vector spaces, U an **open subset** of V, $f : U \subset V \to W$ a function and $x_0 \in U$.

- f is said to be *differentiable at* x_0 if there is a continuous linear map $l : V \to W$ such that:

$$\lim_{h \to 0} \frac{\|f(x_0 + h) - f(x_0) - l(h)\|_W}{\|h\|_V}.$$

- f is said to be *differentiable (on* U*)* if it is differentiable at each point.

When l it is unique and called the differential (or derivative) of f at x_0, we write:

$$l = df_{x_0}.$$

Remark 1.6.3.

- We will do calculus on *open sets*, a reason for this will be given below.
- In finite dimensions the continuity assumption on linear maps is of course redundant, as we have seen.

In the special case $V = \mathbb{R}^n$ and $W = \mathbb{R}^m$, identified with column vectors, then, given our discussion above, we have a canonical representation of df_{x_0} as a matrix using the canonical bases of $\mathbb{R}^n$ and $\mathbb{R}^m$ as described above. This matrix will be called the *Jacobian matrix of* f, written

$$\text{Jac } f(x_0) = [Df](x_0) = \mathbf{D}f(x_0);$$

we will determine how to calculate it later.

Example 1.11.

- Our preliminary discussion shows that a differentiable function $f : (a, b) \to \mathbb{R}$ in the standard sense is differentiable in the new sense (the definitions are equivalent) and:

$$df_{x_0}(h) = f'(x_0)h, \quad h \in \mathbb{R},$$

in particular:

$$df_{x_0}(1) = f'(x_0).$$

- As one should expect, continuous linear maps $u : V \to W$ are differentiable everywhere, indeed for any $x_0 \in V$

$$u(x_0 + h) = u(x_0) + u(h),$$

and therefore:

$$\frac{\|u(x_0 + h) - u(x_0) - u(h)\|_W}{\|h\|_V} = 0, \qquad \text{for any } h \neq 0.$$

In particular:

$$du_{x_0} = u.$$

- We consider a more sophisticated example. Suppose $V = W = M_n(\mathbb{R})$ and use the norm:

$$\|A\| = \max_{X \in \mathbb{R}^n \setminus \{0\}} \frac{\|AX\|_{\mathbb{R}^n}}{\|X\|_{\mathbb{R}^n}},$$

recall that this norm has the property:

$$\|AB\| \leqslant \|A\|\|B\|.$$

Let us consider the map defined on the set of invertible matrices $GL_n(\mathbb{R})$ (which as we have seen is open):

$$\begin{array}{rcl} \phi : GL_n(\mathbb{R}) & \longrightarrow & GL_n(\mathbb{R}) \\ A & \longmapsto & A^{-1}. \end{array}$$

We will show, using the definition that ϕ is differentiable and at every point $A \in M_n(\mathbb{R})$ its differential is given by:

$$d\phi_A(H) = -A^{-1}HA^{-1}, \quad H \in M_n(\mathbb{R}).$$

Remark 1.6.4. This formula generalises to matrices the usual formula $(\frac{1}{x})' = -\frac{1}{x^2}$.

← End Lecture 5
← Start Lecture 6

Note that, for the moment, the definition does not give us a formula for $d\phi_A$ and so we must find a candidate. In general, we can do this by seeking a series (or polynomial) expansion of the function $\phi(A + H)$ and identifying the linear (first order) part in H.

To do this we will use the following fact (for your culture): one can make sense of series with values in a normed vector space V. Using the norm we can imitate the definition for numerical series, when V is finite dimensional (or more generally a Banach space) then we have the following theorem:

Theorem 1.7

If $(v_n)_{n\in\mathbb{N}}$ is a sequence of points of V, then:

$$\sum_{n=0}^{\infty} \|v_n\| < +\infty \Rightarrow \sum_{n\in\mathbb{N}} v_n \text{ converges.}$$

In this case the series is said to converge absolutely.

We return now to our example, recall, from 1 dimensional calculus that

$$\frac{1}{1+x} = \sum_{n=0}^{\infty} (-1)x^n$$

as long as $|x| < 1$. This formula generalises to matrices: if $\|A\| < 1$ then: $(I-A)$ is invertible and

$$(I-A)^{-1} = \sum_{n=0}^{+\infty} A^n,$$

where I is the identity matrix. Formally this makes sense as, supposing the series converges, then:

$$(I-A)\sum_{n=0}^{+\infty} A^n = \sum_{n=0}^{+\infty} A^n - \sum_{n=0}^{+\infty} A^{n+1} = \sum_{n=0}^{+\infty} A^n - \sum_{n=1}^{+\infty} A^n = I.$$

It remains to show that the series converges. For this observe that it follows from $\|AB\| \leqslant \|A\|\|B\|$ that for any $A \in M_n(\mathbb{R})$:

$$\|A^n\| \leqslant \|A\|^n.$$

Hence, by the triangle inequality, for fixed any $N \geqslant 0$:

$$\sum_{n=0}^{N} \|A^n\| \leqslant \sum_{n=0}^{N} \|A\|^n,$$

When $\|A\| < 1$ the standard geometric series shows that $\sum_{n=0}^{+\infty} \|A\|^n < +\infty$, therefore by the comparison theorem for positive series, $\sum_{n=0}^{+\infty} \|A^n\| < +\infty$ and taking $N \to \infty$ above:

$$\sum_{n=0}^{+\infty} \|A^n\| \leqslant \sum_{n=0}^{+\infty} \|A\|^n = \frac{1}{1-\|A\|}.$$

We shall apply this to our present example, let $A \in GL_n(\mathbb{R})$ be fixed and choose $H \in M_n(\mathbb{R})$ with $\left\|A^{-1}H\right\| < 1$, (this makes sense since we are only interested in the limit for H small), then since

$$(A+H) = A(I+A^{-1}H)$$

$(A+H)$ is invertible and:

$$(A+H)^{-1} = (I+A^{-1}H)^{-1}A^{-1} = \sum_{n=0}^{+\infty} (-1)^n (A^{-1}H)^n A^{-1}.$$

Here recall that $(AB)^{-1} = B^{-1}A^{-1}$.

Remark 1.6.5. I have been using, without proof, continuity of matrix multiplication, do you see why?

Taking out the first terms in this series we have:

$$(A+H)^{-1} = I - A^{-1}HA^{-1} + \sum_{n=2}^{+\infty}(-1)^n(A^{-1}H)^nA^{-1}.$$

The first term is first order (linear) in H:

$$-A^{-1}(\lambda H_1 + H_2)A^{-1} = -\lambda A^{-1}H_1A^{-1} - A^{-1}H_2A^{-1}.$$

Since the rest of the terms contain at least two factors of H, they will not be linear in H and so our candidate for the derivative is indeed $d\phi_A(H) = -A^{-1}HA^{-1}$. It remains to see that the linear approximation is good enough, as required by the definition. For this we must estimate, for H small (i.e. $\|AH\|^{-1} < 1$) and non-vanishing:

$$\frac{\|\phi(A+H) - \phi(A) - d\phi_A(H)\|}{\|H\|} = \frac{\left\|\sum_{n=2}^{+\infty}(-1)^n(A^{-1}H)^nA^{-1}\right\|}{\|H\|}.$$

but

$$\left\|\sum_{n=2}^{+\infty}(-1)^n(A^{-1}H)^nA^{-1}\right\| \leqslant \sum_{n=2}^{+\infty}\|A\|^{-n-1}\|H\|^n = \|H\|^2 \underbrace{\sum_{n=0}^{+\infty}\|A\|^{-n-3}\|H\|^n}_{\text{converges for } \|H\| \text{ small enough}}.$$

Hence it follows that:

$$\frac{\left\|\sum_{n=2}^{+\infty}(-1)^n(A^{-1}H)^nA^{-1}\right\|}{\|H\|} \leqslant \|H\|\sum_{n=0}^{+\infty}\|A\|^{-n-3}\|H\|^n \underset{H\to 0}{\longrightarrow} 0.$$

Which proves at the same time differentiability of ϕ at every point A and that the differential is as claimed. ◇

Let us consider one more fundamental example:

Example 1.12. Recall that a map $B : V \times V \to W$ is said to be *bilinear* if it is linear in both variables, i.e.

$$\begin{cases} B(\lambda v_1 + v_2, v_3) = \lambda B(v_1, v_3) + B(v_2, v_3), \\ B(v_1, \lambda v_2 + v_3) = \lambda B(v_1, v_2) + B(v_1, v_3). \end{cases}$$

These maps can be thought of as obeying the same rules as a product (which might not commute). It can be shown that a bilinear map is continuous when there is a constant $M > 0$ such that:

$$\|B(v_1, v_2)\|_W \leqslant M\|v_1\|_V\|v_2\|_V.$$

Remark 1.6.6. This is always the case in finite dimensions.

Remark 1.6.7. One may, for instance, define B to be the map of matrix multiplication: $(A_1, A_2) \in M_n(\mathbb{R}) \times M_n(\mathbb{R}) \mapsto A_1 A_2 \in M_n(\mathbb{R})$.

We will justify that:

> Any continuous bilinear map $B : V \times V \to W$ is differentiable everywhere and at any point $(v_1, v_2) \in V \times V$
>
> $$dB_{(v_1,v_2)}(h_1, h_2) = B(v_1, h_2) + B(h_1, v_2).$$

This general statement is at the root of all the "product rules". In essence, any product will satisfy the product rule.

In this instance, it is convenient to use a different (but equivalent) norm on the product space $V \times V$ namely:

$$\|(v_1, v_2)\|_{V\times V} = \sqrt{\|v_1\|_V^2 + \|v_2\|_V^2}.$$

Now, if $(h_1, h_2) \in V \times V$ then let us calculate:

$$\begin{aligned} B(v_1 + h_1, v_2 + h_2) &= B(v_1 + h_1, v_2) + B(v_1 + h_1, h_2) \\ &= B(v_1, v_2) + \underbrace{B(h_1, v_2) + B(v_1, h_2)}_{\text{linear part in } (h_1, h_2)} + B(h_1, h_2). \end{aligned}$$

We now need to estimate:

$$\|B(v_1 + h_1, v_2 + h_2) - B(v_1, v_2) - (B(h_1, v_2) + B(v_1, h_2))\|_W = \|B(h_1, h_2)\|_W.$$

But:

$$\|B(h_1, h_2)\|_W \leqslant M\|h_1\|_V\|h_2\|_V \leqslant \frac{M}{2}\left(\|h_1\|_V^2 + \|h_2\|_V^2\right) = \frac{M}{2}\|(h_1, h_2)\|_{V\times V}^2.$$

Remark 1.6.8. If $a, b \in \mathbb{R}$, then since $(a - b)^2 \geqslant 0$ it follows that:

$$2ab \leqslant a^2 + b^2.$$

We have used this fact above.

Now we conclude that:

$$\frac{\|B(h_1, h_2)\|_W}{\|(h_1, h_2)\|_{V\times V}} \leqslant \frac{M}{2}\|(h_1, h_2)\|_{V\times V} \underset{(h_1,h_2)\to(0,0)}{\longrightarrow} 0.$$

Which proves again that B is differentiable everywhere and the derivative is as claimed. ◇

1.7 Directional and partial derivatives

1.7.1 Directional derivatives

Let $f : U \subset (V, \|\cdot\|_V) \to (W, \|\cdot\|_W)$ be a function defined on an open set U of the vector space V.

There is another way in which one could generalise the derivative to higher dimensions: one might try to pick some direction defined by a vector $h \in V$ and start to walk in a straight line away from x_0 in the direction defined by h. This means we will walk along the curve φ defined by $\gamma(t) = x_0 + th, t \in \mathbb{R}$.

Observe here that the choice of h is an arbitrary element of V, indeed, whilst the unit vector $\frac{h}{\|h\|_V}$ controls the direction in which one walks, the norm $\|h\|_V$ controls how *fast* we run in that direction and we can certainly run as fast as we want in any direction.

Now whilst the curve makes for all $t \in \mathbb{R}$, $x_0 + th$, it may certainly leave U at some point. However, *since* U *is assumed open*, $x_0 + th$ will be an element of U as long as t is small enough.

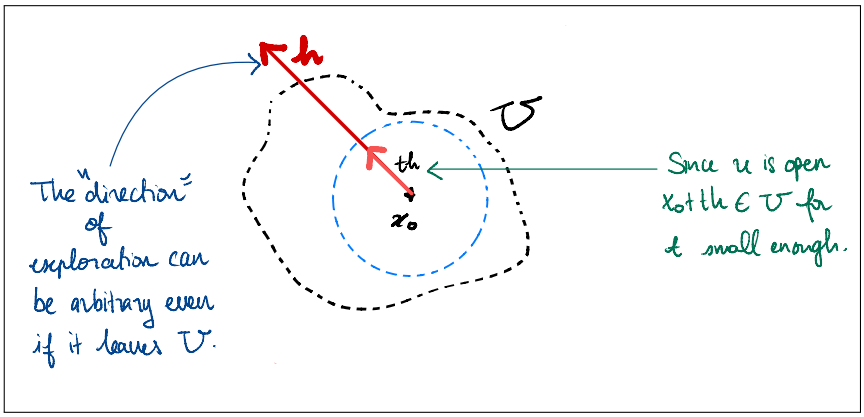


Figure 1.6: Directional derivative

This is because by definition of open sets, if $x_0 \in U$ then there is some open ball $B(x_0, \delta)$ that is entirely contained within U. Hence, $x_0 + th$ will stay in U at least as long:

$$|t| < \frac{\delta}{\|h\|_V}.$$

This is illustrated in Figure 1.6. Observe that this is coherent with the fact that the faster I run along the curve, the less time I will have before I leave U.

Hence, the composition $\varphi(t) = f \circ \gamma$ will make sense as long as $t \in (-\frac{\delta}{\|h\|_V}, +\frac{\delta}{\|h\|_V})$ and we can consider its tangent vector at $t = 0$, this leads to:

Definition 1.13: Directional derivative

Using the notations above, the *directional derivative* in the direction of h of f at x_0 $D_h f(x_0) = \frac{\partial f}{\partial h}(x_0)$ is defined to be:

$$\varphi'(t) = \lim_{t \to 0} \frac{f(x_0 + th) - f(x_0)}{t},$$

if the limit exists.

An important question we must answer is:

Suppose f is *differentiable*, what is the link, if any, between these derivatives and df_{x_0}?

For this we use our composition of limits theorem, when $t \to 0$ when $th \to 0$ so from the definition of the derivative (substituting h for th, where h is now *fixed* and t is the variable), we have:

$$\lim_{t \to 0} \frac{\|f(x_0 + th) - f(x_0) - df_{x_0}(th)\|_W}{\|th\|_V} = 0.$$

Now, by the homogeneity property of norms, it follows that $\|th\|_V = |t|\|h\|_V$. Furthermore, df_{x_0} is a linear function so: $df_{x_0}(th) = t df_{x_0}(th)$. So we can conclude that:

$$\lim_{t \to 0} \left\| \frac{f(x_0 + th) - f(x_0) - t df_{x_0}(h)}{t\|h\|_V} \right\|_W = \lim_{t \to 0} \left\| \frac{f(x_0 + th) - f(x_0)}{t\|h\|_V} - \frac{df_{x_0}(h)}{\|h\|_V} \right\|_W = 0.$$

Since multiplication by a scalar is continuous, and h is fixed, we can multiply the expression in the limit by $\|h\|_V$ and conclude that:

$$\lim_{t \to 0} \left\| \frac{f(x_0 + th) - f(x_0)}{t} - df_{x_0}(h) \right\|_W = 0.$$

Hence, we conclude:

Proposition 1.3

When f is differentiable at x_0, directional derivatives exist in all directions h and:

$$D_h f(x_0) = \frac{\partial f}{\partial h}(x_0) = df_{x_0}(h).$$

Remark 1.7.1.

- An immediate consequence is that we have also justified the uniqueness of df_{x_0} since we can calculate it from its directional derivatives in any direction. To obtain this uniqueness is part of the reason why we assume U open.

- When $\dim V < +\infty$, df_{x_0} is completely determined by the directional derivatives of f in the directions of some basis $(e_1, \ldots, e_n)$.

1.7.2 Partial derivatives for functions defined on $V = \mathbb{R}^n$

When $V = \mathbb{R}^n$, as we have already observed, $\mathbb{R}^n$ possesses a canonical basis given by the vectors:

$$e_i = (0, \ldots, 0, \underbrace{1}_{\substack{\uparrow \\ \text{ith position}}}, 0, \ldots, 0).$$

Definition 1.14: Partial derivatives

Let $f : U \subset \mathbb{R}^n \to W$ be a function defined on an open set U. Let $x_0 \in U$, then the ith partial derivative of f is defined by:

$$\frac{\partial f}{\partial x_i}(x_0) = D_{e_i} f(x_0).$$

When f is differentiable at x_0, we can calculate completely df_{x_0} from the partial derivatives, indeed, if

$$h = (h_1, \ldots, h_n) = \sum_{i=1}^{n} h_i e_i,$$

then, by linearity:

$$df_{x_0}(h) = df_{x_0}(\sum_{i=1}^{n} h_i e_i) = \sum_{i=1}^{n} h_i df_{x_0}(e_i) = \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(x_0).$$

A notation that is often used to reconcile with some usages in the scientific literature is the following.

Define the linear map:

$$\begin{array}{rcl} dx^i : \mathbb{R}^n & \to & \mathbb{R} \\ (x_1, \ldots, x_n) & \mapsto & x_i. \end{array}$$

Then in this case:

$$dx^i(h) = h_i,$$

and plugging this into expression we found for $df_{x_0}(h)$ we arrive at:

$$df_{x_0}(h) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0) dx^i(h).$$

Hence, we can write:

$$df_{x_0} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0) dx^i.$$

⚠ We have shown that *if* f is differentiable at x_0 *then* its partial derivatives exist in every direction and determine df_{x_0}. However, the converse is false: even if all the partial derivatives exist f might not be differentiable. e.g. the linear approximation isn't good enough.

To illustrate this, consider the real valued function defined on $\mathbb{R}^2$ by:

$$f(x,y) = \begin{cases} \frac{xy}{\sqrt{x^2+y^2}} & (x,y) \neq (0,0), \\ 0 & (x,y) = (0,0), \end{cases}$$

then f is continuous on $\mathbb{R}^2$, indeed, for any point $(x,y) \neq (0,0)$ its expression is a rational function and therefore continuous where defined.

The only point that might pose problem is $(0,0)$ but there:

$$0 \leqslant \frac{|xy|}{\sqrt{x^2+y^2}} \leqslant \frac{1}{2}\sqrt{x^2+y^2} \to_{(x,y)\to 0} 0,$$

so

$$\lim_{(x,y)\to(0,0)} f(x,y) = 0 = f(0,0).$$

For the same reason as above, differentiability is clear when $(x,y) \neq (0,0)$, so we only need to study the point $(0,0)$.

Now, for any $t \in \mathbb{R}$, $f(t,0) = f(t,0) = 0$ so the partial derivatives $\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = 0$ both vanish. This means, that *if* f is differentiable at $(0,0)$, $df_{(0,0)} = 0$.

It remains to check that the linear approximation is good enough, as required by the definition, let us compute:

$$\frac{|f(h_1,h_2) - f(0,0)|}{\sqrt{h_1^2+h_2^2}} = \frac{h_1 h_2}{h_1^2+h_2^2},$$

and it is standard exercise to check that this limit does *not* exist. Hence f is not differentiable at the point $(0,0)$.

It turns out however, that if we assume that the partial derivatives are well-behaved, then we recover a partial converse:

Theorem 1.8: Continuous partial derivatives

Let $f : U \subset \mathbb{R}^n \to (W, \|\cdot\|_W)$ be a function with continuous partial derivatives $\frac{\partial f}{\partial x_i}$ on U then f is differentiable on U.

← End Lecture 6

1.7.3 Complement: A more general notion of partial derivative (can be skipped)

In this optional reading segment, we discuss another notion of partial derivative that can sometimes be useful, at least to simplify notation. For instance, in practice, if a function is defined on $\mathbb{R}^n$, then it can sometimes be convient to *group together* the k first variables and then the $n-k$ remaining ones: this amounts to writing $\mathbb{R}^n$ as the product space: $\mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^m$.

Let us therefore study a function f defined on an open subset of the product space $E_1 \times E_2$ of two normed vector spaces, $(E_1, \|\cdot\|_{E_1})$, $(E_2, \|\cdot\|_{E_2})$. As usual we will equip $E_1 \times E_2$ with the norm $\|(x,y)\|_{E_1\times E_2} = \|x\|_{E_1} + \|y\|_{E_2}$. Let V be a third normed vector space $(V, \|\cdot\|_V)$, and consider a function:

$$\begin{array}{rccc} f: & U \subset E_1 \times E_2 & \longrightarrow & V \\ & (x,y) & \longmapsto & f(x,y) \end{array}$$

Now, just like, when x, y are real variables, we can consider the functions obtained by fixing one these vector valued variables. To formalise this, let $(x_0, y_0) \in E_1 \times E_2$ and define two maps:

$$\begin{cases} j^2_{x_0}: & E_2 & \longrightarrow & E_1 \times E_2 \\ & y & \longmapsto & (x_0, y) \\ j^1_{y_0}: & E_1 & \longrightarrow & E_1 \times E_2 \\ & x & \longmapsto & (x, y_0) \end{cases}.$$

If the function $f \circ j^1_{y_0}$ (resp. $f \circ j^2_{x_0}$) ix differentiable at x_0 (resp. y_0), then we define the partial derivative of f at (x_0, y_0) by:

$$\frac{\partial f}{\partial x}(x_0, y_0) = d(f \circ j^1_{y_0})_{x_0}, \quad \left(\text{resp. } \frac{\partial f}{\partial y}(x_0, y_0) = d(f \circ j^2_{x_0})_{y_0}\right).$$

⚠ It is important to observe that although the notation for the partial derivative is the same, these are no longer just vectors and have themselves been upgraded to continuous linear functions ! Recall that the set of continuous linear functions between two vectors spaces $E \to F$ is written $\mathcal{L}(E, F)$, so we have:

$$\frac{\partial f}{\partial x}(x_0, y_0) \in \mathcal{L}(E_1, V), \quad \frac{\partial f}{\partial y}(x_0, y_0) \in \mathcal{L}(E_2, V).$$

The space of continuous linear functions $\mathcal{L}(E,F)$ can be made a normed space with the norm:

$$\|u\|_{\mathcal{L}(E,F)} = \sup_{\substack{x\in E\\ x\neq 0}} \frac{\|u(x)\|_F}{\|x\|_E}.$$

It is a straightforward exercise to show that the maps $j^1_{y_0}$ and $j^2_{x_0}$ are differentiable so it will follows from the chain rule that *if* f is differentiable then it has partial derivatives.

Just as in the previous case if a function f has *continuous* partial derivatives on U that is the maps:

$$\frac{\partial f}{\partial x}: \begin{array}{ccc} E_1\times E_2 & \longrightarrow & \mathcal{L}(E_1,V)\\ (x,y) & \longmapsto & \frac{\partial f}{\partial x}(x,y)\end{array}, \qquad \frac{\partial f}{\partial x}: \begin{array}{ccc} E_1\times E_2 & \longrightarrow & \mathcal{L}(E_2,V)\\ (x,y) & \longmapsto & \frac{\partial f}{\partial y}(x,y)\end{array}$$

are continuous then f is differentiable at U.

Start Lecture 7→

1.8 The chain rule

We are now going to discuss the composition theorem for differentiable functions, which is an extremely important result, both for theory and applications. We state it first in its general form and then we will discuss in detail the specific case for the vector spaces $\mathbb{R}^n$.

Theorem 1.9: The chain rule

Let E_1, E_2, E_3 be normed vector spaces, $U_1 \subset E_1, U_2 \subset E_2$ open subsets and:

$$f: U_1 \subset E_1 \to E_2, \quad g: U_2 \subset E_2 \to E_3,$$

with $f(U_1) \subset U_2$. Assume f is differentiable at x_0 and g is differentiable at $f(x_0)$, then the composition $(g\circ f)$ is differentiable at x_0 and:

$$d(g\circ f)_{x_0} = dg_{f(x_0)} \underbrace{\circ}_{\substack{\uparrow\\ \text{composition}\\ \text{of linear maps}}} df_{x_0}.$$

In some sense, d distributes through $\circ$.

We begin with some examples:

Example 1.13.

- Let $f(a,b) \to (c,d)$ and $g: (c,d) \to \mathbb{R}$ satisfy the hypothesis above, then:
$$df_{x_0} = f'(x_0)dx, \quad dg_{f(x_0)} = g'(f(x_0))dx.$$

Remark 1.8.1. In 1 dimension the dx is just the identity map: $dx(h) = h$.

So by the above theorem:

$$d(g \circ f)_{x_0} = g'(f(x_0))f'(x_0)dx,$$

Since the link between the differential and the standard derivative is $(g \circ f)'(x_0) = d(g \circ f)_{x_0}(1)$, we recover the usual formula:

$$(g \circ f)'(x_0) = g'(f(x_0))f'(x_0).$$

- Let us consider now the case of a *curve* in $\mathbb{R}^2$:

$$\gamma : (a, b) \to \mathbb{R}^2, \quad \gamma(t) = (x(t), y(t)).$$

Remark 1.8.2. These are usually interpreted as representing the trajectory of a particle, here in a plane.

And consider $V : \mathbb{R}^2 \to \mathbb{R}$ a scalar valued function (thought of as a potential of some sort), then:

$$d\gamma_t = (x'(t), y'(t))dt, \quad dV_{(x,y)} = \frac{\partial V}{\partial x}(x, y)dx + \frac{\partial V}{\partial y}(x, y)dy.$$

Remark 1.8.3. Again, in one variable dt is just the identity map $h \mapsto h$, in two variables $dx : (h_x, h_y) \mapsto h_x$ and $dy : (h_x, h_y) \mapsto h_y$. In particular:

$$dx((x(t), y'(t))dt) = x'(t)dt,$$

and similarly for dy.

Therefore, by the chain rule:

$$d(V \circ \gamma)_t = \frac{\partial V}{\partial x}(x, y)x'(t)dt + \frac{\partial V}{\partial y}(x, y)y'(t)dt,$$

evaluating this at $h = 1$, we find that:

$$(g \circ \gamma)'(t) = \frac{d(V \circ \gamma)}{dt}(t) = \frac{\partial V}{\partial x}(x, y)\frac{dx}{dt}(t) + \frac{\partial V}{\partial y}(x, y)\frac{dy}{dt}(t).$$

Remark 1.8.4. $(g \circ \gamma)'(t)$ is interpreted as the instantaneous rate of change of the potential V felt by the particle at time t.

$\diamond$

Derivatives of functions *into* a product space (optional reading)

Assume $(Z, \|\cdot\|_Z)$, $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ are normed vector spaces. Recall that the projection maps: $\pi_V : V \times W \to V$ and $\pi_W : V \times W \to W$ are continuous and linear and so *differentiable everywhere* and:

$$d\pi_V = \pi_V, \qquad d\pi_W = \pi_W.$$

Just as was the case for continuous functions, we have the following characterisation of differentiability for these functions:

Theorem 1.10: Differentiability in product spaces

Let $f : U \subset Z \to V \times W$ be a function defined on an open set U of Z. It can be written:

$$f(z) = (f_V(z), f_W(z)), z \in U$$

where $f_V = \pi_V \circ f$ and $f_W = \pi_W \circ f$. Then f is differentiable at $z_0 \in Z$ if and only if the functions $f_V : U \to V$ and $f_W : U \to W$ are differentiable at z_0, moreover:

$$df_{z_0} = (df_{V z_0}, df_{W z_0}).$$

The moral of the story is that we can differentiate component by component.

Proof. If f is differentiable at z_0 then, by the chain rule, f_V and f_W are differentiable at z_0. Conversely, assume f_V and f_W are differentiable at z_0 working with the product norm $\|(v, w)\|_{V\times W} = \|v\|_V + \|w\|_W$, we estimate:

$$\begin{aligned}
&\frac{\left\|f(z_0+h) - f(z_0) - (df_{V z_0}(h), df_{W z_0}(h))\right\|_{V\times W}}{\|h\|_Z} \\
&\qquad = \frac{\left\|(f_V(z_0+h), f_W(z_0+h)) - (f_V(z_0), f_W(z_0)) - (df_{V z_0}(h), df_{W z_0}(h))\right\|_{V\times W}}{\|h\|_Z}, \\
&\qquad = \frac{\left\|(f_V(z_0+h) - f_V(z_0) - df_{V z_0}(h), f_W(z_0+h) - f_W(z_0) - df_{W z_0}(h))\right\|_{V\times W}}{\|h\|_V}.
\end{aligned}$$

Finally, by definition of the norm $\|\cdot\|_{V\times W}$ this is simply:

$$\frac{\left\|f_V(z_0+h) - f_V(z_0) - df_{V z_0}(h)\right\|_V}{\|h\|_V} + \frac{\left\|f_W(z_0+h) - f_W(z_0) - df_{W z_0}(h)\right\|_V}{\|h\|_V},$$

and both of these terms vanish when $h \to 0$ by assumption, which proves the theorem. □

– End of optional reading

We have the following important result:

Proposition 1.4

If $f : U \subset V \to \mathbb{R}^m$, U open in V then if we write: $f(v) = (f_1(v), \ldots, f_m(v))$, then: f is differentiable at $v_0 \in V$ if and only if the real-valued functions $f_1, \ldots, f_m$ are differentiable at z_0 and:

$$df_{v_0} = (df_{1v_0}, \ldots, df_{mv_0}).$$

One obtains the differential of an $\mathbb{R}^m$ valued map by differentiating each of its components separately.

1.9 The Jacobian matrix

Recall that any linear map $u : \mathbb{R}^n \to \mathbb{R}^m$ can be represented by a matrix A. (See the discussion at the beginning of Section 1.6.1.)

Since, by definition, the differential df_{x_0} of a function $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ is a linear map from $\mathbb{R}^n \to \mathbb{R}^m$ therefore can be described by a matrix. This $\mathbf{m} \times \mathbf{n}$ matrix is known as the **Jacobian matrix**. We are now ready to determine its components. Let us write for any $x = (x_1, \ldots, x_n) \in U$:

$$f(x) = (f_1(x), \ldots, f_m(x))$$

We need to determine the components of $df_{x_0}(e_i) = \frac{\partial f}{\partial x_i}(x_0)$, where e_i is the ith vector in the canonical basis of $\mathbb{R}^n$ but we have just seen that:

$$df_{x_0}(e_i) = (df_{1x_0}(e_i), \ldots, df_{mx_0}(e_i)) = (\frac{\partial f_1}{\partial x_i}(x_0), \ldots, \frac{\partial f_m}{\partial x_i}(x_0).$$

So the components of the Jacobian matrix are:

$$(\text{Jac } f(x_0))_{ij} = \frac{\partial f_i}{\partial x_j}(x_0),$$

or in other words:

$$\text{Jac } f(x_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \frac{\partial f_1}{\partial x_2}(x_0) & \ldots & \frac{\partial f_1}{\partial x_n}(x_0) \\ \frac{\partial f_2}{\partial x_1}(x_0) & \ddots & \vdots & \vdots \\ \vdots & \ddots & \vdots & \vdots \\ \frac{\partial f_m}{\partial x_1}(x_0) & \ldots & \ldots & \frac{\partial f_m}{\partial x_n}(x_0) \end{pmatrix}.$$

Example 1.14. Let $f(x, y, z) = (x^2y, 3x + z, z^2 - 1, y\sin^2(z))$ then:

$$\text{Jac } f(x, y, z) = \begin{pmatrix} 2xy & x^2 & 0 \\ 3 & 0 & 1 \\ 0 & 0 & 2z \\ 0 & \sin^2(z) & 2y\cos(z)\sin(z) \end{pmatrix}$$

◇

Remark 1.9.1. The Jacobian matrix exists as soon as the partial derivatives exist, even when f is not differentiable.

If we identify vectors in $\mathbb{R}^n$ with column vectors, then the differential df_{x_0} is represented by the linear transformation:

$$\mathbf{h} = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \mapsto \text{Jac } f(x_0)\mathbf{h}.$$

So the directional derivatives are obtained by standard matrix multiplication. Similarly the definitions give the following matrix version of the chain rule:

Theorem 1.11: Chain rule with Jacobian matrices

Let $f : U_1 \subset \mathbb{R}^n \to \mathbb{R}^m$, $g : U_2 \subset \mathbb{R}^m \to \mathbb{R}^p$, U_1, U_2 open and $f(U_1) \subset U_2$. Suppose f differentiable at $x_0 \in U$, g differentiable at $f(x_0)$, then $g \circ f$ is differentiable at x_0 and:

$$\text{Jac } (g \circ f)(x_0) = \text{Jac } g(f(x_0))\text{Jac } f(x_0).$$

End Lecture 7→

Remark 1.9.2. The multiplication above is the standard matrix multiplication.

1.10 The gradient vector field, level sets

Start Lecture 8→

We will now proceed to apply the general theory to specific cases that often arise in applications. Our first step is going to be the case of *scalar fields*, which are functions with values in $\mathbb{R}$. These are used to model things like energy potential fields, electrostatic potentials, or the temperature in a region of space. In fact, we will now restrict to the special case of Euclidean spaces E, with their Euclidean norm: $\|u\| = \sqrt{\langle u, u\rangle}$, where we will use the notation $\langle \cdot, \cdot \rangle$ to denote the inner/scalar/dot product.

1.10.1 The gradient of a scalar field in Euclidean space

Feel free to set $E = \mathbb{R}^n$. When $f : U \subset E \to \mathbb{R}$ is a differentiable function, then at any point of x_0, df_{x_0} is a continuous real-valued linear map. In fact, the vector space $\mathcal{L}(E, \mathbb{R})$ of all continuous linear functions is particularly important in some applications (for instance momentum in physics is more naturally an element of the dual) and is known as the **dual space** of E; it is written E'.

In some sense one may think of elements of E' as "measuring tools" that can assign to any point of E a real-value. In a Euclidean space we can use the inner product to

produce elements of E'. Indeed, if $v \in E$ then one can define a linear map by:

$$l_v(x) = \langle v, x\rangle, x \in E.$$

It turns out that any element of E' can be obtained in this way.

Theorem 1.12: Riesz representation theorem

The mapping:

$$\begin{array}{ccc} E & \to & E' \\ v & \mapsto & l_v = \langle v, \cdot\rangle, \end{array}$$

is a vector space isomorphism.
In other words, given any linear map $u : E \to \mathbb{R}$ one can find a unique vector v such that:

$$\forall h \in E, u(h) = \langle v, h\rangle.$$

Example 1.15. You have already encountered this phenomenon in elementary geometry in space. Let $\mathscr{P}$ be a vector plane in $\mathbb{R}^2$, then it is defined by an equation:

$$ax + by + cz = 0.$$

Now let us set:

$$u(x, y, z) = ax + by + cz,$$

then u is a linear function from $\mathbb{R}^3 \to \mathbb{R}$. Indeed:

$$\lambda(x_1, y_1, z_1) + (x_2, y_2, z_2) = (\lambda x_1 + x_2, \lambda y_1 + y_2, \lambda z_1 + z_2),$$

and:

$$u(\lambda x_1+x_2, \lambda y_1+y_2, \lambda z_1+z_2) = a(\lambda x_1+x_2)+b(\lambda y_1+y_2)+c(\lambda z_1+z_2) = \lambda u(x_1, y_1, z_1)+u(x_2, y_2, z_2).$$

However, if $\vec{n} = (a, b, c)$ and we set $\vec{x} = (x, y, z)$ we have:

$$u(x, y, z) = \langle \vec{n}, \vec{x}\rangle.$$

Here the duality between E and E' by the fact that one can either describe the plane $\mathscr{P}$ with its equation, or by specifying a normal vector $\vec{n}$. ◇

In our present context, we will use the duality to replace the differential by a vector field;

Definition 1.15: Gradient of a scalar field

The gradient vector at $x_0 \in U$ of a differentiable scalar field $f : U \subset E \to \mathbb{R}$ is the *unique vector*[a] $\overrightarrow{\nabla} f(x_0) \in E$ such that:

$$\forall \vec{h} \in E, \quad df_{x_0}(\vec{h}) = \langle \overrightarrow{\nabla} f(x_0), \vec{h} \rangle.$$

[a] can also be written $\overrightarrow{\text{grad}} f(x_0)$ or $\nabla f(x_0)$

Remark 1.10.1. The advantage of this definition is that it is basis independent !

Example 1.16. We now study the fundamental example where $E = \mathbb{R}^n$ is the standard Euclidean space. Now if $f : U \subset \mathbb{R}^n \to \mathbb{R}$ is differentiable then we know that:

$$df_{x_0} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0) dx^i,$$

or applied to a vector $\vec{h} = (h_1, \ldots, h_n)$:

$$df_{x_0}(\vec{h}) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0) h^i = \langle \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0) \vec{e}_i, \vec{h} \rangle,$$

where $\vec{e}_i$ are the canonical basis vectors. Therefore we conclude that:

$$\overrightarrow{\nabla} f(x_0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0) \vec{e}_i,$$

or written as a column vector:

$$\overrightarrow{\nabla} f(x_0) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x_0) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x_0) \end{pmatrix}.$$

Remark 1.10.2. One can observe that this is just the transpose of the Jacobian matrix. In fact, if we identify vectors of $\mathbb{R}^n$ with column matrices then the elements of its dual space $(\mathbb{R}^n)'$ are naturally identified with *row* matrices. In this special case, the identification of an element of $(\mathbb{R}^n)'$ with an element of $\mathbb{R}^n$ by the Riesz representation theorem is just obtained by taking the transpose.

$\diamond$

If we consider the map:

$$\begin{array}{rccc} \overrightarrow{\nabla} f : & U & \longrightarrow & \mathbb{R}^n \\ & x & \longmapsto & \overrightarrow{\nabla} f(x) \end{array},$$

it is an important example of a **vector field** on $\mathbb{R}^n$.

It encodes local data about the behaviour of f near each point of U, and is visualised as "attaching an arrow" to each point x_0 of $\mathbb{R}^n$. One could say that it lives in the *tangent space* to $x_0 \in U$, the set of all tangent vectors at x_0 of curves in U, but U being open we can walk in straight lines in any direction so $T_{x_0}U \simeq \mathbb{R}^n$.

For instance, in $\mathbb{R}^2$, the gradient vector field of the function:

$$f(x, y) = e^{-\cos^2 x + \sin^2 y}$$

is represented in Figure 1.7.

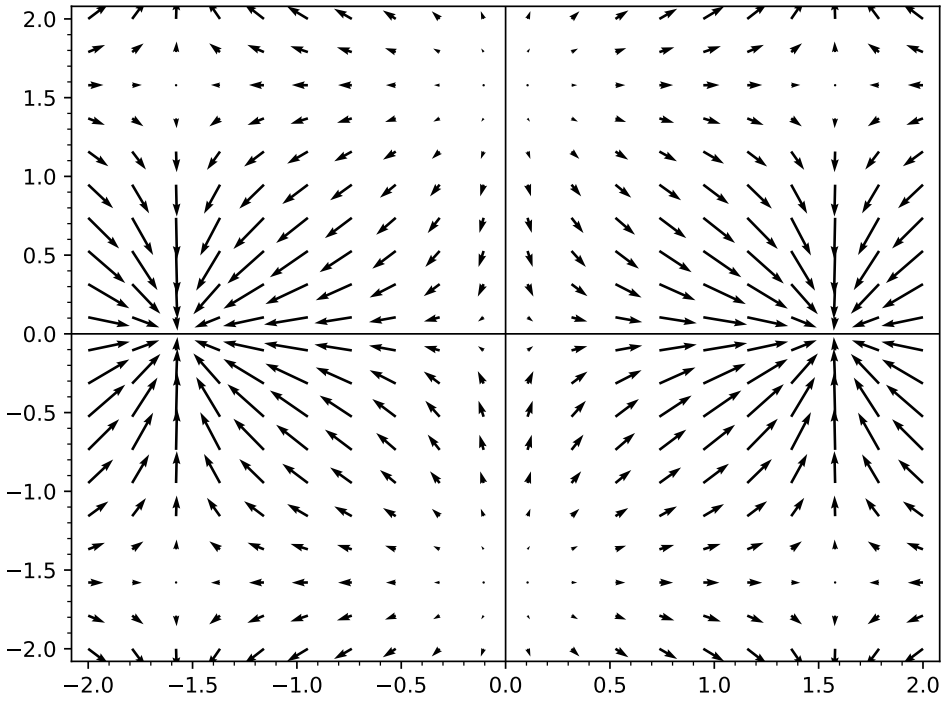


Figure 1.7: Gradient vector field of the scalar field: $f(x, y) = e^{-\cos^2 x + \sin^2 y}$.

1.10.2 Level sets and the interpretation of the gradient vector field

We will now study how to interpret the gradient vector field. Let us reflect on the following question, how can I obtain a visual understanding of a scalar field if I "live" in its domain. For instance, suppose I want to understand the temperature distribution at each point in space, that I model by a function T which assigns to each point (x, y, z) in space the temperature $T(x, y, z)$ at that point. Now this is a function:

$$T : \mathbb{R}^3 \to \mathbb{R},$$

Hence, its graph will live in $\mathbb{R}^3 \times \mathbb{R} = \mathbb{R}^4$, but I cannot draw in 4 dimensions !

Idea: We can look at places in $\mathbb{R}^3$ where T is constant ! These are known as the level sets:

Definition 1.16: Level sets

Let $f : U \subset \mathbb{R}^n \to \mathbb{R}$ be a differentiable scalar field on an open set U of $\mathbb{R}^n$, the sets:

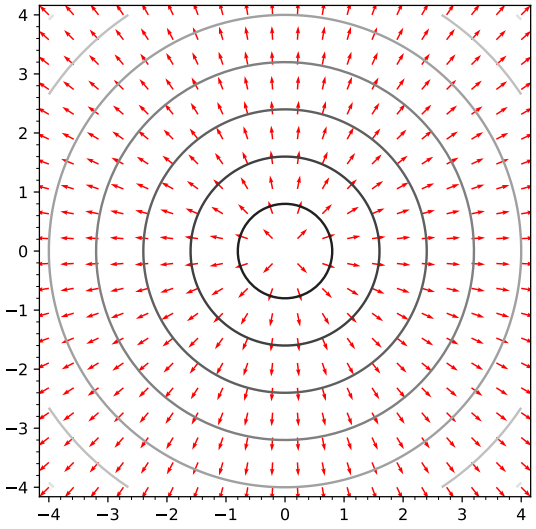
$$f^{-1}(\{c\}) = \{x \in \mathbb{R}^n, f(x) = c\}, \quad c \in \mathbb{R},$$

are known as the *level sets* of f.

Example 1.17. Consider:

$$\begin{array}{rccc} f: & \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ & (x,y) & \longmapsto & \sqrt{x^2+y^2} \end{array},$$

then its level sets are concentric circles as shown below:



◇

We have also represented the gradient vector field of this scalar field on the same diagram. Observe that the arrows representing this vector field appear to be orthogonal to the circles, i.e. perpendicular to the tangent line of the circle at each point.

This means that if $\vec{h}$ is a vector parallel to the tangent line at x_0 then:

$$df_{x_0}(\vec{h}) = \langle \overrightarrow{\nabla} f(x_0), \vec{h} \rangle = 0,$$

i.e. the directional derivative in the direction $\vec{h}$ will vanish. Intuitively, this makes sense, an infinitesimal displacement in the direction of the tangent line will remain on the circle, where f is constant.

Another way of formalising this is to say that the tangent vector $\gamma'(0)$ to any curve $\gamma : (-\varepsilon, \varepsilon) \to \mathbb{R}^2$ with $\gamma(0) = x_0$ that remains restricted the circle, will be parallel to the tangent line to the circle at x_0.

We generalise this as follows:

Definition 1.17: Tangent vector space to level surfaces

Let $\mathscr{L} = \{x \in U, F(x) = c\}$ where $c \in \mathbb{R}$ and $F : U \subset \mathbb{R}^n \to \mathbb{R}$ is a differentiable scalar field with continuous partial derivatives. Let $x_0 \in \mathscr{L}$, we call x_0 a regular point of $\mathscr{L}$ if $\overrightarrow{\nabla} f(x_0) \neq 0$.
In this case, the *tangent vector space* to $\mathscr{L}$ at x_0 $T_{x_0}\mathscr{L}$ is defined to be the set of tangent vectors to curves drawn on $\mathscr{L}$, i.e.

$$\vec{v} \in T_{x_0}\mathscr{L} \Leftrightarrow \exists\, \gamma : \begin{array}{ccc} (-\varepsilon, \varepsilon) & \longrightarrow & \mathbb{R} \\ t & \longmapsto & \gamma(t) \end{array}, \quad \gamma(0) = x_0, \quad \gamma'(t) = \vec{v}.$$

Remark 1.10.3. We will later show that $T_{x_0}\mathscr{L}$ *is* a vector space, but we do not yet have the tools for that.

Proposition 1.5

Let $\mathscr{L}, x_0, F$ satisfy the conditions of the previous theorem, then:

$$T_{x_0}\mathscr{L} = \{\vec{h} \in \mathbb{R}^n, \langle \overrightarrow{\nabla} F(x_0), \vec{h} \rangle = 0\}.$$

Proof. We give only a partial justification, once more the other direction requires a new tool.

Let $\gamma : (-\varepsilon, \varepsilon) \longrightarrow \mathscr{L}$ be some curve with $\gamma(0) = x_0$, then for every $t \in \mathbb{R}$,

$$(F \circ \gamma)(t) = c,$$

hence is constant therefore:

$$(F \circ \gamma)'(t) = 0.$$

However, by the chain rule:

$$0 = (F \circ \gamma)'(0) = d(F \circ \gamma)_0(1) = dF_{x_0}(\gamma'(0)) = \langle \overrightarrow{\nabla} F(x_0), \gamma'(0) \rangle.$$

So this shows that $T_x\mathscr{L} \subset \{\vec{h} \in \mathbb{R}^n, \langle \overrightarrow{\nabla} F(x_0), \vec{h} \rangle = 0\}$; the other inclusion will be shown later. □

← End Lecture 8
← Start Lecture 9

The reader might have observed that I defined the tangent space as a *vector space*. However, it is of course intrinsically an object linked to the point $x_0 \in \mathscr{L}$: it describes all the tangent vectors at that point. For this reason, we actually represent it in $\mathbb{R}^n$ as a plane passing through the point x_0. In other words, we translate the vector plane (which contains the zero vector) to the point x_0. This is also called the *tangent space*, or affine tangent space to distinguish from the tangent vector space: the latter defines the *direction* of the affine space.

Definition 1.18: (Affine) Tangent Space

Let $\mathscr{L}, x_0, F$ satisfy the conditions above, then the affine tangent space $\tilde{T}_{x_0}\mathscr{L}$ at x_0 is defined as follows:

$$x \in \tilde{T}_{x_0}\mathscr{L} \Leftrightarrow x - x_0 \in T_{x_0}\mathscr{L}.$$

Remark 1.10.4. This is completely analogous to the fact that the tangent vector to a point M on a curve is systematically represented as a vector attached to the point M.

Combining this with the characterisation of the vector tangent space in terms of the gradient vector we have:

Proposition 1.6: Equation of the affine tangent plane

$$x \in \tilde{T}_{x_0}\mathscr{L} \Leftrightarrow \langle \vec{\nabla} f(x_0), x - x_0 \rangle = 0.$$

In particular if $x = (x_1, \ldots, x_n)$ and $x_0 = (x_1^0, \ldots, x_n^0)$ then the equation of the plane is given by:

$$\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0)(x_i - x_i^0) = 0$$

Example 1.18.

- Let $f : \mathbb{R}^2 \to \mathbb{R}$.

  The graph of f can be described as a 0 level set of the function $F : (x, y, z) \mapsto f(x, y) - z$, it follows that equation of the (affine) tangent plane at $(x_0, y_0, z_0 = f(x_0, y_0))$ is given by:

  $$\frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) - (z - f(x_0, y_0)) = 0.$$

- The tangent line to the circle $x^2 + y^2 = 1$ at the point $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ is defined by the equation:

  $$(x - \frac{1}{\sqrt{2}}) + (y - \frac{1}{\sqrt{2}}) = 0.$$

$\diamond$

We return now to the discussion about interpreting the gradient. We have established that the gradient vector field $\vec{\nabla}F$ of a scalar field F is, when non-vanishing, *orthogonal* to the level sets of F. By this we mean precisely, that it is at each point x_0 orthogonal to the tangent plane $T_{x_0}F^{-1}(F(x_0))$. The tangent space contains the vectors at x_0 in which the rate of change of F is 0, as these are the directions in which infinitesimal displacements will keep us on the level set to which x_0 belongs. The gradient on the other hand is orthogonal to these and defines the direction in which F is increasing the most, indeed:

$$\langle \vec{\nabla} f(x_0), \vec{h} \rangle = \left\| \vec{\nabla} f(x_0) \right\| \|h\| \cos\theta,$$

where θ is the geometric angle between the two vectors. This is maximal when $\cos\theta = 1$ i.e. $\vec{h}$ is parallel and points the same way as $\vec{\nabla} f(x_0)$. One can also observe from this that:

$$\left\| \vec{\nabla} F(x_0) \right\| = \max_{\vec{h} \neq \vec{0}} \frac{\left| \langle \vec{\nabla} F(x_0), \vec{h} \rangle \right|}{\left\| \vec{h} \right\|} = \max_{\vec{h} \neq \vec{0}} \frac{\left| dF(x_0)(\vec{h}) \right|}{\left\| \vec{h} \right\|}.$$

1.11 The inverse function and implicit function theorems

Above, we have introduced level sets as a means of understanding a scalar field, however, they are geometric objects of interest in their own right. It can often happen that we are interested in subsets of $\mathbb{R}^n$ that are defined by one or more equations:

$$\begin{cases} F_1(x_1, \ldots, x_n) = 0, \\ \qquad \vdots \\ F_m(x_1, \ldots, x_n) = 0, \end{cases}$$

where $F_1, \ldots, F_m$ are real-valued differential functions. Observe that once more we can consider this as the equation:

$$F(x) = 0$$

where $F : \mathbb{R}^n \to \mathbb{R}^m$ and $F(x) = (F_1(x), \ldots, F_m(x))$. For instance, the n-sphere is defined by the equation:

$$x_1^2 + \cdots + x_n^2 - 1 = 0$$

There are two very important theorems that help us understand these sets, and that can be thought of as a generalisation of well known facts in the theory of *linear* equations.

1.11.1 The inverse function theorem

Let us first consider a linear function $F : \mathbb{R}^n \to \mathbb{R}^n$. If $c \in \mathbb{R}^n$, then the equation:

$$F(x) = c,$$

is the same as n-linear equations with n-unknowns. Indeed, F can be represented by a matrix $A = (a_{ij})$ and we can rewrite the equation:

$$A\mathbf{x} = \mathbf{c} \Leftrightarrow \begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = c_1, \\ a_{21}x_1 + \cdots + a_{2n}x_n = c_2, \\ \vdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n = c_n. \end{cases}$$

Now, we know that this can be solved, and has a unique solution, when the matrix A is invertible. The first theorem we will state generalises this to the case where F is differentiable, but the price we have to pay is that we will generically only be able to solve it *locally*, in a neighbourhood of a given point x_0. To make the theorem statement as intuitive as possible let us introduce the following notion:

Definition 1.19: Neighbourhoods

Let $(E, \|\cdot\|)$ be a normed vector space and $x_0 \in E$ then
- a set W *neighbourhood* is called a **neighbourhood of** x_0 if and only if it contains *an* open set U such that $x_0 \in U \subset W$.
- An *open* neighbourhood is an open set such that $x_0 \in U$.

Example 1.19.
- An open set is a neighbourhood of any of its points.
- The closed unit ball $\bar{B}(0,1) = \{(x,y) \in \mathbb{R}^2, x^2 + y^2 \leqslant 1\}$ is a neighbourhood of $(0,0)$ but not of $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$.

$\diamond$

Theorem 1.13: The inverse function theorem

Let $F : U \subset \mathbb{R}^n \to \mathbb{R}^n$ be a differentiable function with *continuous partial derivatives* and $x_0 \in U$.
Assume that the linear transformation dF_{x_0} is *bijective*, or, equivalently, that the Jacobian matrix Jac $f(x_0)$ is an invertible matrix.
Then, there are open neighbourhoods $U_1 \subset U$ and U_2 of x_0 and $F(x_0)$ respectively such that the restriction of F:

$$F|_{U_1} : U_1 \to U_2$$

is invertible (e.g bijective), and the inverse map $F^{-1} : U_2 \to U_1$ is differentiable with continuous derivatives.

The contents of the theorem is summarised schematically in Figure 1.8.

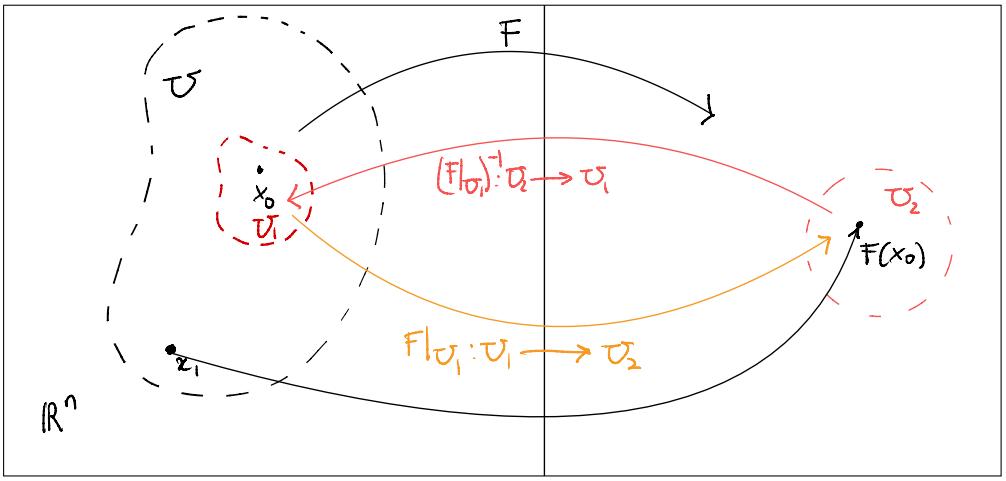


Figure 1.8: Under the hypotheses of the inverse function theorem. One can invert the function F between open neighbourhoods U_1 of x_0 and U_2 of $F(x_0)$ respectively. However, the function may fail to be invertible globally: here x_1 is mapped to $F(x_0)$ too but it is not in U_1.

Remark 1.11.1. The theorem holds if F is a function defined on an open set of a Banach space, into another Banach space.

A differentiable function with *continuous partial derivatives* is said to be of class $\mathcal{C}^1$; we speak of $\mathcal{C}^1$ functions.

We unfortunately do not have the tools to prove this theorem; the proof relies on the completeness of $\mathbb{R}^n$ (every Cauchy sequence converges), and a fixed point type argument.

← End Lecture 9

Remark 1.11.2. (not mentioned in class, for your information) The local inverse function theorem above can be used to prove a global inverse function theorem:

Theorem 1.14: Global inverse function theorem

Let $F : U \subset \mathbb{R}^n \to F(U) \subset \mathbb{R}^m$ be an *injective* (one-to-one) $\mathcal{C}^1$ function, such that dF_{x_0} is invertible at each point $x_0 \in U$, then:

- $F(U)$ is open,
- $F : U \to F(U)$ is invertible and the inverse function F^{-1} is a $\mathcal{C}^1$ function.

Proof. We begin by proving that $F(U)$ is open. Let $y_0 = f(x_0)$, then since dF_{x_0} is invertible we can apply the inverse function theorem and find open neighbourhoods U_1 and U_2 of x_0 and $f(x_0)$ respectively such that $F|_{U_1} : U_1 \to U_2$ is invertible and in particular it is surjective (onto) so the open set U_2 is completely contained $F(U)$. This shows that $F(U)$ is open.

Now, the second statement is quite straightforward to establish: since $F : U \to F(U)$ is bijective, F^{-1} exists and is unique, hence the local inverses must coincide with restrictions of F^{-1}, differentiability is a local property, hence, F^{-1} is a $\mathcal{C}^1$ function. □

Start Lecture 10→

1.11.2 A special case of the implicit function theorem

The inverse function theorem is of great theoretical importance, and has numerous applications throughout mathematics. It is in fact equivalent to another theorem, of equal importance, known as the implicit function theorem. Let us first study the linear version, consider now the case of a linear system of m equations with n unknowns and assume $n > m$; so we have more unknowns than equations. We can again phrase this as an equation $F(x) = c$ where F is a linear map $\mathbb{R}^n \to \mathbb{R}^m$, which can again be represented by a matrix $A = (a_{ij})$:

$$F(x) = c \Leftrightarrow A\mathbf{x} = \mathbf{c} \Leftrightarrow \begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = c_1, \\ a_{21}x_1 + \cdots + a_{2n}x_n = c_2, \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = c_m. \end{cases}$$

In this case, if the matrix A is of full rank, then one can show that m of the unknowns can be expressed in terms of the $n - m$ others.

Let us consider an explicit example where $m = 1$, $n = 3$ and consider an equation:

$$\alpha x + \beta y + \gamma z = c,$$

then the full rank condition simply means in this case that $(\alpha, \beta, \gamma) \neq (0,0,0)$. Assume that $\gamma \neq 0$, then:

$$z = \frac{c}{\gamma} - \frac{\alpha}{\gamma}x - \frac{\beta}{\gamma}y.$$

This principle generalises, locally, to the case where F is a $\mathcal{C}^1$ function. The general theorem requires some notation to make it readable, so we will first state a compact version, in the special case where $n = 3, m = 1$ which is of use in applications.

Theorem 1.15: The implicit function theorem when $n = 3, m = 1$

Let $F : U \subset \mathbb{R}^3 \to \mathbb{R}$ be a $\mathcal{C}^1$ function, and $c \in \mathbb{R}^3$. Let $(x_0, y_0, z_0) \in U$ be such that $F(x_0, y_0, z_0) = c$ and assume that:

$$\frac{\partial F}{\partial z}(x_0, y_0, z_0) \neq 0,$$

then there is an open neighbourhood U_1 of (x_0, y_0, z_0), an open neighbourhood U_2 of $(x_0, y_0) \in \mathbb{R}^2$ and a $\mathcal{C}^1$ function $\varphi : U_2 \to \mathbb{R}$ such that we have the following equivalence:

$$\begin{cases} (x,y,z) \in U_1, \\ F(x,y,z) = c, \end{cases} \Leftrightarrow \begin{cases} z = \varphi(x,y), \\ (x,y) \in U_2. \end{cases}$$

Furthermore, for $(x, y) \in V_2$:

$$\frac{\partial \varphi}{\partial x}(x,y) = -\frac{\frac{\partial F}{\partial x}(x,y,\varphi(x,y))}{\frac{\partial F}{\partial z}(x,y,\varphi(x,y))}, \quad \frac{\partial \varphi}{\partial y}(x,y) = -\frac{\frac{\partial F}{\partial y}(x,y,\varphi(x,y))}{\frac{\partial F}{\partial z}(x,y,\varphi(x,y))}.$$

Remark 1.11.3. It is not actually important that it be the z-variable, it can be any of the three, which will then be expressed as a function of the other two.

The implicit function tells us that, locally, the set of solutions to the equation: $F(x,y,z) = c$ is given by a graph !

Proof. The idea of the proof is to complete the system by adding some extra trivial equations and apply the inverse function theorem. We introduce:

$$\begin{array}{rccc} G: & U \subset \mathbb{R}^3 & \longrightarrow & \mathbb{R}^2 \times \mathbb{R} = \mathbb{R}^3 \\ & (x,y,z) & \longmapsto & (x,y,F(x,y,z)) \end{array}.$$

This amounts to replacing the unique equation $F(x,y,z) = c$ by the system:

$$\begin{cases} x = x, \\ y = y, \\ F(x,y,z) = c. \end{cases}$$

Since F is differentiable, G is differentiable by composition of differentiable maps. Let us compute the Jacobian matrix of G at (x_0, y_0, z_0)

$$\text{Jac } G(x_0,y_0,z_0) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{\partial F}{\partial x}(x_0,y_0,z_0) & \frac{\partial F}{\partial y}(x_0,y_0,z_0) & \frac{\partial F}{\partial z}(x_0,y_0,z_0) \end{pmatrix}.$$

Since $\det(\text{Jac } G(x_0,y_0,z_0)) = \frac{\partial F}{\partial z}(x_0,y_0,z_0) \neq 0$, the inverse function theorem applies. There is therefore some open neighbourhood U_1 of (x_0,y_0,z_0) and an open neighbourhood $U_2 \times I \subset \mathbb{R}^2 \times \mathbb{R}$, ($U_2$ is open in $\mathbb{R}^2$ and I is an open interval, overall this is an open set in $\mathbb{R}^3$) such that, restricted to U_1, G is invertible with $\mathcal{C}^1$ inverse. Let H denote this inverse function. H is necessarily of the form:

$$H(x,y,z) = (x, y, \tilde{\varphi}(x,y,z)).$$

Set $\varphi(x,y) = \tilde{\varphi}(x,y,c)$, $(x,y) \in U_2$, then ϕ is differentiable and for any $(x,y) \in U_2$:

$$(x,y,c) = G(H(x,y,c)) = (x,y,F(x,y,\varphi(x,y))),$$

which proves that $F(x,y,\varphi(x,y)) = c$. This shows that if $z = \varphi(x,y), (x,y) \in U_2$ then $F(x,y,\varphi(x,y)) = c$ as claimed.

Conversely, if we assume that $(x,y,z) \in U_1$, and $F(x,y,z) = c$ then $G(x,y,z) = (x,y,c) \in U_2 \times I$ but then: $H(G(x,y,z)) = H(x,y,c) = (x,y,\varphi(x,y))$ so that $z = \varphi(x,y)$.

This proves that we have found all solutions to the problem locally as claimed. It remains to check the formula for the partial derivatives, but this follows from the chain rule. For instance, we can consider the composition:

$$(x,y) \in U_2 \overset{g}{\mapsto} (x,y,\varphi(x,y)) \overset{F}{\mapsto} F(x,y,\varphi(x,y)) = c$$

The result of this composition is a constant scalar function so its partial derivatives vanish, by the chain rule we then have:

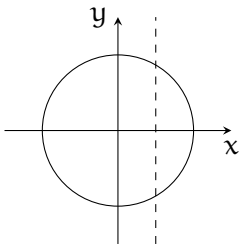
$$(0,0) = \left(\frac{\partial F}{\partial x}(x,y,\varphi(x,y)) \quad \frac{\partial F}{\partial y}(x,y,\varphi(x,y)) \quad \frac{\partial F}{\partial z}(x,y,\varphi(x,y))\right) \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \frac{\partial \varphi}{\partial x}(x,y) & \frac{\partial \varphi}{\partial y}(x,y) \end{pmatrix}.$$

From which the conclusion follows by carrying out the computation. □

Example 1.20. Let's work the theorem out on a standard example. Consider the equation:

$$x^2 + y^2 = 1$$

in $\mathbb{R}^2$, it is well know that the solutions are the points of the unit circle.



The circle cannot globally be the graph $\{y = f(x)\}$ for some function f, as we can see each x would have to be mapped simultaneously to two different values. Nevertheless, if we restrict to the (open) band, $\{(x, y) \in \mathbb{R}^2, -1 < x < 1, y > 0\}$, which is an open neighbourhood of the point $(0, 1)$ then clearly the solutions are given by:

$$y = \sqrt{1 - x^2}.$$

Now, if set $F(x, y) = x^2 + y^2$, then it is clear that $\frac{\partial F}{\partial y}(0, 1) = 2 \neq 0$, so indeed, the implicit function theorem says that we can express y as a function of x in some neighbourhood of $(0, 1)$.

Suppose now that we had considered the point $(1, 0)$, in this case $\frac{\partial F}{\partial y}(1, 0) = 0$ so the implicit function theorem does not apply. Instead however, $\frac{\partial F}{\partial x}(1, 0) = 2 \neq 0$ which means that in a neighbourhood of $(0, 1)$ we can express x as a function of y. Explicitly we see that we will write:

$$x = \sqrt{1 - y^2}.$$

$\diamond$

1.11.3 Example of how the the implicit function theorem can be used in natural sciences.

Unlike in the previous example, in general, we cannot write down the function φ. However, this is not a problem as we can calculate its derivatives from F; the theorem is generally used in this way.

To illustrate this further we shall consider a situation frequently encountered in applications.

Often, it can be that a system is described by variables x, y, z, which are not independent, i.e. there is some relation:

$$F(x, y, z) = 0,$$

between them where $F : \mathbb{R}^3 \to \mathbb{R}$. Assuming for instance that none of the partial derivatives vanish at some point $(x_0, y_0, z_0) \in R^3$ in the parameter space of the system (so $F(x_0, y_0, z_0) = 0$), then according to the implicit function theorem we can use locally around (x_0, y_0, z_0) *any two* of the variables to express the other. Instead of writing φ however we tend to abuse notation and write things like:

$$z(x, y), \quad x(y, z), \quad y(x, z).$$

To help keep track of things it also common to introduce for instance, notation:

$$\left(\frac{\partial z}{\partial x}\right)_y.$$

To indicate that this is the partial derivative of z, viewed as a function of x and y, in the x direction at fixed y. For instance, in a thermodynamics textbook you might see things like:

$$\left(\frac{\partial P}{\partial T}\right)_V,$$

which would mean the partial derivative with respect to T of the pressure P viewed as a function of T and V.

When working this always assume that (x, y, z) satisfies the equation $F(x, y, z)$ so that if we want to calculate the partial derivative, say, we can write:

$$\left(\frac{\partial z}{\partial x}\right)_y (x, y) = -\frac{\frac{\partial F}{\partial x}(x, y, z)}{\frac{\partial F}{\partial z}(x, y, z)}.$$

This is consistent, because, according to the theorem, $z = \varphi(x, y)$ when $F(x, y, z) = 0$.

If we compute in a similar fashion $\left(\frac{\partial x}{\partial y}\right)_z$ and $\left(\frac{\partial y}{\partial z}\right)_x$ then the implicit function theorem shows that:

$$\left(\frac{\partial z}{\partial x}\right)_y \left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x = -1,$$

Which is an identity you will also find in thermodynamics or chemistry textbooks.

1.11.4 The general implicit function theorem

For this we will use the more general notion of partial derivative discussed in Section 1.7.3. If you do not want to read this you should skip ahead to the theorem statement at the end of the section.

We can state the general implicit function theorem as follows:

Theorem 1.16: Implicit function theorem

Let $(E_1, \|\cdot\|_{E_1}), (E_2, \|\cdot\|_{E_2}), (V, \|\cdot\|_V)$ be Banach spaces, $F : U \subset E_1 \times E_2 \to V$ a $\mathcal{C}^1$ map. Let $(a_1, a_2) \in E_1 \times E_2$ and assume that the partial derivative: $\frac{\partial F}{\partial y}(a_1, a_2)$ is an invertible linear map.[a]
Then there are open neighbourhoods U_1 of $(a_1, a_2) \in E_1 \times E_2$, U_2 of $a_1 \in E_1$, U_3 of $F(a_1, a_2) \in V$, and a C^1 map $\tilde{\varphi} : U_2 \times U_3 \to E_2$ such that for all $c \in U_3$:

$$\begin{cases} (x, y) \in U_1, \\ F(x, y) = c, \end{cases} \Leftrightarrow \begin{cases} y = \tilde{\varphi}(x, c), \\ x \in U_2. \end{cases}$$

Furthermore, for fixed $c \in U_3$, if we define $\varphi(x) = \tilde{\varphi}(x, c)$ then:

$$d\varphi_x = -\left[\frac{\partial F}{\partial y}(x, \varphi(x, y))\right]^{-1} \circ \frac{\partial F}{\partial x}(x, \varphi(x, y)).$$

[a] I should say with continuous inverse but this is in fact automatic for Banach spaces.

Remark 1.11.4. Observe that on the last line, this is an equation on linear maps.

The proof follows the same idea as in the case we studied before, so we will not reproduce it here, but will instead explicitly translate this into a compact version in finite dimensions.

Let us start with a function $F : U \subset \mathbb{R}^n = \mathbb{R}^{m-n} \times \mathbb{R}^m \to \mathbb{R}^m$. We should think of this as reorganising the n-real variables $(X_1, \ldots, X_n)$ into two groups: $x_1, \ldots, x_{m-n}$, $y_1, \ldots, y_m$, which we consider as vector-valued variables:

$$x = (x_1, \ldots, x_{m-n}), \qquad y = (y_1, \ldots, y_m).$$

This should be done in such a way that:

$$\frac{\partial F}{\partial y}(x_0, y_0) \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m),$$

is invertible at some point $(x_0, y_0) \in \mathbb{R}^{n-m} \times \mathbb{R}^m$ where $F(x_0, y_0) = c \in \mathbb{R}^m$. This partial derivative is a linear map and can be represented by a square matrix of size $m \times m$, writing: $F(x, y) = (F_1(x, y), \ldots, F_m(x, y))$ it is given at $(x_0, y_0) \in \mathbb{R}^{m-n} \times \mathbb{R}^m = \mathbb{R}^n$ by:

$$\begin{pmatrix} \frac{\partial F_1}{\partial y_1}(x_0, y_0) & \cdots & \frac{\partial F_1}{\partial y_m}(x_0, y_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial F_1}{\partial y_m}(x_0, y_0) & \cdots & \frac{\partial F_m}{\partial y_m}(x_0, y_0). \end{pmatrix}$$

This is a square submatrix of the full Jacobian of F; the assumption of the theorem is therefore that this matrix is invertible i.e. its determinant is non-vanishing.

The theorem then says that we can find a $\mathcal{C}^1$ function $\varphi : U_2 \to \mathbb{R}^m$, defined on some open neighbourhood U_2 of x_0, such that in a neighbourhood U_1 of (x_0, y_0) all the solutions of the equation $F(x, y) = c$ are given by:

$$y = \varphi(x).$$

Translating the final equation in terms of Jacobian matrices we therefore have at any solution point (x, y) (i.e. $F(x, y) = c$ which means $y = \varphi(x)$):

$$\mathbf{Jac}\ \varphi(x) = -\begin{pmatrix} \frac{\partial F_1}{\partial y_1}(x,y) & \cdots & \frac{\partial F_1}{\partial y_m}(x,y) \\ \vdots & \ddots & \vdots \\ \frac{\partial F_1}{\partial y_m}(x,y) & \cdots & \frac{\partial F_m}{\partial y_m}(x,y). \end{pmatrix}^{-1} \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(x,y) & \cdots & \frac{\partial F_1}{\partial x_{n-m}}(x,y) \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial x_1}(x,y) & \cdots & \frac{\partial F_m}{\partial x_{n-m}}(x,y) \end{pmatrix}.$$

ndatory reading

This is summarised in the next theorem:

Theorem 1.17: Implicit function theorem, several equations

Let $F : U \subset \mathbb{R}^{n-m} \times \mathbb{R}^m \to \mathbb{R}^m$ be a $\mathcal{C}^1$ function, assume that:

$$F(\vec{x}, \vec{y}) = (F_1(\vec{x}, \vec{y}), \ldots, F_m(\vec{x}, \vec{y})),$$

let $(\vec{x}_0, \vec{y}_0) \in U$ and suppose that:

$$F(\vec{x}_0, \vec{y}_0) = c \in \mathbb{R}^m.$$

Write:

$$\left[\frac{\partial F}{\partial \vec{y}}(\vec{x}_0, \vec{y}_0)\right] = \begin{pmatrix} \frac{\partial F_1}{\partial y_1}(\vec{x}_0, \vec{y}_0) & \cdots & \frac{\partial F_1}{\partial y_m}(\vec{x}_0, \vec{y}_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial F_1}{\partial y_m}(\vec{x}_0, \vec{y}_0) & \cdots & \frac{\partial F_m}{\partial y_m}(\vec{x}_0, \vec{y}_0), \end{pmatrix}$$

the partial Jacobian matrix with respect to $\vec{y}$ (i.e. the last m variables).
Then assuming that this matrix is *invertible*, there is an open neighbourhood U_1 of $(\vec{x}, \vec{y})$, an open neighbourhood U_2 of $\vec{x}$ and a $\mathcal{C}^1$ function $\varphi : U_2 \to \mathbb{R}^m$ such that:

$$\begin{cases} (\vec{x}, \vec{y}) \in U_1, \\ F(\vec{x}, \vec{y}) = c, \end{cases} \Leftrightarrow \begin{cases} \vec{y} = \varphi(\vec{x}), \\ \vec{x} \in U_2. \end{cases}$$

Furthermore at each point $(\vec{x}, \vec{y} = \varphi(\vec{x}))$

$$\text{Jac}\, \varphi(\vec{x}) = -\begin{pmatrix} \frac{\partial F_1}{\partial y_1}(\vec{x}, \vec{y}) & \cdots & \frac{\partial F_1}{\partial y_m}(\vec{x}, \vec{y}) \\ \vdots & \ddots & \vdots \\ \frac{\partial F_1}{\partial y_m}(\vec{x}, \vec{y}) & \cdots & \frac{\partial F_m}{\partial y_m}(\vec{x}, \vec{y}). \end{pmatrix}^{-1} \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(\vec{x}, \vec{y}) & \cdots & \frac{\partial F_1}{\partial x_{n-m}}(\vec{x}, \vec{y}) \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial x_1}(\vec{x}, \vec{y}) & \cdots & \frac{\partial F_m}{\partial x_{n-m}}(\vec{x}, \vec{y}) \end{pmatrix}.$$

Example 1.21. Let: $F : \mathbb{R}^3 \to \mathbb{R}^2$ be defined by:

$$F(x, y, z) = (x^2 + y^2, 2x + y - z)$$

let $c = (1, 3)$ then:

$$F(x, y, z) = c \Leftrightarrow \begin{cases} x^2 + y^2 = 1 \\ 2x + y - z = 4 \end{cases}$$

which describes the intersection of a cylinder with a plane in $\mathbb{R}^3$.

Observe that, if the implicit function theorem can be applied then we will be able to express two of the variables in function of $3 - 2 = 1$ of the others: hence we recover the fact that if non-empty, it is a curve. Now, $(0, 1, -3)$ belongs to the solutions set. Let us group together (y, z) into a vector valued variable $\vec{y}$ and determine the partial

Jacobian matrix:

$$\left[\frac{\partial F}{\partial \vec{y}}\right](0,1,-3) = \begin{pmatrix} 2 & 0 \\ 1 & -1 \end{pmatrix},$$

which is invertible, therefore, locally $(y,z) = \phi(x)$

$$\begin{pmatrix} \frac{dy}{dx}(0) \\ \frac{dz}{dx}(0) \end{pmatrix} = -\begin{pmatrix} \frac{1}{2} & 0 \\ \frac{1}{2} & -1 \end{pmatrix}\begin{pmatrix} 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}.$$

In this case we can solve the equations explicitly near the point $(0,1,-3)$ and recover the result:

$$z = 2x + y - 3, \quad y = \sqrt{1-x^2}.$$

◇

1.11.5 Application of the implicit function theorem to the geometry of level hypersurfaces

We can also use the implicit function theorem to finish the proof of the fact that if $\mathscr{L} = \{x \in \mathbb{R}^n, F(x) = 0\}$ where F is a $\mathcal{C}^1$ function, then:

$$T_{x_0}\mathscr{L} = \{h \in \mathbb{R}^n, \langle \vec{\nabla} f(x_0), h\rangle = 0\},$$

when x_0 is a regular point. We have already shown that:

$$T_{x_0}\mathscr{L} \subset \{h \in \mathbb{R}^n, \langle \vec{\nabla} f(x_0), h\rangle = 0\}.$$

It remains to establish the other inclusion: this will also prove (for free) that $T_{x_0}\mathscr{L}$ is a vector space !

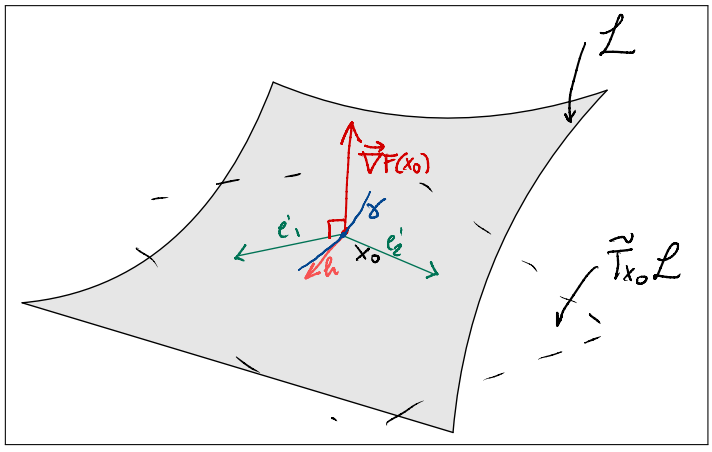


Figure 1.9: Constructing a basis adapted to the problem

The idea is to work in an orthonormal basis $\mathscr{B}' = (e'_1, \dots, e'_n)$ of $\mathbb{R}^n$ in which the last vector is $e'_n = \frac{\overrightarrow{\nabla} f(x_0)}{\left\|\overrightarrow{\nabla} f(x_0)\right\|} \neq 0$; this is always possible (see Figure 1.9). If we write the coordinates of a vector x in this basis $(x'_1, \dots, x'_n)$, then:

$$\frac{\partial F}{\partial x'_n}(x_0) = \langle \overrightarrow{\nabla} F(x_0), e'_n \rangle = \left\| \overrightarrow{\nabla} F(x_0) \right\| \neq 0.$$

It follows that, locally near x_0, the level set is defined by the equation:

$$x'_n = \varphi(x'_1, \dots, x'_{n-1}).$$

Now if h is orthogonal to the gradient vector at x_0 then, $h = h_1 e'_1 + \dots h_{n-1} e'_{n-1}$ and, assuming the coordinates of x_0 are given by $(a'_1, \dots, a'_n)$ in the basis $\mathscr{B}'$ we can define, for small enough t, the curve:

$$\gamma(t) = \sum_{i=1}^{n-1} \underbrace{(a'_i + t h'_i)}_{x'_i(t)} e'_i + \underbrace{\varphi(a'_1 + t h'_1, \dots, a'_{n-1} + t h'_{n-1})}_{x'_n(t)} e'_n.$$

By construction, it is drawn on $\mathscr{L}$ (indeed the coordinate satisfy the equation: $x'_n(t) = \varphi(x'_1, \dots, x'_{n-1})$ for all t) and the tangent vector at $t = 0$ is given by:

$$\gamma'(0) = \sum_{i=1}^{n-1} h'_i e'_i - \frac{1}{\left\|\overrightarrow{\nabla} f(x_0)\right\|} \sum_{i=1}^{n-1} \frac{\partial F}{\partial x'_i}(x_0) h'_i e'_n$$

but $\frac{\partial F}{\partial x'_i}(x_0) = \langle \overrightarrow{\nabla} F(x_0), e'_i \rangle = 0$ for $1 \leqslant i \leqslant n-1$, by definition, hence:

$$\gamma'(0) = h.$$

Remark 1.11.5. Recall that the *partial* derivatives are defined to be the *directional derivatives in the direction of the basis vectors*:

$$\frac{\partial F}{\partial x'_i}(x_0) = D_{e'_i} F(x_0),$$

and when F is differentiable this is linked to the derivative by:

$$dF_{x_0}(e'_i) = D_{e'_i} F(x_0).$$

Finally when F is a scalar field the gradient is *defined* by the property:

$$df_{x_0}(h) = \langle \overrightarrow{F}(x_0), h \rangle.$$

← End Lecture 10

1.12 Applications to the search for extrema

← Start Lecture 11

From 1-variable calculus, you know that the derivative can help us find the maximum and minimum values of numerical functions $f : \mathbb{R} \to \mathbb{R}$. The basic criterion is that these values are reached at *critical points* where the derivative vanishes. However, as the function $x \mapsto x^3$ shows, not all critical points correspond to extrema. Furthermore, the techniques of differential calculus are intrinsically local, so it will only be able to detect x where *local* extrema are reached.

1.12.1 The first derivative test

Let us first generalise this criterion to scalar fields on a normed vector space:

Theorem 1.18: First derivative criterion

Let $f : U \subset E \to \mathbb{R}$ be a real-valued differentiable function, and $x_0 \in U$, then *if* f reaches a local extremum at x_0 *then* $\overrightarrow{\nabla} f(x_0) = 0$.

Proof. Let $h \in E$ and let $I = (-\varepsilon, \varepsilon)$ where ϵ is small enough such that $x_0 + th \in U$ for all $|t| < \varepsilon$. Then, the function $\varphi : t \mapsto f(x_0 + th)$ reaches a local extremum at $t = 0$, and by definition of the directional derivative:

$$\varphi'(0) = \langle \overrightarrow{\nabla} f(x_0), h \rangle.$$

Therefore, by the usual 1-dimensional criterion, $0 = \phi'(0) = \langle \overrightarrow{\nabla} f(x_0), h \rangle = 0$ for any $h \in E$, hence:

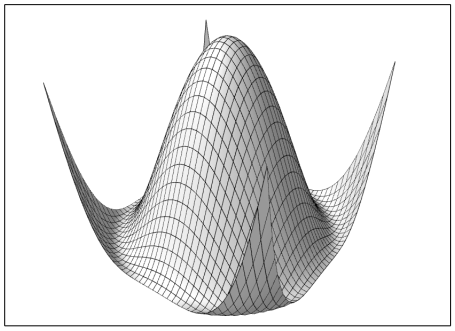
$$\overrightarrow{\nabla} f(x_0) = 0.$$

□

Example 1.22. The Mexican hat potential:

$$H(x, y) = -3(x^2 + y^2) + (x^2 + y^2)^2,$$

has a *local* maximum at $(0, 0)$.



The gradient is given at any point (x, y) by:

$$\overrightarrow{\nabla} f(x,y) = \begin{pmatrix} -6x + 4x(x^2+y^2) \\ -6y + 4y(x^2+y^2) \end{pmatrix},$$

and it is indeed true that:

$$\overrightarrow{\nabla} f(0,0) = 0.$$

However, it is not a *global* maximum as $H(x,y) \to +\infty$ when $\|(x,y)\| \to \infty$. ◇

1.12.2 First derivative test on a constraint: Lagrange multipliers

The above criterion gives us a method for searching for local extrema on an *open set*. However, we are quite often more interested in local extrema on sets defined by certain constraints. These constraints can be written as a set of equations like those we studied in the previous section.

For instance, they might be expressed in the form $F(x) = 0$ and $F : U \subset \mathbb{R}^n \to \mathbb{R}^m$:

$$F(x) = 0 \Leftrightarrow \begin{cases} F_1(x) = 0 \\ \quad \vdots \\ F_m(x) = 0. \end{cases}$$

If F is a $\mathcal{C}^1$ function we have now developed a number of tools and concepts to help us deal with this.

Let us first consider the case $m = 1$, i.e. there is only one constraint. In this case, the constraint corresponds to a level set $\mathscr{L}$ of a scalar function, and we defined the tangent vector space $T_{x_0}\mathscr{L}$ to be, intuitively, the set of directions in which we can make an infinitesimal displacement from x_0 without leaving $\mathscr{L}$; this was characterised as the set of vectors that are orthogonal to gradient vector $\overrightarrow{\nabla} F(x_0)$. Carrying out the same reasoning as before, we see that if a differentiable function $f : U \subset \mathbb{R}^n \to \mathbb{R}$ attains an extrema, subject to the constraint defined by $\mathscr{L}$, at some point $x_0 \in \mathscr{L}$ then:

$$\forall h \in T_{x_0}\mathscr{L}, \quad \langle \overrightarrow{\nabla} f(x_0), h \rangle = 0.$$

We shall call such a point x_0 a *critical point* of $f|_{\mathscr{L}}$ (the restriction of f to $\mathscr{L}$.

Contrary to when we were seeking to optimise the function f on the open set *this does not imply* that $\overrightarrow{\nabla} f(x_0) = 0$! Instead, it just states that it must be orthogonal to all vectors in the vector tangent space, and therefore *parallel* to the gradient of F, in other words:

$$\exists \lambda \in \mathbb{R}, \overrightarrow{\nabla} f(x_0) = \lambda \overrightarrow{\nabla} F(x_0).$$

Such a number is known as a *Lagrange multiplier*.

To illustrate the method in this case, let's consider the following example:

Example 1.23. Let us seek the extrema of $f(x,y) = \sqrt{x^2+y^2}$ subject to the constraint $x^2 + 3y^2 = 1$, i.e. (x,y) lies on an ellipse represented in Figure 1.10. The

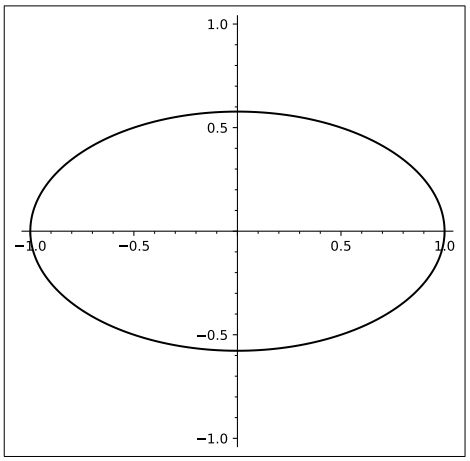


Figure 1.10: Ellipse : $x^2 + 3y^2 = 1$

function f is just value of the radius r from the origin, and we can see that it has its extrema on the coordinate axes. Let us check explicitly that the theory recovers this. Setting $F(x,y) = x^2 + 3y^2 - 1$ then the constraint is the level set $\mathscr{L} = \{(x,y) \in \mathbb{R}^2, F(x,y) = 0\}$. According to the above theory, we need to look for $(x,y) \in \mathbb{R}^2$ and $\lambda \in \mathbb{R}$ that solves the system:

$$\begin{cases} x^2 + 3y^2 = 1 \\ \frac{x}{\sqrt{x^2+y^2}} = 2\lambda x \\ \frac{y}{\sqrt{x^2+y^2}} = 6\lambda y \end{cases}$$

The first equation is just the constraint and the second and third express the condition:

$$\overrightarrow{\nabla} f(x,y) = \lambda \overrightarrow{\nabla} F(x,y).$$

Now, we can observe that if $x \neq 0$ and $y \neq 0$, then we can divide by x and y in the last two equations and would find $2\lambda = 6\lambda \Rightarrow \lambda = 0$, which in turn implies $x = y = 0$ and so is inconsistent. The only solutions are then at points where either $x = 0$ or $y = 0$.

It is straightforward to check that the only solutions are:

$$x = 0, y = \pm\frac{1}{\sqrt{3}}, \lambda = \frac{1}{2\sqrt{3}}, \quad \text{or} \quad x = \pm 1, y = 0, \lambda = \frac{1}{2}.$$

Which are precisely the intersections of the ellipse with the axes.

Observe that these are not critical points of f on its open domain, since $\lambda \neq 0$. ◇

We will now see how to generalise this to problems with several constraints. We already have the general ideas, and the notion of tangent (vector) space extends:

Definition 1.20

Let $F : U \subset \mathbb{R}^n \to \mathbb{R}^m$ be a $\mathcal{C}^1$ function, $c \in \mathbb{R}^m$ and set:

$$\mathscr{L} = \{x \in \mathbb{R}^n, F(x) = c\}.$$

$x_0 \in U$ is said to be regular if dF_{x_0} is surjective or, equivalently, Jac $f(x_0)$ has full rank.
The tangent vector space to a regular point $x_0 \in \mathscr{L}$ is defined to be the set of tangent vectors at x_0 to curves on $\mathscr{L}$ passing through x_0. That is:

$$v \in T_{x_0}\mathscr{L} \Leftrightarrow \exists\ \gamma : \begin{array}{ccc} (-\varepsilon, \varepsilon) & \longrightarrow & \mathscr{L} \subset \mathbb{R}^n \\ t & \longmapsto & \gamma(t) \end{array}, \begin{cases} \gamma(0) = x_0, \\ \gamma'(0) = v. \end{cases}$$

Now, as in the case $m = 1$, if γ is curve defined as in the above definition, then,differentiating the identity:

$$F(\gamma(t)) = c,$$

using the chain rule, it follows that, at $t = 0$.

$$dF_{x_0}(\gamma'(0)) = \text{Jac } F(x_0)\gamma'(0) = 0.$$

So:

$$T_{x_0}\mathscr{L} \subset \{\, \mathbf{h} \in \mathbb{R}^n, \text{Jac } F(x_0)\mathbf{h} = 0\}.$$

Remark 1.12.1.

- If $u : E \to F$ is a linear map, then the kernel or nullspace of u is:

$$\ker u = \{x \in E, u(x) = 0\}.$$

- By extension, A is a matrix then the set of column vectors $X \in \mathbb{R}^n$ such that $AX = 0$ is referred to as the kernel or the nullspace of A.

Adapting the argument in Section 1.11.5, one can show that this is in fact an equality:

Proposition 1.7

Let $F : U \subset \mathbb{R}^n \to \mathbb{R}^m$ be a $\mathcal{C}^1$ function, $c \in \mathbb{R}^m$ and set:

$$\mathscr{L} = \{x \in \mathbb{R}^n, F(x) = c\}.$$

Let $x_0 \in \mathscr{L}$ be a regular point of F then:

$$T_{x_0}\mathscr{L} = \{\, \mathbf{h} \in \mathbb{R}^n, \text{Jac } F(x_0)\mathbf{h} = 0\}.$$

Writing: $F(x) = (F_1(x), \ldots, F_m(x)), x \in U$ we can observe that:

$$\text{Jac } F(x_0)\mathbf{h} = \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(x_0) & \frac{\partial F_1}{\partial x_2}(x_0) & \ldots & \frac{\partial F_1}{\partial x_n}(x_0) \\ \frac{\partial F_2}{\partial x_1}(x_0) & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \frac{\partial F_m}{\partial x_1}(x_0) & \ldots & \ldots & \frac{\partial F_m}{\partial x_n}(x_0) \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ \vdots \\ h_n \end{pmatrix},$$

vanishes if and only if the product of $\mathbf{h}$ with each of the rows of the matrice vanishes. However, the rows are exactly the Jacobian matrices of the scalar fields $F_1, \ldots, F_m$. Observe that the full rank condition implies that these rows are linearly independent. Hence, recalling that for the scalar fields F_i we have the string of identities:

$$dF_{i\,x_0}(h) = \text{Jac } F_i(x_0)\mathbf{h} = \langle \overrightarrow{\nabla} F_i, h\rangle,$$

then:

$$h \in T_{x_0}\mathscr{L} \Leftrightarrow \forall i \in \{1, \ldots, m\}, \quad \langle \overrightarrow{\nabla} F_i(x_0), h\rangle = 0.$$

Now if f attains a local extrema at x_0 on the constraint $\mathscr{L}$, i.e. $f|_{\mathscr{L}}$ admits a local extrema there, then we will conclude that:

$$\forall h \in T_{x_0}\mathscr{L}, df_{x_0}(h) = \langle \overrightarrow{\nabla} f(x_0), h\rangle = 0.$$

As before, this does *not* imply that: $\overrightarrow{\nabla} f(x_0) = 0$! Instead it must be in the orthogonal subset to $T_{x_0}\mathscr{L}$, which as we have seen above is spanned (generated) by the vectors:

$$\overrightarrow{\nabla} F_i(x_0), i \in \{1, \ldots, m\}.$$

Hence, we conclude that:

A regular point x_0 of F is a critical point of $f|_{\mathscr{L}}$ if and only if:

$$\overrightarrow{\nabla} f(x_0) = \sum_{i=1}^{m} \lambda_i \overrightarrow{\nabla} F_i(x_0).$$

End Lecture 11→ The numbers λ_i are again called *Lagrange multipliers*.

ing

Example 1.24. Consider the constraint $\mathscr{L}$ defined by the equations:

$$\begin{cases} xy^3 + 2z^2 - yx = 1, \\ x + z = -1, \end{cases}$$

then one can check that every point of $\mathscr{L}$ is a regular point of the $\mathcal{C}^1$ function:

$$F(x, y, z) = (xy^3 + 2z^2 - yx, x + z).$$

Now set $f(x, y, z) = x$, so that:

$$\overrightarrow{\nabla} f(x, y, z) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

and writing $F_1(x, y, z) = xy^3 + 2z^2 - yx$, $F_2(x, y, z) = x + z$ then:

$$\overrightarrow{\nabla} F_1(x, y, z) = \begin{pmatrix} y(y^2 - 1) \\ x(3y^2 - 1) \\ 4z \end{pmatrix}, \quad \overrightarrow{\nabla} F_2(x, y, z) = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Hence $(x, y, z) \in \mathscr{L}$ is a critical point for $f|_{\mathscr{L}}$ if and only if there are reals λ_1, λ_2, such that:

$$\begin{cases} xy^3 + 2z^2 - yx = 1, \\ x + z = -1, \\ \lambda_1 y(y^2 - 1) + \lambda_2 = 1 \\ \lambda_1 x(3y^2 - 1) = 0 \\ 4\lambda_1 z + \lambda_2 = 0. \end{cases}$$

Observe that, necessarily: $\lambda_1 \neq 0$ and $x \neq 0$. So it follows immediately that:

$$y = \pm \frac{1}{\sqrt{3}}, \quad \mp \frac{2\sqrt{3}}{9}\lambda_1 + \lambda_2 = 1$$

and consequently we can parametrise λ_2, x, z in terms of λ_1:

$$\begin{cases} \lambda_2 = 1 \pm \frac{2\sqrt{3}}{9}\lambda_1 \\ z = -\frac{1}{4\lambda_1} \mp \frac{\sqrt{3}}{18} \\ x = \frac{1}{4\lambda_1} \pm \frac{\sqrt{3}}{18} - 1 \end{cases}$$

after some tedious computation we find that:

$$\lambda_1^2 = \frac{27}{4(55 \mp 12\sqrt{3})}.$$

Showing that $f|_{\mathscr{L}}$ has 4 critical points. ◇

1.12.3 Higher order derivatives

← Start Lecture 12

The first derivative test is not the only means we have to detect values where extrema are attained. In 1-dimensional calculus, we also have a test with the second derivative, but what is the second derivative?

Let $f : U \subset E \to F$ be a differentiable function, then its differential at $x_0 \in U$, df_{x_0} is an element of the vector space of all continuous linear maps between E and F: $\mathcal{L}(E, F)$. This is a normed space, with norm given by:

$$\|u\|_{\mathcal{L}(E,F)} = \sup_{\substack{x \in E \\ x \neq 0}} \frac{\|u(x)\|_F}{\|x\|_E}.$$

Before proceeding lets consider this in two special cases, first, if $E = \mathbb{R}^n, F = \mathbb{R}^m$, then, as we have discussed, any linear map may be represented in matrix form as a map:

$$X \mapsto AX,$$

where A is a $m \times n$ matrix, and we identify $\mathbb{R}^n$ with column vectors. In this case, one might observe that:

$$\|u\|_{\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)} = \sup_{\substack{x \in \mathbb{R}^n \\ x \neq 0}} \frac{\|u(x)\|_{\mathbb{R}^m}}{\|x\|_{\mathbb{R}^n}} = \sup_{\substack{X \in \mathbb{R}^n \\ X \neq 0}} \frac{\|AX\|_{\mathbb{R}^m}}{\|X\|_{\mathbb{R}^n}} = \|A\|_{M_{m,n}(\mathbb{R})}.$$

There is another special case, when E is a Euclidean space and $F = \mathbb{R}$; that is when u is in the dual space. Then the Cauchy-Schwartz inequality shows that:

$$u(h) = \langle v, h \rangle, h \in E, \Rightarrow \|u\|_{E'} = \|v\|_E.$$

In particular, we have for differentiable scalar fields $f : U \subset E \to \mathbb{R}$:

$$\|df_x\|_{E'} = \left\| \vec{\nabla} f(x) \right\|_E$$

The differential of f therefore defines a map between the normed spaces:

$$\begin{array}{rccc} df : & U \subset E & \longrightarrow & \mathcal{L}(E, F) \\ & x & \longmapsto & df_x \end{array}.$$

Definition 1.21

A function $f : U \subset E \to F$ is said to be two times differentiable at x_0 if it is differentiable (on an open neighbourhood of x_0) and the differential $df : U \subset E \to \mathcal{L}(E, F)$ is differentiable at x_0. In this case we write:

$$d^2 f_{x_0} \in \mathcal{L}(E, \mathcal{L}(E, F)).$$

This is starting to look awfully complicated: the second derivative is an object that takes a vector h and assigns a linear map:

$$d^2 f_{x_0}(h) \in \mathcal{L}(E, F).$$

The result is therefore an object that takes a vector k and assigns an element of F:

$$(d^2 f_{x_0}(h))(k) \ \in F.$$

However, we can see that this final expression is linear in both the arguments h and k and therefore be considered to be a continuous bilinear map $E \times E \to F$. Hence, we write instead:

$$d^2 f_{x_0}(k, h) \in F.$$

Example 1.25.

- If $f : E \to F$ is linear, then $d^2 f_{x_0}(k, h) = 0$ at every point. Indeed, the $df_x = f$ for every $x \in E$.
- If $B : E \times E \to F$ is bilinear then: $d^2 B = B$.

◇

We now restrict to the case $E = \mathbb{R}^n$, $F = \mathbb{R}^n$, in this case we can define:

Definition 1.22: Second order partial derivatives

Let $f : U \to \mathbb{R}^n \to \mathbb{R}^m$ be a differentiable function, $x_0 \in U$. Then we define the second order partial derivatives to be the partial derivatives of the partial derivatives:

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right)(x_0),$$

when they exist.

As in the previous first order case, these determine completely the second order differential $d^2 f_{x_0}$ at x_0

Proposition 1.8

Let $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ be a twice differentiable function at x_0, then,

$$d^2 f_{x_0}(e_i, e_j) = \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0).$$

The reader may be concerned that there may be some ambiguity with the order of differentiating in our notation, in fact, luckily for us, it does not matter at all:

Proposition 1.9: Symmetry of partial derivatives

If $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ is twice differentiable at x_0, then:

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x_0),$$

in other words $d^2 f$ is a symmetric bilinear map.

Remark 1.12.2. This symmetry of $d^2 f$ is also true, and actually (in my opinon) less fiddly to prove, when $f : U \subset E \to F$. You can find the proof in André Avez's book, *Calcul différentiel*.

As before:

⚠ A function may have second order partial derivatives and still not be twice differentiable !

Once more, *if* the partial derivatives are continuous, then we recover second order differentiablity.

Proposition 1.10

Let $f : U \subset \mathbb{R}^n \to \mathbb{R}$, Suppose that the partial derivatives

$$x \mapsto \frac{\partial^2 f}{\partial x_i \partial x_j}(x)$$

exist and are continuous functions. Then f is twice differentiable on U.

Remark 1.12.3. Functions that satisfy the conditions of the above Proposition are called $\mathcal{C}^2$ functions.

The theory can be repeated to any order of differentiability $k \in \mathbb{N}$; with the same theorems. If a function is differentiable to any order $k \in \mathbb{N}$ then we say that f is smooth or $\mathcal{C}^\infty$.

1.13 Scalar fields and the Hessian matrix (skipped)

As usual, the case of *real-valued* functions, i.e. scalar fields, is special. Indeed, we saw that the derivative, which was already represented a matrix for all other dimensions, could still be represented by a vector $\vec{\nabla} f$. So we might hope that its second derivative still admits a reasonable representation. This expectation is met: it can be represented by a matrix, known as the Hessian matrix.

Definition 1.23

Let $f : U \subset \mathbb{R}^n \to \mathbb{R}$ be a differentiable function, $x_0 \in U$, then the Hessian matrix of f is defined to be:

$$\mathbf{H}\, f(x_0) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) \right)_{(i,j)\in \{1,\dots,n\}}.$$

When f is twice differentiable the matrix is symmetric.

Remark 1.13.1. When you will have done a bit more linear algebra, you will see that when f is real valued, the second derivative is a symmetric bilinear form (analogous to a scalar product) and the Hessian matrix is just the matrix of this bilinear form when f is twice differentiable.

We still require one further fact in order to discuss extrema, the generalisation of the Taylor formula:

Proposition 1.11: Second order Taylor formula for real-valued functions (matrix form)

Let $f : U \subset \mathbb{R}^n \to \mathbb{R}$ be twice differentiable at x_0, then:

$$f(\mathbf{x_0} + \mathbf{h}) = f(\mathbf{x_0}) + \langle \vec{\nabla} f(x_0), \mathbf{h} \rangle + {}^t\mathbf{h}(\mathbf{H}\, f(x_0))\mathbf{h} + \|\mathbf{h}\|^2 \varepsilon(\mathbf{h}),$$

where $\lim_{\mathbf{h}\to 0} \varepsilon(\mathbf{h}) = 0$. Here: $\mathbf{h} = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \in \mathbb{R}^n$.

1.13.1 The second derivative test (Skipped)

We will require the following fact from Linear algebra. Sorry.
A symmetric matrix is diagonalisable with real eigenvalues: there is an invertible[a] matrix P such that:

$$A = P \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix} P^{-1}$$

The numbers $\lambda_1, \dots, \lambda_n$ are the *eigenvalues* of A. These are numbers $\lambda \in \mathbb{R}$, such that there is a *non-zero* column vector $X \in \mathbb{R}^n$ such that:

$$AX = \lambda X,$$

and are exactly the roots of its *characterstic polynomial*:

$$\chi_A(X) = \det(XI - A).$$

(I is the identity matrix). (This is known as the Cayley-Hamilton theorem.)

[a] actually it can be chosen to be orthogonal i.e. $P^{-1} = {}^tP$

Example 1.26. Consider the symmetric matrix:

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

The characteristic polynomial is:

$$\begin{vmatrix} X-1 & 1 & 2 \\ 1 & X & 0 \\ 2 & 0 & X-1 \end{vmatrix} = X^3 - 2X^2 - 4X + 1.$$

One can check that the three roots are real, 2 are positive and 1 is negative. ◇

If $\mathbf{x_0}$ is a criticial point, i.e. $\vec{\nabla} f(x_0)$, then the Taylor formula shows that:

$$f(\mathbf{x_0} + \mathbf{h}) = f(\mathbf{x_0}) + {}^t\mathbf{h}(\mathbf{H}\, f(x_0))\mathbf{h} + \|\mathbf{h}\|^2 \varepsilon(\mathbf{h}),$$

when $\|\mathbf{h}\|$ is small so that we can neglect the size of the remainder time in front of the first two, one can see that if f

$${}^t\mathbf{h}(\mathbf{H}\, f(\mathbf{x_0}))\mathbf{h} < 0,$$

for all $\mathbf{h}$, then f attains a local maximum at $\mathbf{x_0}$. Indeed, in this case moving in any direction from x_0 causes the value of the function to decrease. When this condition is satisfied we say that the symmetric matrix $\mathbf{H}\ f(\mathbf{x_0})$ is *definite negative*. In the same way, it will attain a local minimum at $\mathbf{x_0}$ if

$$^t\mathbf{h}(\mathbf{H}\ f(\mathbf{x_0}))\mathbf{h} > 0,$$

we say that: $\mathbf{H}\ f(\mathbf{x_0})$ is *definite positive*. If neither conditions are satisfied we are unable to conclude in general, although we say that a critical point $\mathbf{x_0}$ is of **saddle-type** if it is neither definite positive or negative but the matrix $\mathbf{H}\ f(\mathbf{x_0})$ is invertible. In this case, $\mathbf{x_0}$ is a minimum in some directions, and a maximum in others, making the function look like a saddle, which explains the terminology.

The whole criterion relies on the being able to determine whether or not the Hessian matrix is $\mathbf{H}\ f(\mathbf{x_0})$ is definite positive or negative. The eigenvalues enable us to characterise this precisely:

Proposition 1.12

A symmetric matrix A is definite positive (resp. definite negative) if and only if its eigenvalues are strictly positive (resp. strictly negative).

When $n = 2$, recover the standard second derivative test from Calc III. See [1, Theorem 6, p.176].

Calculating the eigenvalues can sometimes be a nuisance. In fact in reality, more than their actual values, we only need to know their sign. The appropriate notion from linear algebra is the *signature* (p, q, r) of the symmetric matrix (p is the number of positive eigenvalues, q the number of negative eigenvalues, r the dimension of the kernel). The interested reader can investigate Sylvester's law of inertia and Gauss's algorithm for quadratic forms. As an example, using Gauss's algorithm, one can easily show that the matrix in the previous example is neither positive definite nor negative definite; in fact, its signature is $(2, 1, 0)$.

Chapter 2

Vector calculus in 3 dimensions

See also Sections 4.3 and 4.4 in Marsden and Tromba 6th edition In this chapter, we return to the setting of a 3-dimensional *oriented* Euclidean space, $(E, \langle \cdot, \cdot \rangle)$. From now on we will use the notation $\langle \cdot, \cdot \rangle$, to denote its inner product. These notions were defined and briefly studied in Chapter 0. As usual, you may replace $E = \mathbb{R}^3$.

This mathematical apparatus is the modern way of modelling the dimensional world described by classical Newtonian physics.

You may be wondering at this point why the sudden return from arbitrary finite dimension n to the special case $n = 3$. In fact, a large part of what we will cover in this chapter does not require us set $n = 3$ and has relatively straightforward generalisations to n-dimensional Euclidean space. However, there is a low dimensional "accident" that makes $n = 3$ slightly special; we have already encountered one of its repercussions when we discussed the cross product. This accident was exploited[1] historically as the theory was being developed, notably to talk of the curl of a vector field, and has shaped the way things are treated in other sciences where $n = 3$ is predominantly the most relevant case. In higher dimensions, there is a sense in which vectors are the "wrong" objects, and we should instead be looking at the dual space...

Nevertheless, it is worthwhile studying how $n = 3$ is treated for applications, if time permits we will discuss how to make the jump back to arbitrary dimension n.

Before we begin, it is important now to fix our ideas about what we mean when we talk about *vector fields*, which will be the main object we wish to study. Despite the fact that we are modelling them using a function: $\vec{X} : U \subset E \to E$, it should really be

[1] probably unwittingly

thought of assigning to each point if space a vector (represented by an arrow). This is represented visually for instance in Figure 2.1.

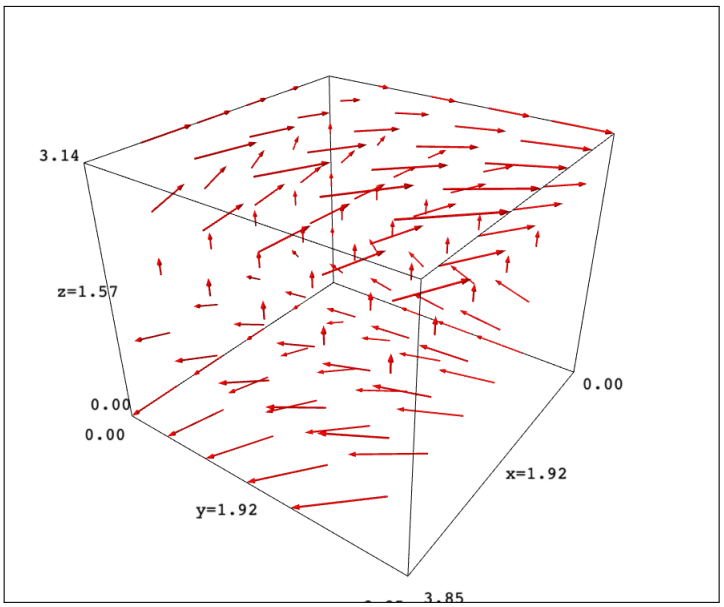


Figure 2.1: A vector field in space.

It *might* be useful for you to distinguish between the E on the left and on the right conceptually. On the left E is modelling space: you can think of it as physical space[2] $\mathscr{E}$ in which we have chosen an origin point O for a reference frame, but have not yet defined axes (or if you take $E = \mathbb{R}^3$ it is the coordinate space $\mathbb{R}^3$ we have obtained after setting up a reference frame: i.e. we have chosen an origin and orthogonal axes).

The E on the right can instead be thought of the *tangent space* T_pU to the open set U at an arbitrary point in space. However, recall that, intuitively, the tangent space is the set of tangent vectors to curves that live in U, but, here this would just be all of E (by definition of open sets), so $T_pU \simeq E$ for any p, so we can identify them all with E. Hence for our purposes, there is no need to distinguish between the tangent space and E.

Just like the example of the gradient vector field of a scalar field f, vector fields are used to describe the local behaviour of some phenomenon near a point in space, which supports the mental image we are constructing.

[2]In classical physics, physical space is modelled by an *affine* space, which can be thought of as a vector space in which we have forgotten the origin. Once we have chosen an origin, it can be identified with a vector space.

2.1 Integral curves of vector field

Now that we agree on what a vector field is, at least mathematically, we can begin discussing some concepts associated with it that will help us to understand the information it contains and, in applications, interpret what it means. The first of these is that of the *flow* of a vector field.

To construct a mental image of this, suppose we are describing the motion of a body of water, like a river. Now, instead of thinking of the river as an unthinkable number of H_2O particles and describing the total motion of each of them individually, instead, we might assimilate the river to a continuous volume, ignoring its actual composition. Now, let us suppose that the river is flowing steadily, so that the movement is the same at any time t. If we mentally imagine placing a test particle at the point (x, y, z) in the river, and then letting it go: it will flow along the river, describing some curve in the volume of water. (See Figure 2.2).

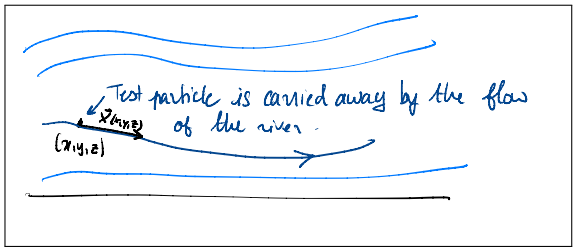


Figure 2.2: A test particle placed in water at some point will describe a curve as it is carried by the flow of the river.

Let us denote by $\vec{X}(x, y, z)$ the velocity vector of this curve at $t = 0$. In this way, we obtain a vector field over the body of water. Describing at each point the velocity a test particle will experience if we place it at a point (x, y, z).

Conversely, if a vector field $\vec{X}$ is given on some open set, we consider a curves, starting at an arbitrary point $p \in E$, such that the tangent point to each curve is given by $\vec{X}$. These are called the **integral curves** or **flow lines** of $\vec{X}$ and are defined by the following differential equation:

$$\begin{cases} \dot{\gamma}(t, p) = \vec{X}(\gamma(t, p)) \\ \gamma(0, p) = p. \end{cases}$$

The theory of ordinary differential equations (ODEs) guarantees that this can always be solved, and has a unique solution, for t in some small interval $(-\varepsilon, \varepsilon)$. We will not assume any knowledge of the theory of ODEs, and will only use this fact. Let us consider an example:

Example 2.1. Let us consider the vector field in $\mathbb{R}^2$ given by:

$$\vec{X}(x, y) = (x, -y)$$

Then the integral curve of $\vec{X}$ through $p = (x_0, y_0)$ is a curve $\gamma(t, p) = (x(t), y(t))$ such that:

$$\begin{cases} \frac{dx}{dt}(t,p) = x(t,p) \\ \frac{dy}{dt}(t,p) = y(t,p) \\ x(0,p) = x_0, y(0,p) = y_0. \end{cases}$$

Here, we can solve everything explicitly and find $x(t,p) = x_0 e^t, y(t,p) = y_0 e^{-t}$. Ignoring the dynamics, we can eliminate the parameter t and see that these are the curves described implicitly by the equation:

$$xy = x_0 y_0.$$

$\diamond$

If the vector field $\vec{X} : U \subset E \to E$ is $\mathcal{C}^1$, then one can show that there is an open subset $\mathscr{D} \subset \mathbb{R} \times M$ with the property that for each $p \in E$, $\mathscr{D}^{(p)} = \{t \in \mathbb{R}, (t,p) \in \mathscr{D}\}$ is an open interval containing 0, and such that the function:

$$\begin{array}{rccc} \gamma : & \mathscr{D} & \longrightarrow & E \\ & (t,p) & \longmapsto & \gamma(t,p) \end{array}$$

is also $\mathcal{C}^1$, this function is often denoted by:

$$(t,p) \mapsto \phi_t^X(p)$$

and is called the *flow* of the vector field, it has many useful properties but we will not study them here.

2.2 An algebraic property of volume

Integral curves will be useful to understand how a vector field might act on volume: perhaps it is describing a compressing motion, or a dilating motion: this will be captured by the notion of *divergence*, which we will derive from the integral curves.

To introduce this concept, recall that the *mixed product* of three vectors $\vec{u}, \vec{v}, \vec{w}$ is interpreted as the volume of the parallelepiped they support. In particular, the mixed product of the vectors of a positive orthonormal basis $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ is 1: this parallelepiped is of unit volume and is used as the reference for computing the volumes of more complicated shapes.

Now, we found that the map: $(\vec{u}, \vec{v}, \vec{w}) \in E \mapsto [\vec{u}, \vec{v}, \vec{w}] \in \mathbb{R}$ had two interesting properties:

1. it was linear in each of its variables,

2. it is alternating, i.e. if any two of the vectors are the same then it is vanishing.

Let us denote by $\mathrm{Vol}(\mathsf{E})$ the set of functions $\omega : \mathsf{E} \to \mathbb{R}$ with these properties then this is naturally a vector space. Indeed, if $\omega_1, \omega_2 \in \mathrm{Vol}(\mathsf{M})$ then the operators are defined by the following for any $\vec{u}, \vec{v}, \vec{w} \in \mathsf{E}, \lambda \in \mathbb{R}$

$$\begin{cases} (\omega_1 + \omega_2)(\vec{u}, \vec{v}, \vec{w}) = \omega_1(\vec{u}, \vec{v}, \vec{w}) + \omega_1(\vec{u}, \vec{v}, \vec{w}) \\ (\lambda\omega_1)(\vec{u}, \vec{v}, \vec{w}) = \lambda\omega(\vec{u}, \vec{v}, \vec{w}). \end{cases}$$

It turns out that it is of dimension 1, in other words, the mixed product is determined by these properties up to a constant.

We can see this by counting the number of free parameters in an arbitrary element $\omega \in \mathrm{Vol}(\mathsf{M})$: if $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ is a *positive orthonormal basis* of E, then decomposing:

$$\vec{u} = \sum_{i=1}^{3} u_i \vec{e}_i, \qquad \vec{v} = \sum_{i=1}^{3} v_i \vec{e}_i, \qquad \vec{w} = \sum_{i=1}^{3} w_i \vec{e}_i,$$

we can write:

$$\begin{aligned} \omega(\vec{u}, \vec{v}, \vec{w}) &= \omega(\sum_{i=1}^{3} u_i \vec{e}_i, \sum_{j=1}^{3} v_j \vec{e}_j, \sum_{i=k}^{3} w_k \vec{e}_k), \\ \textit{by linearity in the first variable} &= \sum_{i=1}^{3} u_i \omega(\vec{e}_i, \sum_{j=1}^{3} v_j \vec{e}_j, \sum_{i=k}^{3} w_k \vec{e}_k), \\ \textit{by linearity in the second variable} &= \sum_{i=1}^{3}\sum_{j=1}^{3} u_i v_j \omega(\vec{e}_i, \vec{e}_j, \sum_{i=k}^{3} w_k \vec{e}_k), \\ \textit{by linearity in the third variable} &= \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3} u_i v_j w_k \omega(\vec{e}_i, \vec{e}_j, \vec{e}_k). \end{aligned}$$

Since ω is assumed to be alternating, the sums vanish unless we have $\{i, j, k\} = \{1, 2, 3\}$: i, j, k must all be distinct and are therefore $1, 2, 3$ in some order. We can represent an ordering of $\{1, 2, 3\}$ as a one-to-one correspondence $\sigma : \{1, 2, 3\} \to \{1, 2, 3\}$ of the set $\{1, 2, 3\}$ with itself: these are known as permutations of $\{1, 2, 3\}$. One may count that there are only $3! = 6$ possible functions, which we represent in the following way:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ \sigma(1) & \sigma(2) & \sigma(3) \end{pmatrix}.$$

The six possibilities are:

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 2 & 3 & 1 \\ 1 & 3 & 2 \end{pmatrix}$$

Collecting these into a set written $\mathfrak{S}_3$, we can rewrite our computation as:

$$\omega(\vec{u},\vec{v},\vec{w}) = \sum_{\sigma\in\mathfrak{S}_n} u_{\sigma(1)}v_{\sigma(2)}w_{\sigma(3)}\omega(\vec{e}_{\sigma(1)},\vec{e}_{\sigma(2)},\vec{e}_{\sigma(3)}).$$

However, we also saw that if a function was alternating, then if we swap any two of its variables then the function is only changed by a minus sign. For instance:

$$\omega(\vec{e}_1,\vec{e}_2,\vec{e}_3) = -\omega(\vec{e}_2,\vec{e}_1,\vec{e}_3).$$

Hence, we can rearrange the order of the arguments of ω so that it is always in the order $1,2,3$, however, the price of this ordering will be a sign ± 1 which we will call $\varepsilon(\sigma)$ that depends on how many vectors we needed to swap. Overall we see that:

$$\omega(\vec{u},\vec{v},\vec{w}) = \omega(\vec{e}_1,\vec{e}_2,\vec{e}_3)\sum_{\sigma\in\mathfrak{S}_n}\varepsilon(\sigma)u_{\sigma(1)}v_{\sigma(2)}w_{\sigma(3)}.$$

The sum $\sum_{\sigma\in\mathfrak{S}_n}\varepsilon(\sigma)u_{\sigma(1)}v_{\sigma(2)}w_{\sigma(3)}$ that has appeared here is known as the *Leibniz formula for determinants* and is actually how determinants are defined. Hence, we have shown that:

$$\omega(\vec{u},\vec{v},\vec{w}) = \omega(\vec{e}_1,\vec{e}_2,\vec{e}_3)\begin{vmatrix} u_1 & v_1 & w_1\\ u_2 & v_2 & w_2\\ u_3 & v_3 & w_3\end{vmatrix} = \omega(\vec{e}_1,\vec{e}_2,\vec{e}_3)[\vec{u},\vec{v},\vec{w}].$$

If ever we stumble across a function ω with the same algebraic properties of the triple product, then there is a constant $\lambda\in\mathbb{R}$ such that:

$$\forall \vec{u},\vec{v},\vec{w}\in E,\quad \omega(\vec{u},\vec{v},\vec{w}) = \lambda[\vec{u},\vec{v},\vec{w}]$$

2.3 Divergence of a vector field

2.3.1 Definition

We will now study how volume, i.e. the triple product is affected, infinitesimally, by the flow of a vector field. Consider the following figure 2.3. The integral lines of some vector field $\vec{X}$ are represented in blue. Let us imagine that they represent the flow of some body of water. Now, consider a fish travelling along some curve in the water such that at the point p it has the velocity vector $\vec{v}$. We can transport this trajectory by the flow of the water for a fixed amount of time t and will find a new curve (in pink-red), which we can imagine would have been the trajectory of the fish had it been travelling a bit further downstream. By the chain rule, the velocity

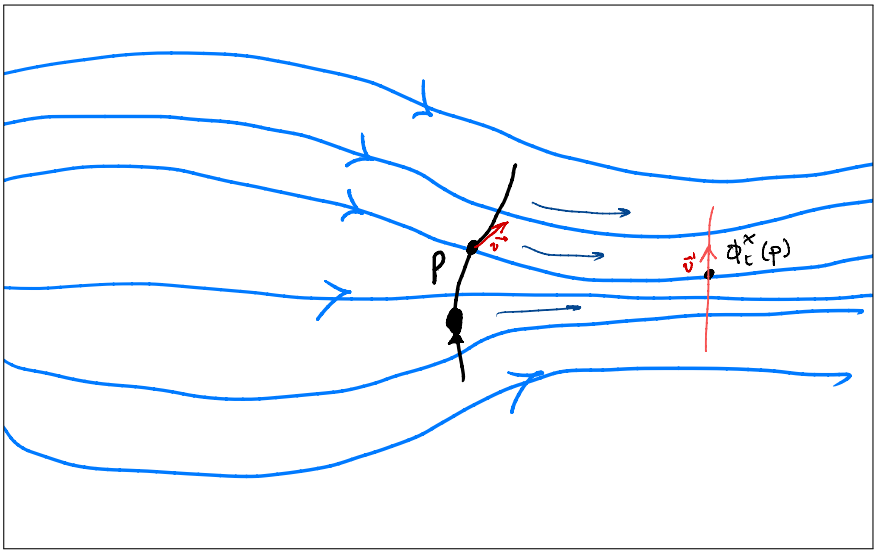


Figure 2.3:

vector $\vec{v'}$ to this curve at the point $\phi_t^X(p)$ is exactly the directional derivative of the function $p \mapsto \phi_t^X(p)$ in the direction $\vec{v}$, or in terms of differentials:

$$\vec{v'} = (d\phi_t^X)_p(\vec{v}).$$

Assume now, that we do this for vectors $\vec{u}, \vec{v}, \vec{w}$. Then we can calculate the triple product:

$$[(d\phi_t^X)_p(\vec{u}), (d\phi_t^X)_p(\vec{v}), (d\phi_t^X)_p(\vec{w})],$$

which will enable us to compare the (oriented) volume of an arbitrary parallelepiped at p, with the volume of the transformed parallelepiped under the flow of the vector field $\vec{X}$. The infinitesimal change at p will then be given by the derivative:

$$\frac{d}{dt}[(d\phi_t^X)_p(\vec{u}), (d\phi_t^X)_p(\vec{v}), (d\phi_t^X)_p(\vec{w})]\Big|_{t=0} = \lim_{t\to 0} \frac{[(d\phi_t^X)_p(\vec{u}), (d\phi_t^X)_p(\vec{v}), (d\phi_t^X)_p(\vec{w})] - [\vec{u}, \vec{v}, \vec{w}]}{t}.$$

Letting $\vec{u}, \vec{v}, \vec{w}$ be arbitrary, this limit can be viewed as defining the derivative of a map defined an open interval I of $\mathbb{R}$ into the vector space $\mathrm{Vol}(E)$. The result is consequently a new element of $\mathrm{Vol}(E)$ and there is a constant, that we write: $\mathrm{div}\,\vec{X}(p)$ and call the *divergence of* X *at* p, such that for any $\vec{u}, \vec{v}, \vec{w}$:

$$\frac{d}{dt}[(d\phi_t^X)_p(\vec{u}), (d\phi_t^X)_p(\vec{v}), (d\phi_t^X)_p(\vec{w})]\Big|_{t=0} = (\mathrm{div}\,\vec{X}(p))[\vec{u}, \vec{v}, \vec{w}].$$

In practice:

Theorem 2.1: The formula for the divergence

Let $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ be a positive orthonormal basis, then decomposing the point $p = \sum_{i=1}^{n} x_i \vec{e}_i$, and the vector field $\vec{X}(p) = \sum_{i=1}^{n} X_i(p)\vec{e}_i$. We have:

$$\operatorname{div} \vec{X}(p) = \sum_{i=1}^{3} \frac{\partial X_i}{\partial x_i}(p).$$

Proof. The (sketch of the) proof is presented an exercise for the motivated student. Without loss of generality you may take $E = \mathbb{R}^3$, it always reduces to this case upon a choice orthonormal basis.

1. First consider the map $A \mapsto \det A$ defined on $M_n(\mathbb{R})$: this is a differentiable map as it is a polynomial in the coefficients of the matrix. Show that the differential of this map at the identity matrix I is given by:

$$d\det_I(H) = \operatorname{tr}(H), \qquad H \in M_n(\mathbb{R})$$

2. Justify that:

$$[(d\phi_t^X)_p(\vec{u}), (d\phi_t^X)_p(\vec{v}), (d\phi_t^X)_p(\vec{w})] = \det(\operatorname{Jac} \phi_t^X)(p)[\vec{u}, \vec{v}, \vec{w}].$$

3. Using the chain rule and combining the above with the facts that: $\phi_0^X(p) = p$ and

$$\frac{d}{dt}\frac{\partial \phi_t^X}{\partial x_i}(p)\bigg|_{t=0} = \frac{\partial}{\partial x_i}\frac{d}{dt}\phi_t^X(p)\bigg|_{t=0} = (\partial_i \vec{X})(p)$$

conclude.

□

In view of the above formula it is common to introduce the del operator:

$$\vec{\nabla} = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix},$$

and write the divergence as a dot/inner product:

$$\operatorname{div} \vec{X} = \vec{\nabla} \cdot \vec{X} = \langle \vec{\nabla}, \vec{X} \rangle.$$

This can be useful as long as we work in Cartesian coordinates.

2.3.2 Interpretation

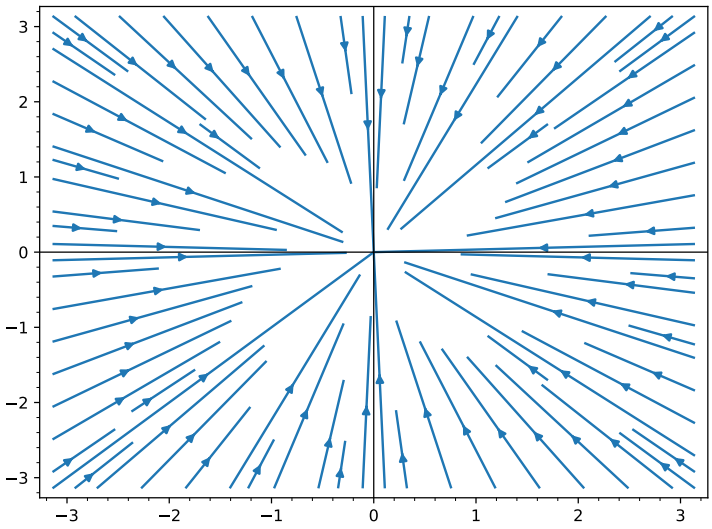
By construction, the divergence is a local measure of the change of volume as we let regions flow along the integral lines of the vector field $\vec{X}$. In particular, its sign indicates whether the flow lines are approaching one another, i.e. there is compression, or if they are spreading out, indicating expansion. In particular, we have the following interpretation of the sign of the divergence:

If a vector field $\vec{X}$ represents the velocity field of a fluid, then:
- The fluid is compressing if $\operatorname{div} \vec{X} < 0$,
- The fluid is expanding if $\operatorname{div} \vec{X} > 0$.

Example 2.2. Consider the (radial) vector field in the plane $z = 0$,

$$\vec{X}(x, y) = -x\vec{e}_x - y\vec{e}_y,$$

then its integral curves are represented as follows:



They are converging towards the point 0, hence, if this was describing a fluid we would imagine it to be compressing, and this intuition is supported by:

$$\operatorname{div} \vec{X} = \frac{\partial}{\partial x}(-x) + \frac{\partial}{\partial y}(-y) = -2 < 0.$$

◇

← End Lecture 12
← Start Lecture 13

2.3.3 Properties of the divergence

The following two properties can be derived from the definition:

1. If $\vec{X}$ and $\vec{Y}$ are two vectors fields, then:

$$\text{div}\,(\vec{X}+\vec{Y}) = \text{div}\,\vec{X} + \text{div}\,\vec{Y}.$$

2. If f is a scalar field, and $\vec{X}$ a vector field then:

$$\text{div}(f\vec{X}) = \langle\overrightarrow{\nabla} f, \vec{X}\rangle + f\text{div}\,\vec{X}.$$

The divergence of a gradient vector field of a $\mathcal{C}^2$-scalar field is a distinguished second order differential operator, known as the Laplace operator:

$$\Delta f = \text{div}\,(\overrightarrow{\nabla} f) = \vec{\nabla}\cdot\vec{\nabla} f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}.$$

2.4 The curl of a vector field

The del notation:

$$\text{div}\,\vec{X} = \overrightarrow{\nabla}\cdot\vec{X},$$

suggests that we can obtain another differential operator if we replace the dot product by the cross product:

$$\overrightarrow{\text{curl}}\,\vec{X} = \overrightarrow{\nabla}\times\vec{X}.$$

This operator is known as the *curl* operator. Whilst this "definition" is ad hoc, quite unsatisfactory and, unlike the divergence, does not immediately help us understand what information it conveys about the behaviour of the vector field, it is a good starting point for computation. It is also at this point where dimension 3 becomes important: the discussion about the divergence operator generalises immediately to arbitrary dimensions, the definition of the curl operator will not as we have used the cross product.

Starting from our definition we will find the following expression for the curl in Cartesian coordinates associated with an orthonormal basis: $(\vec{e}_x, \vec{e}_y, \vec{e}_z)$, writing:

$$\vec{X} = X_x\vec{e}_x + X_y\vec{e}_y + X_z\vec{e}_z,$$

we have:

$$\begin{aligned}\vec{\text{curl}}\,\vec{X} &= \begin{vmatrix} \frac{\partial}{\partial x} & X_x & \vec{e}_x \\ \frac{\partial}{\partial y} & X_y & \vec{e}_y \\ \frac{\partial}{\partial z} & X_z & \vec{e}_z \end{vmatrix} \\ &= (\frac{\partial X_z}{\partial y} - \frac{\partial X_y}{\partial z})\vec{e}_x - (\frac{\partial X_z}{\partial x} - \frac{\partial X_x}{\partial z})\vec{e}_y + (\frac{\partial X_y}{\partial x} - \frac{\partial X_x}{\partial y})\vec{e}_z.\end{aligned}$$

It is important to observe that whilst the divergence of a vector field is a *scalar field*, the curl of a vector is another *vector field*.

Example 2.3. Consider:

$$\vec{X}(x,y,z) = 2xz^2\vec{e}_x + \vec{e}_y + y^3zx\vec{e}_z,$$

then:

$$\vec{\nabla} \times \vec{X} = \begin{vmatrix} \frac{\partial}{\partial x} & 2xz^2 & \vec{e}_x \\ \frac{\partial}{\partial y} & 1 & \vec{e}_y \\ \frac{\partial}{\partial z} & y^3zx & \vec{e}_z \end{vmatrix} = 3y^2zx\vec{e}_x - (y^3z - 4xz)\vec{e}_y.$$

◇

Example 2.4. Let us consider the example of $\vec{X} = \vec{\nabla} f$ for some $\mathcal{C}^2$-function, then we can see that:

$$\overrightarrow{\text{curl}}\,(\vec{\nabla} f) = 0.$$

In particular, the curl can be seen to measure the default of a vector field from being a gradient vector field. ◇

The interpretation of the curl operator will be considered in more detail when we study the integral theorems of vector calculus. It is somewhat related to a local measure of how the integral curves curl around an axis locally, but this is not always easy to visualise.

We borrow the following example 9 from [1, Example 9, p.250], to develop this idea further on an example where it is particularly clear.

Example 2.5. Consider an arbitrary rigid object which is rotating around the $\vec{e}_z$ axis in $\mathbb{R}^3$ with constant angular velocity ω. Defining the vector $\vec{\omega} = \omega\vec{e}_z$, then velocity vector at any point M of the object is given by:

$$\vec{v} = \vec{\omega} \times \overrightarrow{OM},$$

where O is an arbitrary origin on the z axis. Let us consider the vector field on the open set U defining the interior of the object, then, introducing an arbitrary orthonormal basis and writing: $\overrightarrow{OM} = x\vec{e}_x + y\vec{e}_y + z\vec{e}_z$. We find that the velocity vector field is then:

$$\vec{v}(x,y,z) = \omega(x\vec{e}_y - y\vec{e}_x).$$

We compute:

$$\overrightarrow{\text{curl}}\,\vec{v} = (\vec{\nabla} \times \vec{v})(x,y,z) = \omega(\vec{e}_z + \vec{e}_z) = 2\vec{\omega}.$$

The integral curves are contained in the planes $z = \text{cst.}$ and if $\gamma(t) = (x(t), y(t), z_0)$ then:

$$\begin{cases} x'(t) = -\omega y(t) \\ y'(t) = \omega x(t). \end{cases}$$

Defining $\zeta(t) = x(t) + iy(t)$ then we can see that this is equivalent to: $\zeta'(t) = i\omega\zeta(t)$ of which the solutions are:

$$\zeta(t) = \zeta_0 e^{i\omega t} = \zeta_0 \cos(\omega t) + i\zeta_0 \sin(\omega_t),$$

where $\zeta_0 = x_0 + iy_0$ and hence we see that:

$$x(t) = x_0 \cos(\omega t) - y_0 \sin(\omega t), \quad y(t) = y_0 \cos(\omega t) + x_0 \sin(\omega t),$$

which are circles centred at the origin of the (x, y) plane.

However, it is important to note that the rotational is measuring something more than just the shape of the curves as if we consider the vector field on $\mathbb{R}^3 \setminus \{x = y = 0\}$,

$$\vec{X}(x, y, z) = \frac{y\vec{e}_x - x\vec{e}_y}{x^2 + y^2},$$

then the integral curves are contained in planes of constant z defined by the equations:

$$\begin{cases} x'(t) = \frac{y(t)}{x(t)^2 + y(t)^2} \\ y'(t) = -\frac{x(t)}{x(t)^2 + y(t)^2}. \end{cases}$$

However defining, $r(t) = \sqrt{x^2(t) + y^2(t)}$ we see that:

$$r'(t) = \frac{1}{\sqrt{x^2(t) + y^2(t)}}(x'(t)x(t) + y'(t)y(t)) = \frac{1}{r(t)^3}(y(t)x(t) - x(t)y(t)) = 0.$$

So the integral curves of this vector field are also contained in circles, but:

$$\vec{\nabla} \times \vec{X} = -\frac{x^2 - y^2}{(x^2 + y^2)^2}\vec{e}_z + \frac{x^2 - y^2}{(x^2 + y^2)^2}\vec{e}_z = 0.$$

◇

We will try to motivate further the relevance of the curl at a later point. For now, we remark that, similarly to the divergence, it has the following basic properties, if $\vec{X}$ is a vector field and f a *scalar* field then:

1. $\overrightarrow{\text{curl}}\,(\vec{X} + \vec{Y}) = \overrightarrow{\text{curl}}\,\vec{X} + \overrightarrow{\text{curl}}\,\vec{Y}$,
2. $\overrightarrow{\text{curl}}\,(f\vec{X}) = \overrightarrow{\text{grad}}\,f \times \vec{X} + f\overrightarrow{\text{curl}}\,\vec{X}$.

In applications, in particular electromagnetism, the following second order identities are useful:

Proposition 2.1: Second order vector calculus identities

Let $\vec{X}$ be a $\mathcal{C}^2$ vector field, then:

1. $\operatorname{div} \overrightarrow{\operatorname{grad}} f = \Delta f$
2. $\operatorname{div} \overrightarrow{\operatorname{curl}}\, \vec{X} = 0$
3. $\overrightarrow{\operatorname{curl}}\, \overrightarrow{\operatorname{curl}}\, \vec{X} = \overrightarrow{\operatorname{grad}} \operatorname{div} \vec{X} - \Delta \vec{X}$.
4. $\overrightarrow{\operatorname{curl}}\, \overrightarrow{\operatorname{grad}}\, f = 0$

In the above the Laplacian on vector fields is given, in *Cartesian coordinates* by:

$$\Delta \vec{X} = (\Delta X_x)\vec{e}_x + (\Delta X_y)\vec{e}_y + (\Delta X_z)\vec{e}_z.$$

Remark 2.4.1. Using the double cross product formula 0.4 from the beginning of the course, can you guess the third property?

Example 2.6. Maxwell's equation in vacuum are:

$$\begin{cases} \operatorname{div} \vec{E} = 0 \\ \overrightarrow{\operatorname{curl}} \vec{E} = -\frac{\partial \vec{B}}{\partial t}, \\ \operatorname{div} \vec{B} = 0, \\ \overrightarrow{\operatorname{curl}} \vec{B} = \underbrace{\mu_0 \varepsilon_0}_{\frac{1}{c^2}} \frac{\partial \vec{E}}{\partial t}. \end{cases}$$

It is custom to introduce a vector potential $\vec{A}$ such that $\vec{B} = \overrightarrow{\operatorname{curl}}\, \vec{A}$, and a scalar potential ϕ such that: $\vec{E} = \overrightarrow{\operatorname{grad}}\, \phi - \frac{\partial \vec{A}}{\partial t}$. In fact, the potentials $\vec{A}$ and ϕ always exist (at least locally) but are not uniquely determined, this is an example of what is known as a *gauge theory*. This so-called gauge freedom allows us to impose an extra constraint:

$$\operatorname{div} \vec{A} - \frac{1}{c^2} \frac{\partial \phi}{\partial t} = 0,$$

defining what is known as the Lorenz gauge.

We can use the above identities to derive the second order equations satisfied by these potentials: plugging the expression for $\vec{B}$ into the last equation we find that:

$$\overrightarrow{\operatorname{grad}} \operatorname{div} \vec{A} = \Delta \vec{A} - \frac{1}{c^2} \frac{\partial^2 \vec{A}}{\partial t^2} + \frac{1}{c^2} \overrightarrow{\operatorname{grad}} \frac{\partial}{\partial t} \phi.$$

Similarly, since $\operatorname{div} \vec{E} = 0$:

$$\Delta \phi = \frac{\partial}{\partial t} \operatorname{div} \vec{A}.$$

We therefore obtain the famous *wave* equation on both potentials:

$$\begin{cases} \Delta\vec{A} - \frac{1}{c^2}\frac{\partial^2\vec{A}}{\partial t^2} = 0, \\ \Delta\phi - \frac{1}{c^2}\frac{\partial^2\phi}{\partial t} = 0. \end{cases}$$

◇

This is possibly one of the most important partial differential equations in mathematics. Understanding it has lead to the development of many techniques in the their study. It is still studied in many different contexts today.

End Lecture 13→

2.5 Path and line integrals

Start Lecture 14→

Whilst we can immediately appreciate the geometric significance of the divergence, our presentation of the curl operator is relatively obscure and unsatisfying: its interpretation and relevance can only truly be appreciated in the context of Stoke's theorem which, in its modern formulation, is a higher dimensional generalisation of the fundamental theorem of calculus:

$$\int_a^b f'(t)dt = f(b) - f(a).$$

However, here, we do not want to integrate over an interval $[a, b]$, but instead more complicated geometric objects like curves and surfaces. For instance, we might want to compute the work of a force acting on a point-like particle as it follows its trajectory, or make sense of the flux of a field through a surface.

The mathematical question here is *what* objects can be integrated along curves and surfaces. In this section, we begin with the case of curves.

It is useful to make a distinction between the notions of paths and geometric curves: Recall that a *path* in 3-space is described by a function:

$$\gamma : [a, b] \longrightarrow \mathbb{R}^3.$$

However, the *geometric* curve $\mathscr{C}$ is the set of points reached by this function γ:

$$\mathscr{C} = \gamma([a, b]) = \{\gamma(t), t \in [a, b]\},$$

and γ is just a means of describing this set. In particular, one can change the parameter t, for instance, by setting $t = \phi(s)$ for some bijective function $[c, d] \to [a, b]$ and obtain a new parametrisation $\tilde{\gamma}(s) = \gamma(\phi(s))$, which describes the *same set of points* in space.

From a physical and mathematical perspective, the choice of parametrisation should not matter in so much as we only care about the set of points. A choice of a specific parameter is instead associated with questions of dynamics, i.e. how something is

moving along the curve. Hence, a good notion of integral along a curve should be independent of any choice of parameter.

Let $f : \mathbb{R}^3 \to \mathbb{R}$ be a scalar field, we could try to define an integral along the curve as follows

$$\int_a^b f(\gamma(t))dt.$$

However, recall the change of variable formula for a 1D integral:

$$\int_{\phi^{-1}(a)}^{\phi^{-1}(b)} f(\phi(s))\phi'(s)ds = \int_a^b f(t)dt,$$

So if we reparametrise the curve setting $t = \phi(s)$, we will have:

$$\int_a^b f(\gamma(t))dt = \int_{\phi^{-1}(a)}^{\phi^{-1}(b)} f(\tilde{\gamma}(s))\phi'(s)ds \neq \int_{\phi^{-1}(a)}^{\phi^{-1}(b)} f(\tilde{\gamma}(s))ds.$$

This is therefore not a good notion for an integral...

The solution is to find a geometric object associated with a curve that transforms under change of parametrisation like the change of variable formula, it turns out that a *differential form* is such an object... Remember the linear maps dx, dy,dz that popped up when we wrote:

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz?$$

Recall that these are just the projection maps that act on vectors and give the components in the directions $\vec{e}_x$,$\vec{e}_y$ and $\vec{e}_z$ respectively. They form a basis of the vector space all linear functions $\mathcal{L}(\mathbb{R}^3, \mathbb{R})$ which we called the dual space of $(\mathbb{R}^3)'$ and observed that can be identified with the set of all row matrices.

Definition 2.1: Differential one forms

A differential one form on an open set $U \subset \mathbb{R}^3$ is a differentiable map, $\alpha : U \to (\mathbb{R}^3)'$: these are expressions of the form:

$$\alpha_{(x,y,z)} = \alpha_1(x,y,z)dx + \alpha_2(x,y,z)dy + \alpha_3(x,y,z)dz, \quad (x,y,z) \in U$$

where $\alpha_i : \mathbb{R}^3 \to \mathbb{R}$ ($i \in \{1,2,3\}$) are differentiable functions.

You have already encountered an important class of differential forms:

Example 2.7. Let $f : U \subset \mathbb{R}^3 \to \mathbb{R}$ be a differentiable scalar field, then its differential df is the differential one-form:

$$df_{(x,y,z)} = \frac{\partial f}{\partial x}(x,y,z)dx + \frac{\partial f}{\partial y}(x,y,z)dy + \frac{\partial f}{\partial z}(x,y,z)dz.$$

$\diamond$

It turns out these are exactly the objects that can be integrated in a meaningful way along a curve, as long as one has chosen an *orientation of the curve*. We will return to this point later, first let us define:

Definition 2.2: Integral of a differential form along a path

Let $\gamma : [a, b] \to U \subset \mathbb{R}^3$ be a (piecewise) C^1 path and α a differential form on U, then, writing: $\gamma(t) = (x(t), y(t), z(t))$

$$\int_\gamma \alpha = \int_a^b \alpha_{\gamma(t)}(\gamma'(t))dt,$$
$$= \int_a^b \left(\alpha_1(x(t), y(t), z(t))x'(t) + \alpha_2(x(t), y(t), z(t))y'(t) + \alpha_3(x(t), y(t), z(t))z'(t)\right) dt$$

To understand the meaning of $\alpha_{\gamma(t)}(\gamma'(t))$ we can think of α as assigning to each point in space a measuring tool on vectors, and here, we are applying this along the curve.

Using the chain rule, we find that this definition accomplishes (up to a subtlety) the desired parametrisation independence:

Proposition 2.2: Parametrisation independence of the integral of differential forms

Let $\gamma : [a, b] \to \mathbb{R}^3$ be a path and $\phi : [c, d] \to [a, b]$ an *increasing* bijective $\mathcal{C}^1$ map (with $\mathcal{C}^1$ inverse), then defining:

$$\tilde{\gamma} = \gamma \circ \phi,$$

for any differential form α:

$$\int_{\tilde{\gamma}} \alpha = \int_\gamma \alpha.$$

Proof. Performing the change of variable $t = \phi(s)$ in the integral:

$$\int_\gamma \alpha = \int_a^b \alpha_{\gamma(t)}(\gamma'(t))dt = \int_c^d \alpha_{\gamma(\phi(s))}(\gamma'(\phi(s))\phi'(s))ds,$$
$$= \int_c^d \alpha_{\tilde{\gamma}(s)}(\tilde{\gamma}'(s))ds = \int_{\tilde{\gamma}} \alpha,$$

For the first equation, recall that $\alpha_{\gamma(t)}$ is a *linear map*, and the second equation follows from the chain rule. $\square$

Example 2.8. Let $\alpha_1 = y\mathrm{d}x - x\mathrm{d}y$ and $\gamma : t \in [0, 2\pi] \mapsto (\cos(t), \sin(t))$ be a curve in the plane $z = 0$, then:

$$\int_\gamma \alpha_1 = \int_0^{2\pi} (\sin(t)(-\sin(t)) - \cos(t)\cos(t))\,\mathrm{d}t = -2\pi.$$

◇

Example 2.9. Let f be a scalar field and $\gamma : [a, b] \to \mathbb{R}^3$ a *closed* curve, i.e. $\gamma(a) = \gamma(b)$ then:

$$\int_\gamma \mathrm{d}f = \int_a^b \mathrm{d}f_{\gamma(t)}(\gamma'(t))\mathrm{d}t = \int_a^b (f \circ \gamma)'(t)\mathrm{d}t = f(\gamma(b)) - f(\gamma(a)) = 0.$$

◇

Remark 2.5.1. It follows then that α_1 from the example above is *not* $\mathrm{d}f$ for some scalar field f, this motivates the following notion.

Definition 2.3: Exact differential forms

A differential form $\alpha = \alpha_1\mathrm{d}x + \alpha_2\mathrm{d}y + \alpha_3\mathrm{d}z$ is said to be **exact** if there is a scalar field f such that $\alpha = \mathrm{d}f$.

← End Lecture 14
← Start Lecture 15

The subtlety in Proposition 2.2 is that we chose ϕ to be *increasing*, but what would have happened if we had let $\phi : [c, d] \to [a, b]$ be decreasing instead? In this case, we would have: $\phi(d) = a$ and $\phi(c) = b$ and performing the change of variable $t = \phi(s)$ in the integral:

$$\begin{aligned}\int_\gamma \alpha = \int_a^b \alpha_{\gamma(t)}(\gamma'(t))\mathrm{d}t &= \int_d^c \alpha_{\gamma(\phi(s))}(\gamma'(\phi(s))\phi'(s))\mathrm{d}s, \\ &= \int_d^c \alpha_{\tilde\gamma(s)}(\tilde\gamma'(s))\mathrm{d}s = -\int_{\tilde\gamma} \alpha.\end{aligned}$$

This means that the integral of a differential form is sensitive (up to a sign) to way we walk along the curve; hence we should always integrate along *oriented* objects.

Using the above results we can define an integral over a certain class of geometric curve $\mathscr{C}$:

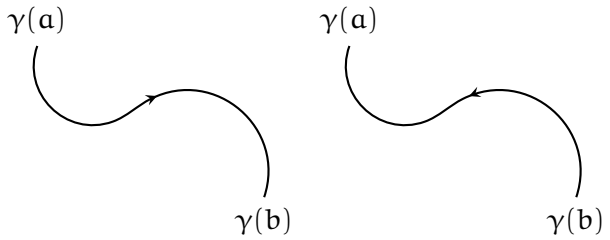
Definition 2.4

A geometric curve $\mathscr{C}$ is said to be **regular** $\mathcal{C}^1$ curve it admits a parametrisation $\gamma : [a, b] \to \mathscr{C}$ such that:

$$\forall t \in [a, b], \gamma'(t) \neq 0,$$

It is said to be **simple** if the parametrisation γ can be chosen to be injective when restricted to the interval $[a, b)$. (So the curve does not have any self intersections).
A curve or path is said to be **closed** if $\gamma(a) = \gamma(b)$.

At each point of a simple regular $\mathcal{C}^1$ curve, the direction of the tangent line is given by $\gamma'(t) \neq 0$. Referring to our discussion about orientation at the beginning of the course, we can orient each tangent line individually by choosing arbitrarily whether $\gamma'(t)$ or $-\gamma'(t)$ is pointing in the positive direction. However, we would like the orientation of the tangent lines to be consistent from point to point and therefore impose that the positive direction move continuously with t. Here, since $\gamma'(t)$ is assumed to be non-vanishing, once we have imposed the orientation at one point it will be automatically determined at all the others. Intuitively, we can assimilate this to deciding if we are walking along the curve from $\gamma(a)$ to $\gamma(b)$ or $\gamma(b)$ to $\gamma(a)$, this is illustrated below:



It therefore makes sense to talk about the integral along an *oriented* simple regular $\mathcal{C}^1$ geometric curve (or arc) $\mathcal{C}$, the sign of this integral will be reversed if the orientation is reversed. We set:

Definition 2.5

Let $\mathcal{C}$ be an oriented simple regular $\mathcal{C}^1$ curve then:

$$\int_{\mathcal{C}} \alpha = \int_{\gamma} \alpha,$$

where $\gamma : [a, b] \to \mathbb{R}^3$ is any parametrisation that maps $[a, b)$ injectively into $\mathscr{C}$ and such that $\gamma'(t) \neq 0$ is positively oriented at each point.

Let us point out a subtlety that explains why we are suddenly concerned about self intersections and the injectivity of the map γ on $[a, b)$. Consider the two parametrised curves:

$$\gamma_1(t) = (\cos(t), \sin(t)), \quad \gamma_2(t) = (\cos(2t), \sin(2t)).$$

Now as t varies between 0 and 2π they describe the same set of points, with the exception that γ_2 goes around the circle twice. Let us consider the differential form:

$$\alpha = -\frac{y\,dx}{x^2+y^2} + \frac{x\,dy}{x^2+y^2}$$

Now:

$$\frac{1}{2\pi}\int_{\gamma_1} \alpha = \frac{1}{2\pi}\int_0^{2\pi} (\sin^2(t) + \cos^2(t))dt = 1,$$

whereas:

$$\frac{1}{2\pi}\int_{\gamma_1} \alpha = \frac{1}{2\pi}\int_0^{2\pi} (2\sin^2(2t) + 2\cos^2(2t))dt = 2!$$

Hence, we should *not* consider these parametrisations to be equivalent: they are not related simply by reparametrisation (in the sense defined above) as the integral is picking up on how many times we are going round the circle. Both integrals are however useful, for instance, in physical applications, if an object is moving and it goes around the circle twice, then one might expect the work done by a force to be doubled, and so the second computation is relevant, but if we are only concerned about the geometric curve, there is no reason to go around the curve multiple times, and so the integral along $\mathscr{C}$ is by convention the first one.

As a motivating example for the notion of integral along a path, I mentioned the notion of work of a force. This involves a vector field $\vec{F}$ describing the force, but we know how to integrate differential forms... Luckily for us, using the inner product of $\mathbb{R}^3$, $\langle , \cdot \rangle$, we can obtain a differential form from $\vec{F}$ by considering the quantity:

$$\langle \vec{F}, \cdot \rangle.$$

Indeed, if $\vec{e}_x, \vec{e}_y, \vec{e}_z$ is our standard basis of $\mathbb{R}^3$ then, writing:

$$\vec{F} = F_x\vec{e}_x + F_y\vec{e}_y + F_z\vec{e}_z,$$

we have:

$$\langle \vec{F}, \rangle = F_x\langle \vec{e}_x, \cdot \rangle + F_y\langle \vec{e}_y, \cdot \rangle + F_z\langle \vec{e}_z, \cdot \rangle,$$

but if $h = h_x\vec{e}_x + h_y\vec{e}_y + h_z\vec{e}_z$ then:

$$\langle e_x, h \rangle = h_x = dx(h),$$

so

$$\langle e_x, \cdot \rangle = dx!$$

Similar observations can be made for the other quantities, so the differential form associated with the vector field F is:

$$\langle \vec{F}, \cdot \rangle = F_x dx + F_y dy + F_z dz,$$

introducing $d\vec{l} = dx\vec{e}_x + dy\vec{e}_y + dz\vec{e}_z$ this is sometimes written:

$$\vec{F} \cdot d\vec{l}.$$

Definition 2.6

The line integral of a vector field $\vec{F} = F_x\vec{e}_x + F_y\vec{e}_y + F_z\vec{e}_z$ along a curve or path parametrised by $\gamma : [a, b] \to \mathbb{R}^3$ is the integral:

$$\int_\gamma \vec{F} \cdot d\vec{l} = \int_\gamma F_x dx + F_y dy + F_z dz.$$

Example 2.10. Consider the vector field:

$$\vec{F}(x, y, z) = -\frac{1}{(x^2 + y^2 + z^2)^{\frac{3}{2}}}(x\vec{e}_x + y\vec{e}_y + z\vec{e}_z)$$

and the path parametrised by $\gamma(t) = (\cos(t), \sin(t), 0), t \in [0, 2\pi]$, then:

$$\begin{aligned}\int_\gamma \vec{F} \cdot d\vec{l} &= -\int_\gamma \frac{x dx}{(x^2 + y^2 + z^2)^{\frac{3}{2}}} + \frac{y dy}{(x^2 + y^2 + z^2)^{\frac{3}{2}}} + \frac{z dz}{(x^2 + y^2 + z^2)^{\frac{3}{2}}} \\ &= \int_0^{2\pi} (\cos(t)\sin(t) - \sin(t)\cos(t))dt = 0\end{aligned}$$

◇

Example 2.11. Let f be a scalar field and $\vec{F} = \vec{\nabla} f$, and $\gamma : [a, b] \to \mathbb{R}^3$, then:

$$\int_\gamma \vec{\nabla} f \cdot d\vec{l} = \int_\gamma \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz = \int_\gamma df = f(\gamma(b)) - f(\gamma(a)).$$

In particular, if the curve is closed then the integral vanishes. ◇

There is another distinguished object that we can integrate along a geometric curve, that is intrinsic to any (piecewise) simple regular $\mathcal{C}^1$ curve:

$$ds = \|\gamma'(t)\|dt = \text{“}\sqrt{dx^2 + dy^2 + dz^2}\text{”}.$$

Recall that the length of the curve between $\gamma(a)$ and $\gamma(b)$ is:

$$L = \int_a^b ds.$$

We can use this object to integrate functions over the line, indeed if $f : U \subset \mathbb{R}^3 \to \mathbb{R}$ is a continuous scalar field and we consider a path parametrised by $\gamma : [a, b] \to \mathscr{C}$, then we can set:

$$\int_\gamma f ds = \int_a^b f(\gamma(t))\|\gamma'(t)\|dt.$$

By the same method as for the line integral, one can check that this integral is parametrisation independent, and does not even depend on orientation !

Example 2.12. Consider the path integral of f along the simple regular curve $\mathscr{C}$ parametrised by:

$$\gamma(t) = (1 + 2t, 3 + 4t, t), t \in [0, 1]$$

of the function $f(x, y, z) = x$.

$$\int_{\mathscr{C}} f ds = \int_0^1 (1 + 2t)\sqrt{4 + 16 + 1}dt = \sqrt{21}\left[\frac{1}{4}(1 + 2t)^2\right]_0^1 = \frac{8\sqrt{21}}{4}.$$

Let us test parametrisation independence by performing the change of variable: $s = 1 - t$:

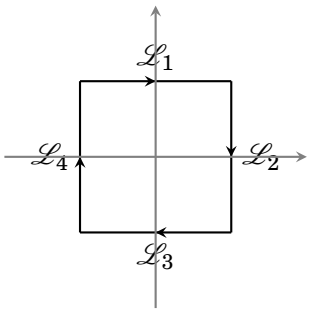
$$\sqrt{21}\int_0^1 (1+2t)dt = \sqrt{21}\int_1^0 (3-2s)(-ds) = \sqrt{21}\int_0^1 (3-2s)ds = -\frac{\sqrt{21}}{4}(1-9) = \frac{8\sqrt{21}}{4}.$$

◇

Remark 2.5.2. If you are aware of the distinction between the Lebesgue and Riemann integral, see below. I am using here the Riemann integral, which is actually oriented, $\int_a^b = -\int_b^a$ as opposed to the Lebesgue integral which is not: it is only concerned with the set $[a, b]$. We write when $a < b$, $\int_{[a,b]} = \int_a^b$. The change of variable formula for a *Lebesgue* integral automatically involves $|\phi'|$ as opposed to ϕ' in the Riemann case.

To avoid complicating things in the Definition and Theorem statement, I have made assumptions about the regularity of the curves and paths. These can of course be relaxed by allowing these assumptions to only be satisfied in a piecewise manner: i.e. that we can break our curve up into a finite number of smaller pieces where the assumptions are satisfied. This means that the assumptions only fail at a finite number of points and in integration theory we will see that these points are of "measure zero" which means the integral is insensitive to them. The definitions then are generalised by using our definitions on each part and adding the results together. This enables us to integrate over rectangles, or paths "with corners" which is often done in physics.

Example 2.13. Consider in a plane the square curve $\mathscr{S}$ with vertices $(-1, 1)$, $(1, 1)$, $(1, -1)$ and $(-1, -1)$ represented below and oriented in the clockwise direction as shown:



and the differential form:

$$\alpha = dx + xdy$$

Then we split up the curve into four lines that we parametrise (for example) as so:

$$\begin{cases} \mathscr{L}_1 : & t \in [0,1] \ \mapsto (-1+2t, 1), \\ \mathscr{L}_2 : & t \in [0,1] \mapsto (1, 1-2t), \\ \mathscr{L}_3 : & t \in [0,1] \mapsto (1-2t, -1), \\ \mathscr{L}_4 : & t \in [0,1] \ \mapsto (-1, -1+2t). \end{cases}$$

Then:

$$\begin{aligned} \int_{\mathscr{S}} \alpha &= \sum_{i=1}^{4} \int_{\mathscr{L}_i} \alpha \\ &= \int_0^1 (2 + (-2) + (-2) + (-2)) dt = -4. \end{aligned}$$

◇

Example 2.14. Check that the integral over the square of $\alpha = df$ (where f is a $\mathcal{C}^1$ scalar field), vanishes. ◇

End Lecture 15→

2.6 A very brief review of integration in $\mathbb{R}^n$

Start Lecture 16→

2.6.1 The basic concepts of integration in the sense of Lebesgue

Nowadays, especially with the modern development of probability theory and its importance in a number of different domains and applications, it is not recommendable to be completely unaware of the modern theory of Lebesgue integration. The purpose of this section is to give you a very brief account of some of the ideas of Lebesgue integration, that you can research further in books on Measure or Probability theory. See for example [4, 3].

The basic stage for Lebesgue integration is an arbitrary space X, equipped with a distinguished collection of sets $\mathscr{A}$ which are referred to as: *measurable*. You can

think of X as being $\mathbb{R}^2$ and $\mathscr{A}$ will in this case be the collection of subsets of $\mathbb{R}^2$ for which one will be able to define its area.

The collection $\mathscr{A}$ of subsets is not arbitrary and satisfies a number of rules that are designed to satisfy our intuitive idea of area or volume. Indeed, one expects to be able to define the total size of the full space X, (even if its infinite), so we always assume:

1. $X \in \mathscr{A}$.

If $(A_n)_{n\in\mathbb{N}}$ are a *countable* selection of measurable sets, then we also expect their union to be measurable i.e.

2. $(A_n)_{n\in\mathbb{N}} \in \mathscr{A}^{\mathbb{N}} \Rightarrow \bigcup_{n\in\mathbb{N}} A_n \in \mathscr{A}$.

Finally, if A is measurable then the subset $A^c = X \setminus A$ should be measurable too, i.e.

3. $A \in \mathscr{A} \Rightarrow X \setminus A \in \mathscr{A}$.

These rules are summarised by saying that $\mathscr{A}$ is a σ-algebra.

The couple $(X, \mathscr{A})$ is known as a *measurable space.* (Probability theorists sometime say a probabilisable space)

The third and final ingredient for integration theory according to Lebesgue is the notion of *measure*. This is a function:

$$\mu : \mathscr{A} \to [0, +\infty],$$

that can take the value $+\infty$ and which assigns to each *measurable set* $A \in \mathscr{A}$, its size or measure. It is also assumed to have some basic intuitive properties:

1. $\mu(\emptyset) = 0$,

and, if $(A_n)_{n\in\mathbb{N}}$ is a countable family of pairwise disjoint sets, then:

2. $\mu(\bigcup_{n\in\mathbb{N}} A_n) = \sum_{n=0}^{\infty} \mu(A_n)$.

This second assumption is what justifies our method of computing volumes and area by splitting complicated objects up into smaller measurable parts.

Remark 2.6.1. When doing probability theory, the size of the total space is always chosen $\mu(X) = 1$.

These assumptions imply other intuitive properties, for instance if $A \subset B$ are two measurable subsets of X then:

$$\mu(A) \leqslant \mu(B).$$

Indeed: $\mu(B) = \mu(A) + \underbrace{\mu(B \setminus A)}_{\geqslant 0}$.

Another property is that if $\mu(A) < +\infty$, for some measurable subset A then:

$$\mu(X) = \mu(A) + \mu(X \setminus A) \Rightarrow \mu(X \setminus A) = \mu(X) - \mu(A).$$

This is especially useful when $\mu(X) < +\infty$, for instance, when $\mu(X) = 1$ we have that:

$$\mu(X \setminus A) = 1 - \mu(A).$$

Remark 2.6.2. As long as we avoid expressions like $+\infty - \infty$, we can allow things to be infinite, it is custom to agree on the following rules:

- $+\infty + \infty = +\infty$
- $+\infty + \alpha = +\infty$, $\alpha \in \mathbb{R}$
- $0 \times (+\infty) = 0$
- $\alpha \times (+\infty) = \operatorname{sgn}(\alpha)\infty$

These rules are *not* continuous.

Example 2.15. Let $X = \mathbb{N}$, one might take $\mathscr{A} = \mathscr{P}(X)$, the set of all subsets of X, (which satisfies all the properties) and then for any $n \in X$ one can define a measure:

$$\delta_n(A) = \begin{cases} 1 & \text{if } n \in A \\ 0 & \text{if } n \notin A. \end{cases}$$

One can check that $\delta_n(\emptyset) = 0$. If $A = \bigcup_{i \in \mathbb{N}} A_i$ and (A_i) are pairwise disjoint, then either $n \notin A$ in which case $\mu(A) = 0$ and, since $A = \bigcup_{i \in \mathbb{N}} A_i$ it follows that:

$$\forall i \in \mathbb{N}, n \notin A_i \ \Rightarrow \forall i \in \mathbb{N}, \mu(A_i) = 0,$$

therefore the equality

$$\mu(A) = 0 = \sum_{i=1}^{\infty} \mu(A_i),$$

holds, or, $n \in A$, which means that there is some $i_0 \in \mathbb{N}$ such that $n \in A_{i_0}$. However, since the (A_i) are pairwise disjoint, i.e. $A_i \cap A_j = \emptyset$ when $i \neq j$ and it follows that: $n \notin A_i$ when $i \neq i_0$. Hence,

$$\mu(A_i) = \begin{cases} 0 & \text{when } i \neq i_0 \\ 1 & \text{when } i = i_0 \end{cases}$$

thus, once more:

$$1 = \mu(A) = \sum_{i=1}^{\infty} \mu(A_i).$$

This measure is known as the Dirac measure at n. ◇

Example 2.16. Let us consider once more the measurable space $(\mathbb{N}, \mathscr{P}(\mathbb{N}))$ from the previous example. We can define a measure $c : \mathscr{P}(\mathbb{N})$ known as the *counting measure*, which to a set assigns the cardinality (number of elements) of the set:

$$c(A) = \begin{cases} \text{card(A)} & \text{if } A \text{ finite} \\ +\infty & \text{if } A \text{ is infinite} \end{cases}$$

One can check that this is a measure. ◇

A key idea is to try to choose $\mathscr{A}$ as large as possible to include any reasonable set whose *size* you want to measure, although it turns out that, in general, unlike in the previous example, it will not include all possible subsets of X (but those which are excluded will in general be so fantastical that you only run into them if you try to.)

When a set A satisfies $\mu(A) = 0$ it is said to be *negligible*: these sets are invisible for all means and purposes in terms of integration theory.

Once you have the three ingredients, $(X, \mathscr{A}, \mu)$ defining a *measured space* (probability theorists might say a probability space), you can define the integral of a class of so-called *measurable functions*, which include (phew!) continuous functions when $X = \mathbb{R}^n$ with the standard $\mathscr{A}$ we choose there.

The notion of integral we end up with is very flexible – notice, for example, that we are always integrating over the whole space X – and is meant to directly generalise the notion of *mean*.

One first defines the integral of positive *simple* functions. These are functions f that can be written:

$$f(x) = \sum_{i=1} c_i \mathbf{1}_{A_i}(x), x \in X,$$

where $A_i \in \mathscr{A}$, and are pairwise disjoint, $c_i \in \mathbb{R}_+$ and $\mathbf{1}_{A_i}$ is called the *characteristic function* of the measurable set A_i and defined by:

$$\mathbf{1}_{A_i}(x) = \begin{cases} 1 & \text{if } x \in A_i \\ 0 & \text{otherwise.} \end{cases}$$

The integral is then defined as:

$$\int_X f d\mu = \sum_{i=1}^{n} c_i \mu(A_i) \in [0, +\infty]$$

Observe that this may be infinite. If $\mu(X) = 1$ then this is precisely the *expected* value of the positive simple function f.

Remark 2.6.3. The representation of a simple function as a sum:

$$\sum_{i=1}^{n} c_i \mathbf{1}_{A_i}$$

may not be unique. For instance, one might choose to divide A_i into smaller parts, or perhaps $c_i = c_j$ for some $i \neq j$. It is important to check that the integral does not depend on how we split things up. Assume that:

$$f(x) = \sum_{i=1}^{n} c_i \mathbf{1}_{A_i} = \sum_{j=1}^{m} \tilde{c}_j \mathbf{1}_{B_j},$$

without loss of generality we may assume $c_i \neq 0$, $\tilde{c}_j \neq 0$, and then it follows that:

$$\bigcup_{i=1}^{n} A_i = \bigcup_{j=1}^{m} B_j = \{x \in A, f(x) \neq 0\},$$

but then:

$$\sum_{i=1}^{n} c_i \mu(A_i) = \sum_{i=1}^{n} c_i \mu\left(\bigcup_{j=1}^{m} (A_i \cap B_j)\right) = \sum_{i=1}^{n} \sum_{j=1}^{m} c_i \mu(A_i \cap B_j)$$

and on $A_i \cap B_j$ we have $c_i = \tilde{c}_j$, hence:

$$\sum_{i=1}^{n} c_i \mu(A_i) = \sum_{i=1}^{n} \sum_{j=1}^{m} \tilde{c}_i \mu(A_i \cap B_j) = \sum_{j=1}^{m} \tilde{c}_j \mu(\bigcup_{i=1}^{n} A_i \cap B_j) = \sum_{j=1}^{m} \tilde{c}_j \mu(B_j).$$

Which shows that our definition of the integral is coherent.

Alternatively, we could have also defined a simple function $f : X \to \mathbb{R}_+$ to be a "measurable" function such that $f(X)$ is a finite set, $\{c_1, \ldots, c_m\}$, then it is clear that the representation:

$$f(x) = \sum_{i=1} c_i \mathbf{1}_{A_i},$$

where $A_i = \{x \in X, f(x) = c_i\}$ is uniquely determined. The fact that the A_i are measurable is hidden in the assumption that f is measurable. For completeness, I will just mention that a function $f : X \to [-\infty, +\infty]$ is measurable if for any $t \in \mathbb{R}$, $\{x \in X : f(x) \leqslant t\} \in \mathscr{A}$. Observe that we allow the function to take infinite values, but to be able to do anything useful this should only happen on negligible sets for the measure.

If A is a measurable subset of X, then one says that $f : A \to [-\infty, +\infty]$ is measurable under the same condition, that is for any $t \in \mathbb{R}$, $\{x \in A : f(x) \leqslant t\} \in \mathscr{A}$. Such a function can always be extended to a measurable function $f : X \to [-\infty, +\infty]$ by setting $f(x) = 0$ on $X \setminus A$ (which is in $\mathscr{A}$). Indeed, if $t \in \mathbb{R}$, and $t < 0$, $\{x \in X : f(x) \leqslant t\} = \{x \in A : f(x) \leqslant t\}$ and so is in $\mathscr{A}$, if $t \geqslant 0$ then:

$$\{ x \in X, f(x) \leqslant t\} = \{x \in A, f(x) \leqslant t\} \cup \underbrace{\{x \in X \setminus A, f(x) \leqslant t\}}_{= X \setminus A},$$

so is in $\mathscr{A}$. So for simplicity, we can just consider mesurable functions on the whole space.

Observe that by *definition*:

$$\int_X \mathbf{1}_A d\mu = \mu(A).$$

To extend the integral to other types of functions, we essentially do approximation by simple functions and this is where analysis takes over. For instance, the integral of a positive function f defined by:

$$\int_X f d\mu = \sup \left\{ \int s d\mu, s \text{ simple positive and } s \leqslant f \right\} \in [0, +\infty]$$

We allow the value $+\infty$. For an arbitrary measurable function we say that f is *integrable* if:

$$\int_X |f| d\mu < +\infty$$

and then it turns out we can define:

$$\int_X f d\mu = \int_X f_+ d\mu - \int_X f_- d\mu.$$

where f_+ and f_- are the positive and negative parts of $f = f_+ - f_-$, which are both defined to be positive functions.

$$f_+ = \frac{f + |f|}{2}, \quad f_- = \frac{|f| - f}{2}.$$

This is the point where analysis takes over completely (and we stop). From here, the natural path for the theory is to go on to derive approximation theorems (Fatou's lemma, the monotone convergence theorem (Beppo-Levi), the dominated convergence theorem) which enable us to reduce hard theorems to the case of simple functions.

To define an integral over a measurable subset A, we just set:

$$\int_A f d\mu = \int_X f \mathbf{1}_A(x) d\mu.$$

If a set A is negligible:

$$\int_A f d\mu = 0.$$

Example 2.17. Let $X = \mathbb{N}$, $\mathscr{A} = \mathscr{P}(X)$, be equipped with the Dirac measure at some n. Consider the function:

$$f(x) = 2\mathbf{1}_{\{1,2,3\}} + 3\mathbf{1}_{\{n,4,5\}}$$

then:

$$\int_{\mathbb{N}} f(x) d\delta_n = 2\delta_n(\{1,2,3\}) + 3\delta_n(\{n,4,5\}) = 3 = f(n).$$

In fact, using the limit theorems we have not given, this is true for any measurable function. ◇

Example 2.18. Consider now the counting measure c on the measurable space $(\mathbb{N}, \mathscr{P}(\mathbb{N}))$, then one can in fact show that for *any* measurable function $f : \mathbb{N} \to \mathbb{R}$:

$$\int_{\mathbb{N}} f dc = \sum_{n=0}^{+\infty} f(n).$$

So Lebesgue theory puts integrals and series on an equal footing. (This means that Fubini's theorem can be used on series, or for swapping Σ and $\int$, the numerous limit theorems work too...) ◇

2.6.2 The Lebesgue measure on $X = \mathbb{R}^n$

When $X = \mathbb{R}^n$, one of the first questions one asks is:

Can one find $\mathscr{A}$ and a measure, often written, λ such that:
a) $\mathscr{A}$ contains all my usual sets, (e.g. open sets ...)
b) λ generalises in a unique way the notion of volume from n-dimensional rectangles :

$$\lambda([a_1, b_1] \times \cdots \times [a_n, b_n]) = (b_1 - a_1) \times \cdots \times (b_n - a_n),$$

to any of these sets?

The answer is *yes* for any n: the corresponding $\mathscr{A}$ is then known as the Lebesgue σ-algebra and λ as the Lebesgue measure. The Lebesgue σ-algebra is enormous, it contains all open sets, closed sets, compact sets, we will not seek to describe it further. Although, as a cultural remark, using the axiom of choice (which turns out to be necessary), one can show that there are sets that are not Lebesgue measurable, see[3, Example 1.4.8, Theorem 1.4.9].

An interesting point is that singletons (sets with one point) are always of zero Lebesgue measure:

$$\forall x \in \mathbb{R}^n, \lambda(\{x\}) = 0,$$

this is the reason why when it comes to, integration we can often allow our assumptions, to break down on countable sets. Of course, we can actually allow them to break down on any negligible set for the Lebesgue measure (and these can be large, $\mathbb{Q}$ is negligible in $\mathbb{R}$ for the Lebesgue measure, see also the Cantor set [3, Proposition 1.4.7].)

When $X = \mathbb{R}$ ($n = 1$), we recover (not without some work) the standard theory of integration in a more flexible framework. In particular, we obtain the usual computation rules for integrals (fundamental theorem of calculus, integration by parts) that you are aware of. The theorems are obtained in a slightly more general setting than continuous functions, but for our purposes we will just consider that everything we have learnt before should stay learnt. I should point out that although in principle one should write ($a \leqslant b$):

$$\int_{[a,b]} f d\lambda,$$

we continue to write:

$$\int_a^b f(x)dx = \int_{[a,b]} f d\lambda,$$

and define:

$$\int_b^a f(x)dx = -\int_{[a,b]} f d\lambda.$$

In order to take $X = \mathbb{R}^n$ with $n > 1$ we need to do some more work: we need to be able to actually calculate integrals. In analysis, this work involves the notion of product

measure $\mu \otimes \nu$ of two measures over distinct spaces X and Y, and there are some subtleties that we can sweep under the rug. The result is summarised in what is commonly referred to as Fubini's theorem (actually the Fubini-Tonnelli theorem). It can be applied iteratively and, in practice, it boils down to[3], the following statements (lets look at $\mathbb{R}^2$):

- When $f : \mathbb{R}^2 \to \mathbb{R}$ is a (Lebesgue) *positive* measurable function, for instance, continuous or piecewise continuous, then we always have:

$$\int_{\mathbb{R}^2} f d\lambda = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} (f(x,y)\lambda(dx)) \right) \lambda(dy) = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} (f(x,y)\lambda(dy)) \right) \lambda(dx).$$

  This should be understood to mean that if one of these quantities is infinite then they all are, and if one is finite they are all finite and have the same value.

- When f is an *integrable* measurable function, then:

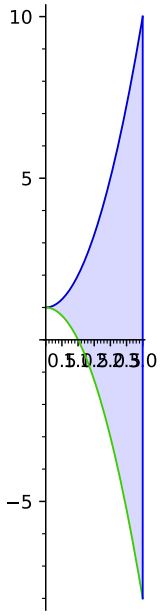
$$\int_{\mathbb{R}^2} f d\lambda = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} (f(x,y)\lambda(dx)) \right) \lambda(dy) = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} (f(x,y)\lambda(dy)) \right) \lambda(dx).$$

Remark 2.6.4. The λ in $\lambda(dx)$, which indicates we are using the Lebesgue measure, will eventually disappear.

When f is a (piecewise) continuous function, and D is a bounded measurable set, then Fubini's theorem will in fact always apply. Therefore, for our purposes, we are just going to practice computing integrals and the skill of how to "set the bounds" in the iterated integrals. This really is simply just about translating the meaning of multiplication by the function $\mathbf{1}_D$ which ahs the effect of "cutting off" the function outside of D.

Example 2.19. Let us consider the region $D = \{-x^2 + 1 \leqslant y \leqslant x^2 + 1, 0 \leqslant x \leqslant 3\}$, represented as the shaded blue region below:

[3] in reality these statements are a bit incorrect although they are essentially true



It has an exceptional property, if one fixes x, then the set:

$$D_x = \{\, y \in \mathbb{R}, (x,y) \in D\},$$

has a surprisingly simple description:

$$D_x = \begin{cases} \{-1 + x^2 \leqslant y \leqslant x^2 + 1\} & \text{if } 0 \leqslant x \leqslant 3 \\ \emptyset & \text{otherwise} \end{cases}$$

If f is a continuous function on $\mathbb{R}^2$, then according to Fubini's theorem:

$$\int_D f dx dy = \int_{\mathbb{R}^2} f(x,y)\mathbf{1}_D(x,y) dx dy = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x,y)\mathbf{1}_D(x,y) dy \right) dx.$$

Now, if x is fixed, $\mathbf{1}_D(x,y) = 1$ if and only if $y \in D_x$, so in fact:

$$\int_D f dx dy = \int_{\mathbb{R}} \left(\int_{D_x} f(x,y) dy \right) dx$$

But $D_x = \emptyset$ if $x \notin [0,3]$ so:

$$\int_D f dx dy = \int_0^3 \left(\int_{-1+x^2}^{1+x^2} f(x,y) dy \right) dx.$$

Let us work out the bounds if we instead want to integrate with respect to x first. For this we need to determine, $D_y = \{x \in \mathbb{R}, (x,y) \in D\}$

← End Lecture 16
← Start Lecture 17

Let us first observe that if $(x,y) \in D$ then, $-8 \leqslant y \leqslant 10$, we shall restrict to those values of y. Now if y is fixed;

$$(x,y) \in D \Leftrightarrow 0 \leqslant x \leqslant 3, 1-x^2 \leqslant y \leqslant 1+x^2,$$

but this means that $1-y \leqslant x^2$ and $y-1 \leqslant x^2$ or, in other words:

$$x^2 \geqslant |y-1|, 0 \leqslant x \leqslant 3$$

Hence we conclude that:

$$(x,y) \in D, y \in [-8,10] \Leftrightarrow \sqrt{|y-1|} \leqslant x \leqslant 3$$

and therefore:

$$\int_D f(x,y)dxdy = \int_{-8}^{10} \left(\int_{\sqrt{|y-1|}}^{3} f(x,y)dx \right) dy.$$

◇

Let us consider another example:

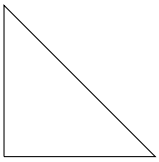
Example 2.20. If $f : [a,b] \to \mathbb{R}$ is a positive (continuous) function, Fubini's theorem allows us to *prove* that the integral is equal to the area under the curve (as defined by the Lebesgue measure) which is usually how a 1D integral is "defined". Indeed, the region under the curve is:

$$D = \{x \in [a,b], 0 \leqslant y \leqslant f(x)\}.$$

$$\lambda(D) = \int_{\mathbb{R}^2} \mathbf{1}_D(x,y)dxdy = \int_a^b \left(\int_0^{f(x)} dy \right) dx = \int_a^b f(x)dx.$$

◇

Example 2.21. Let D be the triangular region T whose vertices are $(0,0), (0,2), (2,0)$



To describe the region we can observe that, $0 \leqslant y \leqslant 2$ and that for fixed y, T_y is given by:

$$T_y = \{x \in \mathbb{R}, (x,y) \in T\} = [0, 2-y].$$

So:

$$\int_T f(x,y)dx = \int_0^2 \left(\int_0^{2-y} f(x,y)dx \right) dy.$$

The other way round, we can observe that $0 \leqslant x \leqslant 2$ and if x is fixed:

$$T_x = [0, 2-x]$$

◇

Fubini's theorem can be applied iteratively, using $\mathbb{R}^3 = \mathbb{R}\times\mathbb{R}^2 = \mathbb{R}\times(\mathbb{R}\times\mathbb{R})$ enabling us to reduce integrals over regions of $\mathbb{R}^3$ to integrals over $\mathbb{R}$. More explicitly, we have for positive and integrable measurable functions, for instance:

$$\int_{\mathbb{R}\times\mathbb{R}^2} f d\lambda = \int_{\mathbb{R}} \left(\int_{\mathbb{R}^2} f(x,y,z) d\lambda(y,z) \right) \lambda(dx) = \int_{\mathbb{R}^2} \left(\int_{\mathbb{R}} f(x,y,z)\lambda(dx) \right) d\lambda(y,z).$$

One then applies[4] Fubini's theorem again on the integral over $\mathbb{R}^2$. Of course, it works no matter how you break up $\mathbb{R}^3$ into $\mathbb{R}$ and $\mathbb{R}^2$, and you will have up $3! = 6$ possible permutations of all the integrals.

Example 2.22. Once more, we can *prove* using Fubini's theorem that an integral of a positive measurable function on $\mathbb{R}^2$ is exactly the volume under the surface. Indeed, set $A = \{(x,y,z) \in \mathbb{R}^3, 0 \leqslant z \leqslant f(x,y)\}$. The measurability assumption garanties that this is measurable and so we can look at:

$$\begin{aligned}\lambda(A) = \int_{\mathbb{R}^3} \mathbf{1}_A d\lambda &= \int_{\mathbb{R}^2} \left(\int_{\mathbb{R}} \mathbf{1}_A \lambda(dz) \right) d\lambda(x,y) \\ &= \int_{\mathbb{R}^2} \left(\int_0^{f(x,y)} \lambda(dz) \right) d\lambda(x,y) = \int_{\mathbb{R}^2} f d\lambda.\end{aligned}$$

If one only wanted to consider (x,y) in a smaller range D where D is a measurable subset of $\mathbb{R}^2$, then, we would just modify the definition of A:

$$A = \{(x,y,z) \in \mathbb{R}^3, (x,y) \in D, 0 \leqslant z \leqslant f(x,y),$$

or cut off the function f by multiplying by $\mathbf{1}_D$, and arrive at the same conclusion. ◇

When integrating over a bounded domain, the game is the same as in $\mathbb{R}^2$ except it can be harder to visualise and we need to trust our ability to write down the conditions defining D in *equivalent* ways.

Remark 2.6.5. We have never really discussed this explicitly in class, but it is important to understand the meaning of the symbol $\Leftrightarrow$. When we write $A \Leftrightarrow B$ we *really* mean that if A is true then B is true *and* if B is true then A. Both of these should be checked (mentally) before using the symbol $\Leftrightarrow$ between two mathematical statements. Hence for $x \in \mathbb{R}$:

$$x^2 = 1 \Leftrightarrow x = 1$$

is a false statement (and a common error made by students, even of advanced calculus), whilst:

$$x^2 = 1 \Leftrightarrow x \in \{-1, 1\},$$

is the correct statement ($x \in \mathbb{R}$). Hence, it means that when doing computations, to ensure we have an equivalence, we must check that at each step we can go backwards.

[4] essentially, there are some subtleties that we can ignore

So

$$x = xy \Leftrightarrow x(x-y) = 0 \Leftrightarrow x = 0 \text{ or } x = y$$

but whilst:

$$x - y = 0 \Rightarrow x(x-y) = 0,$$

of course $x(x-y) = 0 \Rightarrow x = y$ is false.

Let us consider a region, V of $\mathbb{R}^3$ defined by:

$$V = \{x \leqslant z \leqslant 1, 1-x \leqslant y \leqslant 1, 0 \leqslant x \leqslant 1\}.$$

Given the current description of this set, we can see that the easiest order of integration one can write is to integrate with respect to x last and with respect to y and z first. In particular, if $x \in [0,1]$:

$$V_x = \{(y,z) \in \mathbb{R}^2, (x,y,z) \in \mathbb{R}^2\} = [1-x,1] \times [x,1].$$

This is a rectangle and Fubini's theorem shows immediately that the order of integration with respect to y and z does not matter. So we have:

$$\int_V f(x,y,z)\,dx\,dy\,dz = \int_0^1 \left(\int_x^1 \left(\int_{1-x}^1 f(x,y,z)\,dy \right) dz \right) dx = \int_0^1 \left(\int_{1-x}^1 \left(\int_x^1 f(x,y,z)\,dz \right) dy \right) dx.$$

Suppose we wanted to integrate with respect to z first. We can see that z can take any value between $[0,1]$ within the region. So let us fix z and try to find V_z. We must have:

$$0 \leqslant x \leqslant z \quad \text{and} \quad 1-x \leqslant y \leqslant 1.$$

This is an implication, but one can check that:

$$\begin{cases} x \leqslant z \leqslant 1, \\ 1-x \leqslant y \leqslant 1, \\ 0 \leqslant x \leqslant 1, \end{cases} \Leftrightarrow \begin{cases} 0 \leqslant x \leqslant z, \\ 1-x \leqslant y \leqslant 1, \\ 0 \leqslant z \leqslant 1. \end{cases}$$

This means that if *all* of the statements on the left-hand side are true **then** *all* of the statements on the right-hand side true, **and vice versa**. This is what *guarantees* that we are describing the same set on both sides and allows us to rewrite the order of integration:

$$\int_0^1 \left(\int_0^z \left(\int_{1-x}^1 f(x,y,z)\,dy \right) dx \right) dz.$$

We could also try to swap the order between y and x, observing that:

$$\begin{cases} 0 \leqslant x \leqslant z, \\ 1-x \leqslant y \leqslant 1, \\ 0 \leqslant z \leqslant 1 \end{cases} \Leftrightarrow \begin{cases} 1-z \leqslant y \leqslant 1, \\ 1-y \leqslant x \leqslant z, \\ 0 \leqslant z \leqslant 1, \end{cases}$$

we have:

$$\int_0^1 \left(\int_{1-z}^1 \left(\int_{1-y}^z f(x,y,z)dx \right) dy \right) dz.$$

Once more the key point is the equivalence: $\Leftrightarrow$. If you mess this up, for example, most commonly, by forgetting a condition in your reasoning by implication which leads to you describing a larger set than the original set, the sentence is *immediate.* This is one equation where some rigour is essential in order to help us not make a mistake, especially when the domain we are trying to describe is hard to visualise.

Let us check that these integrals are all identical when $f = 1$.

- $\int_0^1 \left(\int_x^1 \left(\int_{1-x}^1 dy \right) dz \right) dx = \int_0^1 (1-x)x dx = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}$
- $\int_0^1 \left(\int_{1-x}^1 \left(\int_x^1 dz \right) dy \right) dx = \frac{1}{6}$.
- $\int_0^1 \left(\int_0^z \left(\int_{1-x}^1 f(x,y,z)dy \right) dx \right) dz = \int_0^1 \int_0^z x dx dz = \int_0^1 \frac{z^2}{2} dz = \frac{1}{6}$.
- $\int_0^1 \left(\int_{1-z}^1 \left(\int_{1-y}^z dx \right) dy \right) dz = \int_0^1 \int_{1-z}^1 (z+y-1)dy dz = \int_0^1 \frac{z^2}{2} dz = \frac{1}{6}$.

2.6.3 The change of variable formula

We quote without proof the higher dimensional change of variable formula:

Theorem 2.2: Change of variables in $\mathbb{R}^n$

Let U, V be open subsets of $\mathbb{R}^n$ and $\phi : U \to V$ be a bijection of U onto V such that both ϕ and ϕ^{-1} are $\mathcal{C}^1$ [a], then if $f : V \to \mathbb{R}$ is integrable then:

$$\int_V f d\lambda = \int_U f(\phi(x)) |\det(\text{Jac } \phi(x))| d\lambda$$

[a] this is what is called a $\mathcal{C}^1$-diffeomorphism

Example 2.23. Consider: $\phi : (r, \theta, \varphi) \mapsto (r\cos\varphi\sin\theta, r\sin\varphi\sin\theta, r\cos\theta)$ where: $U = \mathbb{R}^* \times (0,\pi) \times (0, 2\pi)$, $V = \phi(U)$, then it is a simple exercise to check that ϕ is a C^1 diffeomorphism. Now, $\mathbb{R}^3 \setminus V$ is actually a negligible set so, in reality (and this is what saves us), it is the same to integrate over $\mathbb{R}^3$ or V. The above theorem then says:

$$\int_{\mathbb{R}^3} f(x,y,z)dxdydz = \int_U f(r\cos\phi\sin\theta, r\sin\phi\sin\theta, r\cos\theta) r^2 \sin\theta dr d\theta d\varphi.$$

In practice we then split the second integral up using Fubini's theorem:

$$\int_{\mathbb{R}^3} f(x,y,z)dxdydz = \int_0^{+\infty} \left(\int_0^{\pi} \left(\int_0^{2\pi} f(r\cos\phi\sin\theta, r\sin\phi\sin\theta, r\cos\theta) r^2 \sin\theta d\varphi \right) d\theta \right) dr,$$

and can swap the order of integration as long as f is integrable.

Let us use this to compute the volume of a ball B of radius R:

$$\begin{aligned}\lambda(B) &= \int_B d\lambda = \int_{\mathbb{R}^3} \mathbf{1}_B(x)d\lambda \\ &= \int_0^{+\infty} \left(\int_0^{\pi} \left(\int_0^{2\pi} \mathbf{1}_B(r\cos\phi\sin\theta, r\sin\phi\sin\theta, r\cos\theta) r^2 \sin\theta d\varphi \right) d\theta \right) dr,\end{aligned}$$

Since $\mathbf{1}_B(r\cos\phi\sin\theta, r\sin\phi\sin\theta, r\cos\theta) = 1$ if and only if $r \leqslant R$ it follows that:

$$\lambda(B) = \int_0^R \left(\int_0^{\pi} \left(\int_0^{2\pi} r^2 \sin\theta d\varphi \right) d\theta \right) dr = \left(\int_0^R r^2 dr \right) \left(\int_0^{\pi} \sin\theta d\theta \right) \left(\int_0^{2\pi} d\varphi \right) = \frac{4\pi R^3}{3}.$$

◇

Example 2.24. Let $f(x,y) = e^{-x^2-y^2}$ then, using the change of variables $x = r\cos\theta, y = r\sin\theta$, we have:

$$\int_{\mathbb{R}^2} e^{-x^2-y^2} d\lambda = \int_0^{2\pi} \left(\int_0^{+\infty} e^{-r^2} r dr \right) d\theta = \pi.$$

However, by using Fubini's theorem in the first integral we have:

$$\int_{\mathbb{R}^2} e^{-x^2-y^2} d\lambda = \left(\int_{\mathbb{R}} e^{-x^2} dx \right) \left(\int_{\mathbb{R}} e^{-y^2} dy \right) = \left(\int_{\mathbb{R}} e^{-x^2} dx \right)^2.$$

Therefore:

$$\int_{\mathbb{R}} e^{-x^2} dx = \sqrt{\pi}.$$

◇

2.7 The Green-Riemann theorem

We are now ready to state the first of three integral theorems (which are, in fact, in reality all the same theorem).

Theorem 2.3: Green-Riemann

Let Ω be a bounded connected open set of $\mathbb{R}^2$ whose boundary $\partial\Omega$ is a (piecewise) $\mathcal{C}^1$ simple closed regular curve. Orient the curve $\partial\Omega$ counterclockwise, then:

$$\int_{\partial\Omega} P dx + Q dy = \int_{\Omega} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy.$$

Here P and Q are $\mathcal{C}^1$ functions defined an open neighbourhood of $D = \Omega \cup \partial\Omega \subset \mathbb{R}^2$

Remark 2.7.1. Our assumptions about the boundary $\partial\Omega$ guarantee that it is of Lebesgue measure zero and therefore:

$$\int_{\Omega} = \int_{D}.$$

There are many other ways to state the theorem, allowing for "rougher" domains D.

The right hand side of the equation may seem a little unnatural. As an initial observation, let us consider the differential form $P dx + Q dy$ as a differential form in $\mathbb{R}^3$ that does not depend on the third coordinate z. We can consider the dual vector field, $\vec{F}(x, y, z) = P(x, y)\vec{e}_x + Q(x, y)\vec{e}_y$ and compute its curl:

$$\overrightarrow{\text{curl}}\,\vec{F} = \begin{vmatrix} \partial_x & P & \vec{e}_x \\ \partial_y & Q & \vec{e}_y \\ \partial_z & 0 & \vec{e}_z \end{vmatrix} = \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\vec{e}_z.$$

In the Marsden-Tromba textbook [1], the quantity $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}$ is called the *scalar curl* for this reason.

Proof. We will prove it in a special case where D is a rectangle $[a, b] \times [c, d]$ and $Q = 0$. We can treat the case $P = 0$ in the same way and, by linearity, add the result together. Here:

$$\begin{aligned} \int_{\partial\Omega} P dx &= \int_a^b P(t, c)dt - \int_a^b P(t, d)dt \\ &= -\int_a^b [P(t, d) - P(t, c)]dt \\ &= -\int_a^b \left(\int_c^d \frac{\partial P}{\partial y}(t, s)ds\right) dt = -\int_{[a,b]\times[c,d]} \frac{\partial P}{\partial y}(x, y)dxdy. \end{aligned}$$

Where in the last equality we have renamed the variables t, s and used Fubini's theorem. □

Example 2.25. Let $Q(x, y) = x, P(x, y) = 0$, then by the Green-Riemann theorem:

$$\int_{\partial\Omega} x dy = \int_{\Omega} dxdy = \text{Area}(\Omega),$$

similarly, with $P(x, y) = y$, $Q(x, y) = 0$:

$$-\int_{\partial\Omega} y dx = \text{Area}(\Omega).$$

Hence:

$$\text{Area}(\Omega) = \frac{1}{2}\int_{\partial\Omega} x dy - y dx.$$

This can be a useful formula to determine the area of some region. ◇

Example 2.26. Let $a > 0$, determine the area enclosed by a cardioide $\mathscr{C}$ which admits the following parametrisation:

$$\begin{cases} x(t) = 2a(1-\cos t)\cos t, \\ y(t) = 2a(1-\cos t)\sin t, \end{cases} \quad t \in [0, 2\pi].$$

Let A denote the area we wish to calculate, then:

$$\begin{aligned} A &= \frac{1}{2}\int_{\mathscr{C}} x\,dy - y\,dx \\ &= \frac{1}{2}\int_0^{2\pi} \Big[2a(1-\cos t)\cos t(2a\sin^2 t + 2a(1-\cos t)\cos t) \\ &\qquad - 2a(1-\cos t)\sin t(2a\sin t\cos t - 2a(1-\cos t)\sin t)\Big]dt \\ &= 2a^2\int_0^{2\pi}(1-\cos t)^2 dt. \end{aligned}$$

Before we evaluate the remaining integral, observe that combining both formulae for the area into one, helped us simplify the expression considerably. Now:

$$\int_0^{2\pi}(1-\cos t)^2 dt = \int_0^{2\pi} dt - \underbrace{2\int_0^{2\pi}\cos t\,dt}_{=0} + \int_0^{2\pi}\cos^2 t\,dt = 3\pi.$$

Hence:

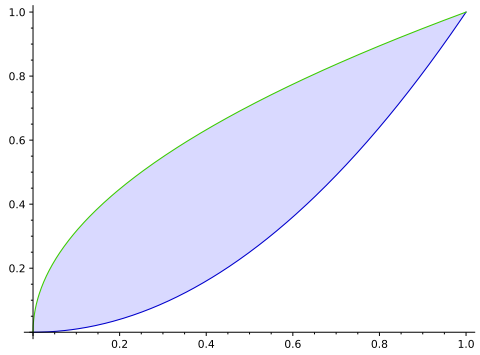
$$A = 6\pi a^2.$$

End Lecture 17→
Start Lecture 18→

◇

Example 2.27. Let D be the bounded domain enclosed by the curves $y = x^2$ $x = y^2$, and $\mathscr{C}$ its boundary, oriented in the counterclockwise sense. Using the Green-Riemann theorem we can calculate:

$$\int_{\mathscr{C}} (2xy - x^2)dx + (x + y^2)dy$$



First, from the diagram we can see that:

$$D = \{(x, y) \in \mathbb{R}^2, 0 \leqslant x \leqslant 1, x^2 \leqslant y \leqslant \sqrt{x}\},$$

and by the Green-Riemann theorem we have:

$$\int_{\mathscr{C}} (2xy - x^2)dx + (x + y^2)dy = \int_D (1 - 2x)\, dxdy.$$

By Fubini's theorem:

$$\begin{aligned}\int_D (1-2x)dxdy &= \int_0^1 (1-2x)\left(\int_{x^2}^{\sqrt{x}} dy\right) dx,\\ &= \int_0^1 (1-2x)(\sqrt{x} - x^2)dx,\\ (\text{set } x = t^2) &= 2\int_0^1 t^2(1-2t^2)(1-t^3)dt = \frac{1}{30}.\end{aligned}$$

◇

One final example to illustrate how Green-Riemann's theorem can be used to replace integrals over domains by integrals over curves. Let

$$D = \{(x, y) \in \mathbb{R}^2, x \geqslant 0, y \geqslant 0, \frac{x^2}{a^2} + \frac{y^2}{b^2} \leqslant 1\},$$

and suppose we want to calculate:

$$\int_D (2x^3 - y)dxdy.$$

Then setting $P = \frac{1}{2}y^2$, $Q = \frac{1}{2}x^4$, we have:

$$\int_D (2x^3 - y)dxdy = \int_{\partial D} \left(\frac{1}{2}y^2 dx + \frac{1}{2}x^4 dy\right).$$

∂_D can be split into 3 parts: the line segments $[OA]$, $[BO]$ where $O = (0, 0)$, $A = (a, 0)$, $B = (0, b)$ and an arc of an ellipse from A to B. We parametrise as follows (making sure to run along the boundary counterclockwise):

$$\begin{cases}\gamma_1(t) = (ta, 0), t \in (0, 1),\\ \gamma_2(t) = (a\cos t, b\sin t), t \in [0, \frac{\pi}{2}],\\ \gamma_3(t) = (0, (1-t)b).\end{cases}$$

Therefore:

$$\begin{aligned}\int_{\partial D} \left(y^2 dx + x^4 dy\right) &= \int_0^1 0dt + \int_0^{\frac{\pi}{2}} (-ab^2 \sin^3 t + ba^4 \cos^5 t)dt + \int_0^1 0dt\\ &= \frac{4}{15}a^4 b - \frac{ab^2}{3}.\end{aligned}$$

The details are left as an exercise.

2.8 Parametrised surfaces in $\mathbb{R}^3$

2.8.1 Definition

Green-Riemann's marvellous formula relates an integrals over "reasonable" domains of $\mathbb{R}^2$ with integrals over their boundaries which, not incidentally, are curves: objects of one dimension less. It is natural to wonder if the formula is a low dimensional accident or if it can be generalised to higher dimensions. Nothing about the proof of the theorem supports the idea that it is an accident: the main tool of our (simplified) proof is Fubini's theorem; which is valid in all dimensions. In higher dimensions, there is more room for different types of objects what might want to integrate over. For instance, in $\mathbb{R}^3$, there are, of course, curves, but, we have also encountered surfaces. Can integrals over "nice" bounded open sets in $\mathbb{R}^3$ be related to integrals over their boundaries, which would now be a surface? How do we integrate over a surface? and what object can be integrated over a surface? Before we venture into these questions, it is useful to introduce another way of describing a surface in $\mathbb{R}^3$. That ressembles how we have been discussing curves.

Definition 2.7: Parametrised surfaces

A subset S is called a *parametrised surface* in $\mathbb{R}^3$ if it is the image of a $\mathcal{C}^1$ map:

$$\begin{array}{rccc} \Phi: & U \subset \mathbb{R}^2 & \longrightarrow & \mathbb{R}^3 \\ & (u,v) & \longmapsto & \Phi(u,v)\,, \end{array}$$

where U is an open set in $\mathbb{R}^2$ (we normally assume it is connected, so in "one piece").

A *reparametrisation* of S is a bijective $\mathcal{C}^1$ map $\phi : V \to U$ with $\mathcal{C}^1$ inverse. The map: $\tilde{\Phi} = \Phi \circ \phi$ then describes the same surface S.

A parametrised surface is said to be **regular** at $p = \Phi(u,v)$ if:

$$\frac{\partial \Phi}{\partial u}(u,v) \times \frac{\partial \Phi}{\partial v}(u,v) \neq 0,$$

it is said to be regular if it is regular at every point.

Remark 2.8.1.

- This definition of a parametrised surface is especially weak: it can allow for strange phenomena like corners, self-intersections... that, for some purposes, we would like to disallow.
- Furthermore, one might remark that, in reality, it would be more precise to call parametrised surface the set of all parameterisations that are related by reparametrisation, rather than the shared image set. Indeed, we will always need a parametrisation to talk about many of the concepts we discuss below.

- It is not forbidden that $\Phi(u,v) = \Phi(u',v')$ for $(u',v') \neq (u,v) \in U$, this can lead to ambiguities. The problem can be avoided imposing that Φ be injective.
- The condition $\frac{\partial \Phi}{\partial u}(u,v) \times \frac{\partial \Phi}{\partial v}(u,v) \neq 0$ guarantees that $d\Phi_{(u,v)}$ is an injective linear map. (Equivalently that $(\text{Jac } \Phi(u,v))X = 0 \Rightarrow X = 0$, for any $X \in \mathbb{R}^3$). In this case Φ is called an *immersion*.
- Many of the surfaces we naturally consider in $\mathbb{R}^3$ satisfy the following more restrictive definition that can be found in many textbooks on the differential geometry of curves and surfaces:

  $S \subset \mathbb{R}^3$ is called regular $\mathcal{C}^1$ surface, if for any point $p \in S$ one can find an open neighbourhood V or $p \in \mathbb{R}^3$, an open set $U \subset \mathbb{R}^3$, and a bijective $\mathcal{C}^1$ map: $\Phi : U \to \mathbb{R}^3$ such that $\Phi(U) = V \cap S$, the map: $\Phi : U \to S \cap V$ has continuous inverse and:

  $$\frac{\partial \Phi}{\partial u}(u,v) \times \frac{\partial \Phi}{\partial v}(u,v) \neq 0.$$

  The map Φ is called a local parametrisation, and Φ^{-1} is a local chart.
- In some situations it is useful to allow the parametrisation to be defined on a set D that is not open. In this case, it is understood that ϕ is $\mathcal{C}^1$ if it is the restriction to D of a $\mathcal{C}^1$ defined on an open set containing D.

  There are many advantages to this definition, which avoids many nasty phenomena and enables us to do calculus on S. However, these advantages cannot be appreciated at our current level of exposition, which is why we will stick to parametrised surfaces.

Just as for curves, properties of the surface should be stable under reparametrisation.

Example 2.28. The unit sphere $S^2 = \{(x,y,z) \in \mathbb{R}^3, x^2 + y^2 + z^2 = 1\}$ in $\mathbb{R}^3$ can be viewed as a parametrised surface using:

$$\Phi(\theta,\phi) = (\cos\phi \sin\theta, \sin\phi \sin\theta, \cos\theta), \qquad \theta \in [0,\pi], \quad \phi \in [0,2\pi]\,.$$

Let us check if it is regular:

$$\partial_\theta \Phi(\theta,\phi) = \begin{pmatrix} \cos\phi\cos\theta \\ \sin\phi\cos\theta \\ -\sin\theta \end{pmatrix}, \quad \partial_\phi \Phi(\theta,\phi) = \begin{pmatrix} -\sin\phi\sin\theta \\ \cos\phi\sin\theta \\ 0, \end{pmatrix}$$

and:

$$\partial_\theta \Phi(\theta,\phi) \times \partial_\phi \Phi(\theta,\phi) = \sin\theta \begin{pmatrix} \cos\phi\sin\theta \\ \sin\phi\sin\theta \\ \cos\theta \end{pmatrix}.$$

As long as $\theta \notin \{0, \pi\}$ (North and south poles), the points are regular.

Remark 2.8.2. This is a situation where the notion of local parametrisation is useful: to catch all points by using different parametrisations. In fact, here, if we want to catch all points in a regular fashion we need several local parametrisations. A sphere is a regular surface in the sense of the local definition. However, for our purposes, which are mainly integration, if we miss a small (i.e. Lebesgue measure 0) number of points in our parametrisation then it does not matter.

Observe that, the vector $\partial_\theta \Phi(\theta, \phi) \times \partial_\phi \Phi(\theta, \phi)$ is normal to the tangent plane of $\Phi(\theta, \phi)$. ◇

Example 2.29. Let $0 < r < R$ then the parametrisation:

$$\Phi(u, v) = ((R - r\cos v)\cos u, (R - r\cos v)\sin u, r\sin v), \qquad u, v \in [0, 2\pi],$$

describes a torus. Observe that, one can eliminate the parameters u, v above and find that these points satisfy the equation:

$$(\sqrt{x^2 + y^2} - R)^2 + z^2 = r^2.$$

Let us see where this parametrisation is regular:

$$\partial_u \Phi(u, v) = \begin{pmatrix} -(R - r\cos v)\sin u \\ (R - r\cos v)\cos u \\ 0 \end{pmatrix}, \quad \partial_v \Phi(u, v) = \begin{pmatrix} r\sin v \cos u \\ r\sin v \sin u \\ r\cos v \end{pmatrix},$$

and:

$$\partial_u \Phi(u, v) \times \partial_v \Phi(u, v) = \begin{pmatrix} r(R - r\cos v)\cos u \cos v \\ r(R - r\cos v)\sin u \cos v \\ -r(R - r\cos v)\sin v \end{pmatrix} = r(R - r\cos v) \begin{pmatrix} \cos u \cos v \\ \sin u \cos v \\ -\sin v \end{pmatrix}.$$

◇

So this is regular everywhere.

Example 2.30. Consider the 1-sheeted hyperboloïd defined by the equation:

$$\{(x, y, z) \in \mathbb{R}^3,\ x^2 + y^2 - z^2 = 1\}$$

then this can be realised as a parametrised surface by setting:

$$\Phi(u, v) = (\cosh u \cos v, \cosh u \sin v, \sinh u).$$

Once more we check for regularity:

$$\partial_u \Phi(u, v) = \begin{pmatrix} \sinh u \cos v \\ \sinh u \sin v \\ \cosh u \end{pmatrix}, \quad \partial_v \Phi(u, v) = \begin{pmatrix} -\cosh u \sin v \\ \cosh u \cos v \\ 0 \end{pmatrix},$$

$$\partial_u \Phi(u,v) \times \partial_v \Phi(u,v) = \begin{pmatrix} -\cosh^2 u \cos v \\ -\cosh^2 u \sin v \\ \sinh u \cosh u \end{pmatrix} = \cosh u \begin{pmatrix} -\cosh u \cos v \\ -\cosh u \sin v \\ \sinh u \end{pmatrix}.$$

This is regular everywhere. ◇

Example 2.31. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a $\mathcal{C}^1$ function, then its graph: $z = f(x,y)$ can be realised as a parametrised surface with parametrisation:

$$\Phi(u,v) = (u,v,f(u,v)).$$

◇

Example 2.32. The plane $x + y + z = 2$ can be described as a parametrised surface with parametrisation:

$$\Phi(u,v) = (2-u-v,u,v),$$

it can also be reparametrised as:

$$\Phi(u,v) = (u,v,2-y-v),$$

and so on. ◇

Remark 2.8.3. I will point out that all of the above examples are also surfaces in the more restricted sense I mentioned above. It is also no accident that many of the sets we considered can also be described (locally) implicitly by an equation or several equations; the implicit function theorem can be used to show that for surfaces in the restricted sense these points of view are equivalent (locally).

Example 2.33. Consider the rectangular region of a plane given by:

$$\Phi(u,v) = (u,v,0), \qquad u,v \in [0,1]$$

$$\partial_u \Phi(u,v) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \partial_v \Phi(u,v) = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Here problems arise by the fact we are not working an open set. According to our language this surface is regular, whilst it is clear it has corners on the boundary. ◇

2.8.2 Tangent spaces

When we discussed surfaces defined by level sets in Section 1.10.2, we defined the tangent (vector) space at a regular point p to be the set of directions in which one can travel from p and remain on the surface. This description is sufficiently geometric to carry over to the present situation verbatim, once more, at regular points. The explicit description using the parametrisation is of course different than before.

Although, if a surface can be described implicitly and parametrically we expect the tangent space to be the same.

The parametrisation Φ offers a direct way to determine the tangent space. Suppose without loss of generality that $p = \Phi(0,0)$, then the coordinate axis $v = 0$, $u = 0$ in the parameter space $U \subset \mathbb{R}^2$, are mapped to curves that lie on the surface (see Figure 2.4), and parametrising them as the curves $t \mapsto \Phi(t,0)$, $t \mapsto \Phi(0,t)$, their tangent vectors at p are the partial derivatives: $\partial_u\Phi(0,0)$ and $\partial_v\Phi(0,0)$. At a regular point, by assumption, they are linearly independent and span a two dimensional vector space. To prove that no other vectors can be tangent to the surface at p requires a discussion that we shall skip (to avoid it entirely, one might just *define* the tangent space to be the vector space spanned by these vectors). We will take for granted that the tangent space at a regular point to a surface is a plane, spanned by the partial derivatives of the parametrisation at that point. Additionally, the vector $\partial_u\Phi(u,v) \times \phi_v\Phi(u,v)$ will be a *normal vector* to the surface at $\Phi(u,v)$. Finally, just as before, the affine tangent plane is defined by translating the vector space from the origin to the point $\Phi(u,v)$.

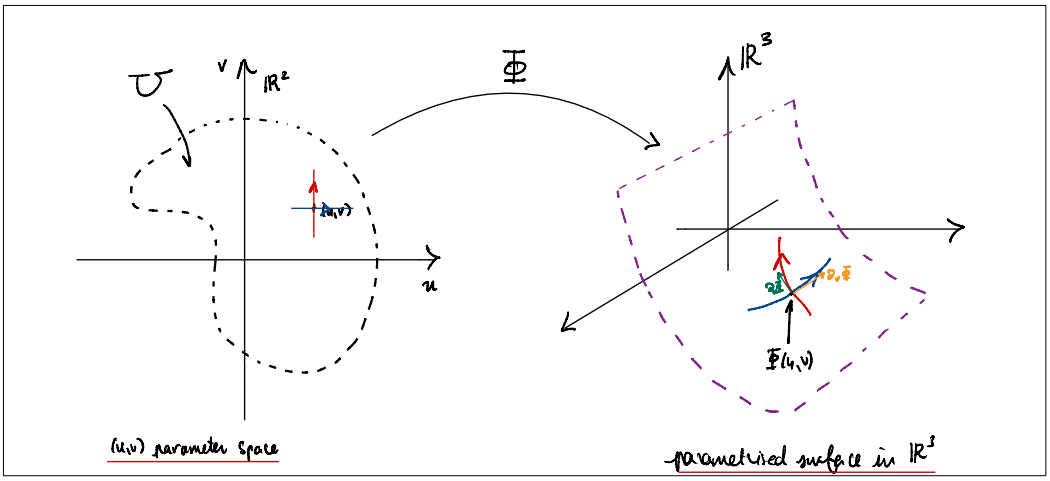


Figure 2.4: Schematic representation of parameter space being mapped to a surface in $\mathbb{R}^3$.

Example 2.34. Let us consider the standard parametrisation of sphere:

$$\Phi(\theta,\phi) = (\cos\phi\sin\theta, \sin\phi\sin\theta, \cos\theta),$$

for which we saw that:

$$\frac{\partial_\theta\Phi(\theta,\phi) \times \partial_\phi\Phi(\theta,\phi)}{\|\partial_\theta\Phi(\theta,\phi) \times \partial_\phi\Phi(\theta,\phi)\|} = \begin{pmatrix}\cos\phi\sin\theta\\ \sin\phi\sin\theta\\ \cos\theta\end{pmatrix} = \Phi(\theta,\phi).$$

Observe that this unit vector does not vanish even at the north or south poles. The affine tangent space is then:

$$\{\, M \in \mathbb{R}^3, \langle \overrightarrow{\Phi(\theta,\phi)M}, \Phi(\theta,\phi)\rangle = 0\},$$

and coincides with the affine tangent space when described implicitly. (Even at the north and south poles !) ◇

Example 2.35. Let $0 < r < R$ and consider the torus parametrised by:

$$\Phi(u,v) = ((R - r\cos v)\cos u, (R - r\cos v)\sin u, r\sin v), \qquad u, v \in\ (0, 2\pi)\,,$$

a unit normal vector at $\Phi(u,v)$ is given by:

$$N(u,v) = \frac{\partial_u\Phi(u,v)\ \times \partial_v\Phi(u,v)}{\|\partial_u\Phi(u,v)\ \times \partial_v\Phi(u,v)\|} = \begin{pmatrix} \cos u\cos v \\ \sin u\cos v \\ -\sin v \end{pmatrix},$$

and the tangent space vector space is the set of vectors that are orthogonal to it. Eliminating the parameters, the normal vector at $(x,y) = \Phi(u,v)$ is given by:

$$N = \begin{pmatrix} -\frac{x}{r\sqrt{x^2+y^2}}\left(\sqrt{x^2+y^2} - R\right) \\ -\frac{y}{r\sqrt{x^2+y^2}}\left(\sqrt{x^2+y^2} - R\right) \\ -\frac{z}{r} \end{pmatrix},$$

this is consistent with computing the gradient of the function $F(x,y,z) = (\sqrt{x^2+y^2} - R)^2 + z^2 - r^2$ (the torus is defined by $F(x,y,z) = 0$, so the gradient will be normal to the surface). The results differ by an overall factor, hence they will lead to equivalent descriptions of the tangent plane. ◇

← End Lecture 18
← Start Lecture 19

2.8.3 The first fundamental form of a surface, area of a surface

The ambient Euclidean space $\mathbb{R}^3$ is equipped with a scalar product, $\langle\cdot,\cdot\rangle$. At a regular point $\Phi(u,v)$ of a parametrised one can restrict this scalar product to vectors of the tangent space, making the tangent space itself into a Euclidean space. This is known as the *first fundamental form* and written I_p. It is a fundamental object that encodes the local geometry of the surface.

Any vector in the tangent space at $p = \Phi(u,v)$ is a linear combination of $\partial_u\Phi(u,v)$ and $\partial_v\Phi(u,v)$. Let:

$$v_1 = \alpha_1\partial_u\Phi + \beta_1\partial_v\Phi, \quad v_2 = \alpha_2\partial_u\Phi + \beta_2\partial_v\Phi,$$

be two vectors in the tangent vector space then:

$$\begin{aligned} I_p(v_1,v_2) &= \langle \alpha_1\partial_u\Phi + \beta_1\partial_v\Phi, \alpha_2\partial_u\Phi + \beta_2\partial_v\Phi\rangle \\ &= (\alpha_1, \beta_1\ ) \begin{pmatrix} \langle\partial_u\Phi, \partial_u\Phi\rangle & \langle\partial_u\Phi, \partial_v\Phi\rangle \\ \langle\partial_v\Phi, \partial_u\Phi\rangle & \langle\partial_v\Phi, \partial_v\Phi\rangle \end{pmatrix} \begin{pmatrix} \alpha_2 \\ \beta_2 \end{pmatrix}. \end{aligned}$$

Therefore we can determine I_p for any vector if we know the *symmetric* matrix:

$$\begin{pmatrix} \langle \partial_u \Phi, \partial_u \Phi \rangle & \langle \partial_u \Phi, \partial_v \Phi \rangle \\ \langle \partial_v \Phi, \partial_u \Phi \rangle & \langle \partial_v \Phi, \partial_v \Phi \rangle \end{pmatrix} = \begin{pmatrix} E & F \\ F & G \end{pmatrix}.$$

By abuse of terminology, the matrix is also referred to as the first fundamental form. Then notations E, F, G are due to Gauss.

As an exercise, we ask the reader to show (using Eq. (0.1))that:

$$\|\partial_u \Phi \times \partial_v \Phi\| = \sqrt{EG - F^2} = \sqrt{\det I_p}.$$

We can now define the area of a parametrised surface:

Definition 2.8: Area of a parametrised surface

Let $\Phi : U \subset \mathbb{R}^2 \to \mathbb{R}^3$ parametrise a surface, then we define:

$$\mathcal{A} = \int_U \sqrt{EG - F^2} du dv.$$

Before we discuss why this is a reasonable definition, we check that it is invariant under reparametrisation and consider a few examples.

First, let $\begin{array}{rccc} \phi : & V \subset \mathbb{R}^2 & \longrightarrow & U \subset \mathbb{R}^2 \\ & (u', v') & \longmapsto & (u, v) \end{array}$ be a reparametrisation of the surface and write $\tilde{\Phi} = \Phi \circ \phi$. By the chain rule, if $\phi(u', v') = (u, v)$, then the change of basis matrix from $(\partial_u \Phi(u, v), \partial_v \Phi(u, v))$ to $(\partial_{u'} \tilde{\Phi}(u', v'), \partial_{v'} \tilde{\Phi}(u', v'))$ is exactly the Jacobian matrix of ϕ at (u', v'). (Why ?)

It then follows that:

$$\begin{pmatrix} E' & F' \\ F' & G' \end{pmatrix} \equiv \begin{pmatrix} \langle \partial_{u'} \tilde{\Phi}, \partial_{u'} \tilde{\Phi} \rangle & \langle \partial_{u'} \tilde{\Phi}, \partial_{v'} \tilde{\Phi} \rangle \\ \langle \partial_{v'} \tilde{\Phi}, \partial_{u'} \tilde{\Phi} \rangle & \langle \partial_{v'} \tilde{\Phi}, \partial_{v'} \tilde{\Phi} \rangle \end{pmatrix} = {}^t\text{Jac } \phi(u', v') \begin{pmatrix} E & F \\ F & G \end{pmatrix} \text{Jac } \phi(u', v').$$

Taking the determinant we conclude that:

$$\sqrt{E'G' - F'^2} = |\text{Jac } \phi(u', v')| \sqrt{EG - F^2}.$$

Hence, using the general change of variable formula:

$$\int_U \sqrt{E'G' - F'^2} du' dv' = \int_U |\text{Jac } \phi(u', v')| \sqrt{EG - F^2} du' dv' = \int_V \sqrt{EG - F^2} du dv.$$

Example 2.36. (Surface area of a sphere of radius R)
Recall that $S^2 = \{(x, y, z) \in \mathbb{R}^3, x^2 + y^2 + z^2 = R^2\}$) and can be parametrised by:

$$\Phi(\theta, \phi) = (R\cos\phi\sin\theta, R\sin\phi\sin\theta, R\cos\theta),$$

and:

$$\partial_\theta\Phi(\theta,\phi) = R\begin{pmatrix}\cos\phi\cos\theta\\ \sin\phi\cos\theta\\ -\sin\theta\end{pmatrix}, \quad \partial_\phi\Phi(\theta,\phi) = R\begin{pmatrix}-\sin\phi\sin\theta\\ \cos\phi\sin\theta\\ 0,\end{pmatrix}$$

So: $E = R^2$, $G = R^2\sin^2\theta$, $F = 0$.

$$\mathcal{A} = R^2\int_0^{2\pi}\int_0^{\pi}\sin\theta d\theta d\varphi = 4\pi R^2.$$

◇

Example 2.37. (Surface area of a torus)
Let $0 < r < R$ then recall that:

$$\Phi(u, v) = ((R - r\cos v)\cos u, (R - r\cos v)\sin u, r\sin v), \qquad u, v \in [0, 2\pi],$$

parametrises a torus. Recall that $\sqrt{EG - F^2} = \|\partial_u\Phi \times \partial_v\Phi\|$ and we computed that:

$$\partial_u\Phi \times \partial_v\Phi = r(R - r\cos v)\begin{pmatrix}\cos u\cos v\\ \sin u\cos v\\ -\sin v\end{pmatrix}$$

so:

$$\sqrt{EG - F^2} = r(R - r\cos v).$$

Hence:

$$\mathcal{A} = \int_0^{2\pi}\int_0^{2\pi} r(R - r\cos v)dudv = 2\pi\int_0^{2\pi} r(R - r\cos v)dv = 4\pi^2 rR.$$

◇

The above computations recover the classical formulae for tori and spheres.

Example 2.38. (Surface area of a graph over an open disk $D = B(0, r) \subset \mathbb{R}^2$) We parametrise this as:

$$\Phi(u, v) = (u, v, f(u, v)), (u, v) \in D,$$

then the first fundamental form is given by:

$$\begin{pmatrix}1 + (\partial_u f)^2 & \partial_u f\partial_v f\\ \partial_u f\partial_v f & 1 + (\partial_v f)^2\end{pmatrix},$$

hence:

$$\sqrt{EG - F^2} = \sqrt{1 + (\partial_u f)^2 + (\partial_v f)^2},$$

and:

$$\mathcal{A} = \int_D \sqrt{1 + (\partial_u f)^2 + (\partial_v f)^2} du dv.$$

◇

Example 2.39. Just as for curves, in applications, it can be interesting to consider surfaces that, although not parametrised surfaces themselves, can be split up into pieces that are parametrised surfaces; this is particularly the case of polygons. To practice, let us parametrise the pieces of a pyramid with a rectangular base. Suppose its apex is at $D = (0, 0, h)$ and the points of the base are $O = (0, 0, 0)$, $A = (a, 0, 0)$, $B = (a, b, 0)$ and $C = (0, b, 0)$ for some positive numbers a, b, h.

- The rectangular basis $OABC$ can be parametrised by:

$$\Phi_1(u, v) = (u, v, 0), \qquad (u, v) \in [0, a] \times [0, b].$$

- The triangular face (OAD) can be parametrised by:

$$\Phi_2(u, v) = ((1 - v)a, (1 - u - v)b, vh), \quad (u, v) \in \{u \geqslant 0, v \geqslant 0, u + v \leqslant 1\} = T$$

- The triangular face (BCD) can be parametrised similarly by:

$$\Phi_3(u, v) = (ua, (u + v)b, (1 - (u + v))h), \quad (u, v) \in \{u \geqslant 0, v \geqslant 0, u + v \leqslant 1\}.$$

The remaining two faces are parametrised in the same fashion.

Let us check that the surface area of the triangle (OAD) (which is of course contained in a plane) coincides with what we expect.

$$\partial_u \Phi_2(u, v) = \begin{pmatrix} 0 \\ -b \\ 0 \end{pmatrix}, \qquad \partial_v \Phi_2(u, v) = \begin{pmatrix} -a \\ -b \\ h \end{pmatrix}.$$

Hence:

$$\partial_u \Phi_2(u, v) \times \partial_v \Phi_2(u, v) = (-bh, 0, -ab),$$

Thus:

$$\begin{aligned} \mathcal{A} = \int_T b\sqrt{a^2 + h^2} du dv &= b\sqrt{a^2 + h^2} \int_0^1 \int_0^{1-v} du dv \\ &= b\sqrt{a^2 + h^2} \int_0^1 (1 - v) du dv \\ &= \frac{b\sqrt{a^2 + h^2}}{2}. \end{aligned}$$

One can also consider the triangular face (BCD), this time we compute, E, F, G (which are all constants):

$$\partial_u \Phi_3(u, v) = (a, b, -h), \qquad \partial_v \Phi_3(u, b) = (0, b, -h),$$

$$E = a^2 + b^2 + h^2, \qquad G = b^2 + h^2, \qquad F = b^2 + h^2,$$

thus:

$$EG - F^2 = (b^2 + h^2)^2 + a^2(b^2 + h^2) - (b^2 + h^2)^2 = a^2(b^2 + h^2).$$

Finally:

$$\mathcal{A} = \int_T a\sqrt{b^2 + h^2}\,du dv = \frac{a\sqrt{b^2 + h^2}}{2}.$$

$\diamond$

← End Lecture 19
← Start Lecture 20

Flux of a vector field through a surface

When we discussed integrals over curves it was possible to identify a general type of object that one could integrate, provided we had specified an orientation along the curve: these were differential (one)-forms. A clear duality:

$$\vec{e}_x \leftrightarrow \langle \vec{e}_x, \cdot \rangle = dx, \qquad \vec{e}_y \leftrightarrow \langle \vec{e}_y, \cdot \rangle = dy, \qquad \vec{e}_z \leftrightarrow \langle \vec{e}_z, \cdot \rangle = dz,$$

between one forms and vector fields enabled us to define the *circulation of a vector field along a curve*.

For surfaces, one could (should?) take a similar path: the appropriate objects to integrate are called differential 2-forms, they look like this[5]:

$$B = B_z dx \wedge dy - B_y dx \wedge dz + B_x dy \wedge dz.$$

Unfortunately, whilst conceptually satisfying, this path would take us on a journey that requires some more linear algebra (for instance, to understand what $\wedge$ means) and hence more time than we can afford at our current pace. Furthermore, although there is a natural way to construct a 2-form from a vector field in 3D, using what is known as the Hodge dual[6], $\star$, explaining these algebraic constructions is, much to my dismay, also a distraction we cannot afford. Fortunately, we can explain the notion of surface integral of a vector field, or flux of a vector field through a surface in a relatively natural way. This notion will depend on orientation (like the integral of differential forms) so before we define the integral let us briefly discuss orientation of surfaces.

Induced orientation of a vector plane in $\mathbb{R}^3$. Recall that $\mathbb{R}^3$ is oriented by the so-called right-hand rule, such that the standard basis $(\vec{e}_x, \vec{e}_y, \vec{e}_z)$ is positive. Using this orientation, and an additional choice, we will be able to equip any vector plane in $\mathbb{R}^3$ with an orientation. Let us consider the plane defined by the equation:

$$ax + by + cz = 0,$$

[5] looks pretty cool, right ?

[6] How this works is suggested by the choice of notation in the above example. More explicitly we will have: $\star(dx \wedge dy) = dz$, $\star(dy \wedge dz) = dx$ and $\star(dx \wedge dz) = -dy$. We will also have $\star^2 = \mathrm{id}$

where $(a, b, c) \neq (0, 0, 0)$. There are two unit normal vectors to this plane,

$$\vec{n}_{\pm} = \pm \frac{1}{\sqrt{a^2 + b^2 + c^2}} \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

Given a choice $\vec{n}$ of one of these unit normal vectors, one can use the ambient orientation of $\mathbb{R}^3$ to deduce an orientation on the plane in the following manner:

> A basis $(\vec{u}, \vec{v})$ of the vector plane is positively oriented if and only if $(\vec{u}, \vec{v}, \vec{n})$ is positively oriented in $\mathbb{R}^3$ (See Figure 2.5).

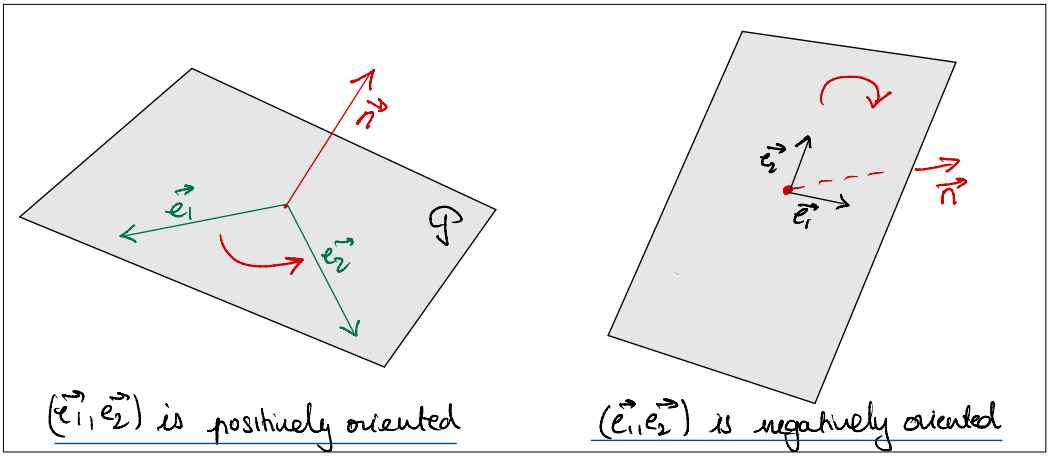


Figure 2.5: Orienting a plane with a choice of normal vector.

Now, at each point of a regular parametrised surface, a choice of unit normal vector will determine an orientation of the tangent plane at that point. However, as was the case with curves, we would like the orientation of these different spaces to be "continuous" as the point moves; this amounts to making choice of normal vector that depends continuously on the point.

For a surface parametrised by[7] $\Phi : U \subset \mathbb{R}^2 \to \mathbb{R}^3$ we have two natural choices that fullfill these conditions:

$$\pm \frac{\partial_u \Phi \times \partial_v \Phi}{\|\partial_u \Phi \times \partial_v \Phi\|}.$$

A choice of either of these normals will determine an *orientation* of the parametrised surface. A parametrised surface with a choice of normal will be said to be *oriented*.

[7] We implicitly assume that U is connected, in other words, in one piece.

Example 2.40. Let us consider the sphere S^2, parametrised by:

$$\Phi(\theta, \phi) = (\cos\phi \sin\theta, \sin\phi \sin\theta, \cos\theta),$$

for which we saw that:

$$\vec{N}(\theta, \phi) = \frac{\partial_\theta \Phi(\theta, \phi) \times \partial_\phi \Phi(\theta, \phi)}{\|\partial_\theta \Phi(\theta, \phi) \times \partial_\phi \Phi(\theta, \phi)\|} = \begin{pmatrix} \cos\phi \sin\theta \\ \sin\phi \sin\theta \\ \cos\theta \end{pmatrix} = \Phi(\theta, \phi).$$

This choice of normal makes the basis $(\partial_\theta \Phi(\theta, \phi), \partial_\phi \Phi(\theta, \phi))$ positively oriented. It is also known as the *outward pointing normal.* ◇

The reader can check that a reparametrisation $\phi : V \to U$ of a parametrised surface will be orientation reversing if Jac $\phi(u, v) < 0$.

Defining the flux of $\vec{F}$ through a surface. Let $\vec{F}$ be a vector field, we would like to measure how much of the vector field is "passing through" a given surface $\mathscr{S}$ (in the direction given by the orientation). In applications, $\vec{F}$ might, for instance, represent a current density. Let us begin our discussion by stating the definition:

Definition 2.9: Surface integral of a vector field

Let $\mathscr{S}$ be an oriented parametrised surface, oriented by a normal vector $\vec{n}$. $\vec{F}$ a vector field in $\mathbb{R}^3$, then we define:

$$\int_{\mathscr{S}} \vec{F} \cdot d\vec{S} = \int_U \langle \vec{F}(\Phi(u, v)), \partial_u \Phi \times \partial_v \Phi \rangle du dv\,,$$

where: $\Phi : U \to \mathbb{R}^3$ is any parametrisation of $\mathscr{S}$ such that $\partial_u \Phi \times \partial_v \Phi$ is in the direction of the normal vector.

Remark 2.8.4. We can also write:

$$\int_{\mathscr{S}} \vec{F} \cdot d\vec{S} = \int_U \langle \vec{F}(\Phi(u, v)), \vec{n} \rangle \|\partial_u \Phi \times \partial_v \Phi\| du dv = \int_U \langle \vec{F}(\Phi(u, v), \vec{n} \rangle \sqrt{EG - F^2} du dv.$$

We leave it as an exercise to the reader to check (using the chain rule and the change of variable formula) that this is invariant under a reparametrisation $\phi : V \to U$ such that Jac $\phi > 0$ but changes sign if Jac $\phi < 0$.

One can observe that if $\vec{F}$ is orthogonal to $\vec{n}$ at each point, then the flux is vanishing. Intuitively, this seems reasonable, no current can flow through the surface if it does not have a transverse component to it.

Example 2.41. Let us consider a radial vector field in $\mathbb{R}^3$ with constant magnitude $|F|$, $\vec{F} = |F|\vec{e}_r$. We recall that:

$$\vec{e}_r = \frac{x}{\sqrt{x^2 + y^2 + z^2}} \vec{e}_x + \frac{y}{\sqrt{x^2 + y^2 + z^2}} \vec{e}_y + \frac{z}{\sqrt{x^2 + y^2 + z^2}} \vec{e}_z.$$

Let us calculate the flux of $\vec{F}$ flowing *out* of a sphere of radius R. This means we should use the *outward pointing normal vector field to* S^2 which is exactly $\vec{e}_r$.

Hence, in this simple case:

$$\int_{S^2} \vec{F} \cdot d\vec{S} = 4\pi R^2 |F|.$$

◇

Example 2.42. Sticking to the sphere, however instead oriented with the *inward pointing normal vector field*, and consider a different vector field:

$$\vec{F} = x\vec{e}_x + y\vec{e}_y,$$

Now the usual parametrisation:

$$\Phi(\theta, \phi) = (\cos\phi \sin\theta, \sin\phi \sin\theta, \cos\theta),$$

is negatively oriented since the vector, since $\frac{\partial_\theta \Phi \times \partial_\phi \Phi}{\|\partial_\theta \Phi \times \partial_\phi \Phi\|} = -\vec{n}$ and:

$$\int_{-S^2} \vec{F} \cdot d\vec{S} = -\int_0^{2\pi} \int_0^{\pi} \sin^3\theta d\theta d\phi = -\frac{8\pi}{3}.$$

◇

Remark 2.8.5. This integral is a good excuse to illustrate a useful tool for integration of trigonometric functions. Here we can linearise $\sin^3\theta$ using de Moivre's Formula. This goes as follows:

$e^{3i\theta} = \cos 3\theta + i \sin 3\theta = (\cos\theta + i\sin\theta)^3 = (\cos^3\theta - 3\cos\theta\sin^2\theta) + i(3\cos^2\theta\sin\theta - \sin^3\theta)$.

Identifying the imaginary parts, we arrive at:

$$\sin 3\theta = 3\cos^2\theta \sin\theta - \sin^3\theta \Leftrightarrow \sin^3\theta = \frac{1}{4}(3\sin\theta - \sin 3\theta),$$

which is easy to integrate.

$$\int_0^{\pi} \sin^3\theta d\theta = \frac{1}{4}\left(\int_0^{\pi} (3\sin\theta - \sin 3\theta) d\theta\right) = \frac{3}{2} - \frac{1}{6} = \frac{4}{3}.$$

2.9 The divergence theorem

We now have the tools to state the next theorem of vector calculus.

Theorem 2.4: (Gauss-Ostrogradsky)

Let $\vec{F}$ be a $\mathcal{C}^1$ vector field and ∂V a closed surface oriented by the *outward pointing normal* and enclosing a bounded connected open set $V \subset \mathbb{R}^3$, then:

$$\int_V \operatorname{div} \vec{F}\, dV = \int_{\partial V} \vec{F} \cdot d\vec{S}.$$

End Lecture 20→
Start Lecture 21→

Example 2.43. Suppose a charge distribution ρ satisfies Maxwell's equations, i.e. $\mathrm{div}\vec{E} = \frac{\rho}{\epsilon_0}$, then the total charge Q is exactly:

$$Q = \int_V \rho dV = \epsilon_0 \int_V \mathrm{div}\vec{E} dV = \epsilon_0 \int_{\partial V} \vec{E} \cdot \, d\vec{S}.$$

This is Gauss' law. ◇

Example 2.44. The divergence theorem is behind a lot of conservation laws, which often take the form:

$$\frac{\partial \rho}{\partial t} + \mathrm{div}\, \vec{j} = 0.$$

In electromagnetism (for definiteness, but this sort of *transport equation* is widespread in applications), ρ is the charge density, and $\vec{j}$ the current density, by the divergence theorem we have:

$$\frac{\partial Q}{\partial t} = \int_V \frac{\partial \rho}{\partial t} dV = -\int_V \mathrm{div}\, \vec{j} = -\int_{\partial V} \vec{j} \cdot d\vec{S},$$

where Q is the total charge. ◇

Example 2.45. Consider:

$$\vec{F}(x, y, z) = xy^2\vec{e}_x + yz^2\vec{e}_y + x^2z\vec{e}_z,$$

use the divergence theorem to calculate:

$$\int_S \vec{F} \cdot d\vec{S},$$

where S is a sphere of radius R centred at the origin and the surface is oriented using the outward pointing normal vector.

In this case, $\mathrm{div}\, \vec{F} = y^2 + z^2 + x^2$, hence by the divergence theorem:

$$\int_B (x^2 + y^2 + z^2) dx dy dz = \int_S \vec{F} \cdot d\vec{S},$$

where $B = \{(x, y, z) \in \mathbb{R}^3, x^2 + y^2 + z^2 \leqslant R\}$. Switching to spherical coordinates:

$$x = r\cos\phi \sin\theta, \quad y = r\sin\phi\sin\theta, \quad z = r\cos\ \theta,$$

we have by the general change of variable formula:

$$\int_B (x^2 + y^2 + z^2) dx dy dz = \int_0^{2\pi} \int_0^{\pi} \int_0^R r^4 \sin\theta dr d\theta d\phi = \frac{4\pi R^5}{5}.$$

Let us set up the surface integral, to see what computations we were able to avoid. Parametrise the sphere, as usual, by:

$$\Phi(\theta, \phi) = (R\cos\phi\sin\theta, R\sin\phi\sin\theta, R\cos\theta),$$

then:

$$\int_S \vec{F} \cdot d\vec{S} = \int_0^{2\pi} \int_0^{\pi} \langle \vec{F}(\Phi(\theta,\phi)), (\partial_\theta \Phi \times \partial_\phi \Phi)(\theta,\phi) \rangle d\theta d\phi.$$

Now:

$$\vec{F}(\Phi(\theta,\phi)) = R^3 \cos\phi \sin^2\phi \sin^3\theta \vec{e}_x + R^3 \sin\phi \sin\theta \cos^2\theta \vec{e}_y + R^3 \cos^2\phi \sin^2\theta \cos\theta \vec{e}_z.$$

Recall that:

$$\partial_\theta \Phi(\theta,\phi) \times \partial_\phi \Phi(\theta,\phi) = R^2 \sin\theta \left(\cos\phi \sin\theta \vec{e}_x + \sin\phi \sin\theta \vec{e}_y + \cos\theta \vec{e}_z\right).$$

It follows that the surface integral is:

$$R^5 \int_0^{2\pi} \int_0^{\pi} \left(\cos^2\phi \sin^2\phi \sin^5\theta + \sin^2\phi \sin^3\theta \cos^2\theta + \cos^2\phi \sin^3\theta \cos^2\theta\right) d\theta d\phi.$$

Plugging this into your favourite computer algebra software, or by using frantically your trigonometry, you can check that this confirms our previous result. ◇

Example 2.46 (Green's identities)**.** Recall that the Laplacian of a $\mathcal{C}^2$ function $f : \mathbb{R}^3 \to \mathbb{R}$ is defined to be $\Delta f = \operatorname{div} \overrightarrow{\operatorname{grad}} f$. Let S be a parametrised surface that encloses a bounded open set V of $\mathbb{R}^3$. Let f, g be two functions, then:

$$\operatorname{div}(g\overrightarrow{\operatorname{grad}} f) = \langle \overrightarrow{\operatorname{grad}} g, \overrightarrow{\operatorname{grad}} f \rangle + g\Delta f,$$

by the product rule for the divergence. Then, by the divergence theorem we then arrive at:

$$\int_S g(\overrightarrow{\operatorname{grad}} f \cdot d\vec{S}) = \int_V \left(g\Delta f + \langle \overrightarrow{\operatorname{grad}} g, \overrightarrow{\operatorname{grad}} f \rangle\right) dV.$$

This is the first of two important identities attributed to Green. Swapping the roles played by f, g and taking the difference we arrive at Green's second identity:

$$\int_S (g \cdot \overrightarrow{\operatorname{grad}} f - f \cdot \overrightarrow{\operatorname{grad}} g) \cdot d\vec{S}) = \int_V (g\Delta f - f\Delta_g)\, dV.$$

◇

2.10 Kelvin-Stokes theorem

In some sense, the divergence theorem is the most immediate generalisation of the Green-Riemann theorem in two dimensions to three dimensions: we relate an integral over a domain in $\mathbb{R}^3$ to an integral over a surface, an object which as one dimension less than the total space. In 3D there is room for another generalisation: can we relate surface integrals over oriented surfaces in $\mathbb{R}^3$ to an integrals over a curve in $\mathbb{R}^3$? The answer is positive and this is the content of the Kelvin-Stokes theorem relating flux of the curls and circulation of vector fields.

Vaguely, the result is going to be:

$$\int_S (\overrightarrow{\text{curl}}\,\vec{F}) \cdot d\vec{S} = \int_{\partial S} \vec{F} \cdot d\vec{l}.$$

Of course, there are some difficulties we need to address in order to interpret correctly this formula:

- What surfaces should be allowed ?
- What is the boundary of a surface?
- What is the orientation on the boundary ?

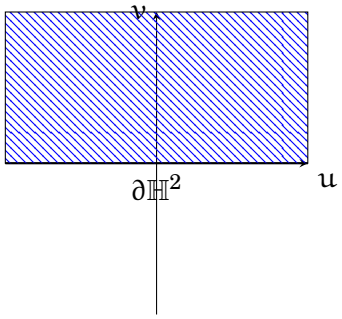
We shall address these issues by means of examples. First of all, let us consider an example which clearly has a boundary (and even corners):

Example 2.47. Let us consider, in $\mathbb{R}^3$, the half plane $y \geqslant 0, z = 0$ parametrised as follows:

$$\Phi : (u, v) \mapsto (u, v, 0), \qquad u \in \mathbb{R}, y \geqslant 0.$$

Whilst, up to now, we have been a bit cavalier with our parameters, (whereas we should have been restricted to an open set), here, we are going to be more careful about what we mean when $y = 0$.

The parameters (u, v) take their values in the half space $\mathbb{H}^2$ of $\mathbb{R}^2$:



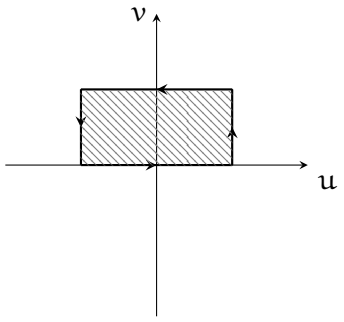
As we saw at the beginning of the course, $\mathbb{H}^2$ certainly has a boundary, which is the x-axis. We will identify the image points of this boundary with the boundary of the parametrised surface.

Now let us test Stoke's theorem in this example in $\mathbb{R}^3$, orienting the surface with the normal in the positive z direction. Take an arbitrary vector field $\vec{F} = F_x\vec{e}_x + F_y\vec{e}_y + F_z\vec{e}_z$ which we assume is zero outside of $[-R, R] \times [-R, R] \times [-R, R]$ for some $R > 0$, then:

$$\int_S \overrightarrow{\text{curl}}\,\vec{F} \cdot d\vec{S} = \int_0^R \int_{-R}^R \left(\frac{\partial F_y}{\partial x}(u, v, 0) - \frac{\partial F_x}{\partial y}(u, v, 0) \right) du dv.$$

This has very nicely taken the form of Green's theorem, where $Q(u,v) = F_y(u,v,0)$ and $P(u,v) = F_x(u,v,0)$, and hence, if we call B the boundary of this rectangle in the (u,v) plane it follows that:

$$\int_S \overrightarrow{\text{curl}}\,\vec{F} \cdot d\vec{S} = \int_B F_x du + F_y dv.$$



Taking R large enough, we can assume $\vec{F}$ vanishes on all the sides of the rectangle save the one intersecting $\partial\mathbb{H}^2$), so the integral is:

$$\int_{\Phi(\partial\mathbb{H}^2)} \vec{F} \cdot d\vec{l}.$$

◇

The following example illustrates that we need to be careful about just thinking of the boundary as the boundary of the parameter domain:

Example 2.48. Recall that the unit sphere is parametrised by:

$$\Phi(\theta,\phi) = (\cos\phi\sin\theta, \sin\phi\sin\theta, \cos\theta),$$

Now, one might allow the parameters θ, ϕ to range as follows $[0,\pi] \times [0,2\pi)$, in order to capture all the points on the sphere, but in this case, we might conclude that half the great circle in the xz plane would be a boundary for the sphere. However, geometrically there is no reason why this should be considered a boundary, and even if it were, it is not a boundary in the intuitive sense: one can cross it. Furthermore, rotating the coordinate axis and reparametrising the sphere we can make half of any great circle into a "boundary", indicating that it is not invariant under reparametrisation and therefore not a geometric property.

Intuitively, the sphere does *not* have a boundary (and mathematically, when things are properly defined, it does not). Accordingly, the integral over the sphere of the curl of any vector field should vanish in our proposed Stoke's theorem. Let us check this, working in spherical coordinates. To make this computation as simple as possible we will work in spherical coordinates in $\mathbb{R}^3$, and we will need a formula for the curl, this is:

Let $\vec{F} = F_r\vec{e}_r + F_\theta\vec{e}_\theta + F_\phi\ \vec{e}_\phi$ be the expression in spherical coordinates of a $\mathcal{C}^1$ vector field, then:

$$\overrightarrow{\operatorname{curl}}\,\vec{F} = \begin{pmatrix}\vec{e}_r & \vec{e}_\theta & \vec{e}_\phi\end{pmatrix}\begin{pmatrix}\frac{1}{r\sin\theta}\left(\frac{\partial(\sin\theta F_\phi)}{\partial\theta} - \frac{\partial F_\theta}{\partial\phi}\right)\\ -\frac{1}{r}\left(\frac{\partial(rF_\phi)}{\partial r} - \frac{1}{\sin\theta}\frac{\partial F_r}{\partial\phi}\right)\\ \frac{1}{r}\left(\frac{\partial(rF_\theta)}{\partial\, r} - \frac{\partial F_r}{\partial\theta}\right)\end{pmatrix}.$$

Now let us orient the *unit* sphere with the outward pointing normal $\vec{e}_r$ and observe that, then:

$$\begin{aligned}\int_S \overrightarrow{\operatorname{curl}}\,\vec{F}\cdot d\vec{S} &= \int_0^{2\pi}\int_0^{\pi}\left(\frac{\partial(\sin\theta F_\phi)}{\partial\theta} - \frac{\partial F_\theta}{\partial\phi}\right)d\theta d\phi\\ &= \int_0^{2\pi}\left[\sin\theta F_\phi\ \right]_0^{\pi}d\phi - \int_0^{\pi}\left[F_\theta\right]_0^{2\pi}d\theta\\ &= 0.\end{aligned}$$

The first integral vanishes honestly, and the second vanishes because, at fixed θ, $\phi = 0$ and $\phi = 2\pi$ define the same point (and F is continuous).

Hence, Stoke's theorem is consistent with the intuition that a sphere has no boundary. ◇

The first example inspires the following more general situation:

Example 2.49. Let $D \subset U$ be a bounded domain (with boundary ∂D) to which Green's theorem can be applied contained in an open set of $\mathbb{R}^2$ and $\Phi : U \to \mathbb{R}^3$ an injective differentiable piecewise $\mathcal{C}^2$ map. Let $S = \Phi(D)$ be the parameterised surface, and convene that $\partial S = \Phi(\partial D)$, and the orientation of S is defined by the unit normal determined from Φ, then:

$$\int_S (\overrightarrow{\operatorname{curl}}\,\vec{F})\cdot d\vec{S} = \int_{\partial S}\vec{F}\cdot d\vec{l},$$

holds, where ∂S is given the induced orientation.

Let us establish this in a simple case:

Since $\overrightarrow{\operatorname{curl}}$ is linear, we can begin with $\vec{F} = F_x\vec{e}_x = F\vec{e}_x$, then $\overrightarrow{\operatorname{curl}}\,\vec{F} = -\frac{\partial F_x}{\partial y}\vec{e}_z + \frac{\partial F_x}{\partial z}\vec{e}_y$. Let us write $\Phi = (\Phi_x,\ \Phi_y, \Phi_z)$, then:

$$\begin{aligned}\int_S \overrightarrow{\operatorname{curl}}\,\vec{F}\cdot d\vec{S} &= \int_D\left[-\frac{\partial F}{\partial y}\left(\partial_u\Phi_x\partial_v\Phi_y - \partial_u\Phi_y\partial_v\Phi_x\right) - \frac{\partial F}{\partial z}\left(\partial_u\Phi_x\partial_v\Phi_z - \partial_u\Phi_z\partial_v\Phi_x\right)\right]du dv\\ &= \int_D\left[-\frac{\partial(F\circ\Phi)}{\partial v}\partial_u\Phi_x + \frac{\partial(F\circ\Phi)}{\partial u}\partial_v\Phi_x\right]du dv\\ &= \left[-\frac{\partial(F\circ\Phi\partial_u\Phi_x)}{\partial v} + \frac{\partial(F\circ\Phi\partial_v\Phi_x)}{\partial u}\right]du dv.\end{aligned}$$

For the second equality, we use the chain rule, and the last equality we use the product rule with the symmetry of partial derivatives.

Now, by assumption, we can apply Green's theorem in the (u, v) plane to obtain:

$$\int_S \overrightarrow{\text{curl}}\, \vec{F} \cdot d\vec{S} = \int_{\partial D} (F \circ \Phi)\partial_u \Phi_x du + (F \circ \Phi)\partial_v \Phi_x dv.$$

Let us assume that $\partial D = \bigcup_{i=1}^n \partial D_i$ where ∂D_i are simple regular curves, and suppose that $\gamma_i : [0,1] \to \mathbb{R}^3$ parametrises ∂D_i, oriented so that ∂D is travelled along counterclockwise. In this case, $\partial S = \bigcup_{i=1}^n \Phi(\partial D_i)$ and $c_i(t) = (\Phi \circ \gamma_i)(t) = (x_i(t), y_i(t), z_i(t))$ parametrises $\Phi(\partial D_i)$. Furthermore, by the chain rule:

$$(\Phi \circ \gamma_i)' = ((\partial_u \Phi) \circ \gamma_i)u_i' + ((\partial_v \Phi) \circ \gamma_i)v_i',$$

if $\gamma_i = (u_i, v_i)$, but then:

$$\begin{aligned}
&\int_{\partial D} (F \circ \Phi)\partial_u \Phi_x du + (F \circ \Phi)\partial_v \Phi_x dv \\
&\qquad = \sum_{i=1}^n \int_0^1 \left((F \circ \Phi)(\gamma_i(t))\partial_u \Phi_x(\gamma_i(t))u'(t) + (F \circ \Phi)(\gamma_i(t))\partial_v \Phi_x(\gamma_i(t))v'(t) \right) dt \\
&\qquad\qquad = \sum_{i=1}^n \int_0^1 F(c(t))x_i(t)dt = \int_{\partial D} \vec{F} \cdot d\vec{l}.
\end{aligned}$$

Similar computations for the other components lead to the Stoke's formula in this case. ◇

End Lecture 21→

Observe that Stoke's theorem recovers the fact that if $\vec{F} = \overrightarrow{\nabla} f$ for some function f then the integral over a closed curve is vanishing.

Example 2.50. An application in physics, recall from Maxwell's equations that: $\overrightarrow{\text{curl}}\, \vec{B} = \mu_0(\vec{j} + \varepsilon_0 \frac{\partial \vec{E}}{\partial t})$, then:

$$\int_S \left(\mu_0 \vec{j} + \frac{1}{c^2} \frac{\partial \vec{E}}{\partial t} \right) \cdot d\vec{S} = \int_{\partial S} \vec{B} \cdot d\vec{l},$$

linking flux of currents to the circulation of the magnetic field. ◇

Now for a computational example:

Example 2.51. Consider the cylinder in $\mathbb{R}^3$ defined by the equation: $x^2 + y^2 = r^2$, for some $r > 0$. We will work in cylindrical coordinates (ρ, θ, z):

$$x = \rho \cos\theta, \quad y = \rho \sin\theta.$$

Observe that the equation of the cylinder in these coordinates is just $\rho = r$. We remind the reader that there is a standard (moving) orthonormal basis $(\vec{e}_\rho, \vec{e}_\theta, \vec{e}_z)$ associated with these coordinates where:

$$\vec{e}_\rho = \cos\theta \vec{e}_x + \sin\theta \vec{e}_y, \qquad \vec{e}_\theta = -\sin\theta \vec{e}_x + \cos\theta \vec{e}_y.$$

Given the geometry of the problem it will be useful for us to know the expression of the $\overrightarrow{\text{curl}}$ in this case:

Consider a vector field $\vec{F}$ whose expression in *cylindrical* coordinates is given by:

$$\vec{F} = F_\rho \vec{e}_\rho + F_\theta \vec{e}_\theta + F_z \vec{e}_z,$$

then:

$$\overrightarrow{\text{curl}}\, \vec{F} = \begin{pmatrix} \vec{e}_\rho & \vec{e}_\theta & \vec{e}_z \end{pmatrix} \begin{pmatrix} \frac{1}{\rho}\frac{\partial F_z}{\partial \theta} - \frac{\partial F_\theta}{\partial z} \\ -\left(\frac{\partial F_z}{\partial \rho} - \frac{\partial F_\rho}{\partial z}\right) \\ \frac{1}{\rho}\left(\frac{\partial (\rho F_\theta)}{\partial \rho} - \frac{\partial F_\rho}{\partial \theta}\right). \end{pmatrix}$$

We will compute the (curved) surface area of a finite portion S of this cylinder, say between $z = -\frac{h}{2}$ and $z = \frac{h}{2} > 0$, using Stoke's theorem. For this, we consider the vector field: $\vec{F} = -z\vec{e}_\theta$, then, according to the above formula:

$$\overrightarrow{\text{curl}}\, \vec{F} = \vec{e}_\rho - \frac{z}{\rho}\vec{e}_z.$$

Orienting S using the outward pointing unit normal vector, i.e. $\vec{e}_\rho$, we will have:

$$\int_S \vec{F} \cdot d\vec{S} = \int_S dS = \mathcal{A}.$$

Intuitively, the boundary of the portion S of the cylinder is two disjoint circles on either end; let us confirm that this is consistent with Stoke's theorem.

Now the cylinder is readily parametrised by:

$$\Phi(\theta, z) = (r\cos\theta, r\sin\theta, z), \theta \in (0, 2\pi), z \in \mathbb{R}.$$

Since $\partial_\theta \Phi = r\vec{e}_\theta$ and $\partial_z \Phi = \vec{e}_z$, we conclude that the normal in this parametrisation is compatible with our choice of orientation on the cylinder. The orientation of the curve for Stoke's theorem is always given by the outward pointing normal vector to the curve in the surface. Here, for the circle $\mathscr{C}_1$ in the $z = \frac{h}{2}$ this is $\vec{e}_z$ and this means we need to go around the circle in a *clockwise motion* around the z-axis. However, for the circle $\mathcal{C}_2$ in the $z = -\frac{h}{2}$ plane it is now $-\vec{e}_z$ and so we need to go the other way around ! This is confirmed by the computation:

$$\int_{\partial S} \vec{F} \cdot \, d\vec{l} = \int_{\mathscr{C}_1} \vec{F} \cdot d\vec{l} + \int_{\mathscr{C}_2} \vec{F} \cdot d\vec{l}.$$

$$\int_{\mathscr{C}_1} \vec{F} \cdot d\vec{l} = \int_{\mathscr{C}_1} \frac{zy}{\sqrt{x^2+y^2}} dx - \frac{zx}{\sqrt{x^2+y^2}} dy = -\frac{h}{2} r \int_0^{2\pi} (-\sin^2\theta - \cos^2\theta) d\theta = \pi r h,$$

here the minus sign comes from the orientation. Next:

$$\int_{\mathscr{C}_2} \vec{F} \cdot d\vec{l} = \int_{\mathscr{C}_2} \frac{zy}{\sqrt{x^2+y^2}} dx - \frac{zx}{\sqrt{x^2+y^2}} dy = \frac{-h}{2} r \int_0^{2\pi} (-\sin^2\theta - \cos^2\theta) d\theta = \pi r h,$$

and here the minus sign comes from $z = -\frac{h}{2}$. The sum is $2\pi r h$, which is the expected result for the surface area. ◇

The orientation of the boundary can be explained in similar terms as that of the surface. Recall that choosing a normal vector enabled us to endow each of the tangent planes with an orientation in a systematic matter. For Stoke's theorem, we must do this again: inside each tangent plane, we must specify an orientation of the curve. For this we always use the outward pointing normal vector $\vec{n}$ to the boundary, within the surface, and declare that the positive unit tangent vector $\vec{t}$ (the way in which we should move) is the one such that $(\vec{n}, \vec{t})$ is a positive basis for the orientation of the surface.

Example 2.52. In the above example, the boundary is also the boundary to another surface. Namely the "surface" S formed by the two disjoint discs in the $z = -\frac{h}{2}$ and $z = \frac{h}{2}$. We can orient the top disk (D_1) with the upward pointing normal vector, and the bottom disk (D_2) with the downward pointing normal vector.

Remark 2.10.1. This is a consistent choice of orientation because the surface splits up into two disjoint pieces.

We can check that the surface integral of $\overrightarrow{\text{curl}}\, \vec{F}$ through this surface is almost the same as above.

$$\begin{aligned}\int_S \overrightarrow{\text{curl}}\, \vec{F} \cdot d\vec{S} &= \int_{D_1} \overrightarrow{\text{curl}}\, \vec{F} \cdot d\vec{S} + \int_{D_2} \overrightarrow{\text{curl}}\, \vec{F} \cdot d\vec{S} \\ &= \int_0^{2\pi} \int_0^r (-\frac{h}{2\rho} \rho d\rho d\theta) + \int_0^{2\pi} \int_0^\rho -\frac{-h}{2\rho} (-\vec{e}_z \cdot \vec{e}_z) \rho d\rho d\theta = -2\pi r h.\end{aligned}$$

This illustrates that in order to have the same integral, the surface should have been oriented in the "opposite" way to have the same boundary with the same orientation. In this example, the orientation of the boundary need for Stoke's theorem has been reversed. ◇

Let us consider one last example:

Example 2.53. Let us consider a portion of a torus with parameters r, R parametrised by:

$$\Phi(u, v) = ((R - r\cos v)\cos u, (R - r\cos v)\sin u, r\sin v), \qquad u, v \in [0, 2\pi],$$

described when u varies between 0 and $\frac{\pi}{2}$. Recall that the unit normal, determined from the parametrisation, is given by:

$$\vec{N}(u,v) = \begin{pmatrix} \cos u \cos v \\ \sin u \cos v \\ -\sin v \end{pmatrix},$$

a drawing will convince you that this is the *inward* pointing normal vector. Define, $\vec{F} = -2y\vec{e}_x$ so that $\overrightarrow{\text{curl}}\,\vec{F} = 2\vec{e}_z$. The flux can be computed directly, we have:

$$\int_S \overrightarrow{\text{curl}}\,\vec{F} \cdot d\vec{S} = -2\int_0^{\frac{\pi}{2}} \int_0^{2\pi} r(R - r\cos v)\sin v \, du dv = 0.$$

The boundary of S is again two disjoint circles in the xz and yz planes respectively. For the application of Stoke's theorem, the first one must be travelled along with the direction of the tangent vector given by $-\partial_v \Phi(u = 0, v)$ and the second must be travelled along with tangent vector $+\partial_v \Phi(u = \frac{\pi}{2}, v)$.

We compute now the circulation of $\vec{F}$ along this boundary:

$$\begin{aligned} \int_{\partial S} \vec{F} \cdot d\vec{l} &= \int_{\mathscr{C}_1} \vec{F} \cdot d\vec{l} + \int_{\mathscr{C}_2} \vec{F} \cdot d\vec{l}, \\ &= \underbrace{0}_{y = 0 \text{ on this disk}} + \underbrace{0}_{dx = 0 \text{ on this disk}} \end{aligned}$$

◇

2.11 Integrals of functions over curves and surfaces

Up to now, we have mainly been concerned with the oriented integrals over our curved objects that are required to state the three integral theorems of vector calculus we have studied.

There are other types of geometric objects that can naturally be integrated over curves and surfaces (oriented or not), these objects are called *densities*. In fact, orientation provides an identification between differential n-forms, where n is the dimension of your space, and densities. The space of densities is locally finitely generated of rank 1: in other words locally any density is of the form $f\omega$ where f is a function and ω a non-vanishing density. Using the geometric objects at our disposal, there is a canonical choice for ω enabling us to identify densities with functions and providing a natural way to integrate functions.

Definition 2.10

Let $f : \mathbb{R}^3 \to \mathbb{R}$ be a (continuous) function, $\mathscr{C}$ a (piecewise) $\mathcal{C}^1$ simple regular curve (not necessarily closed), parametrised by $\gamma : [a, b] \to \mathbb{R}^3$, and $\mathscr{S}$ a regular parametrised surface parametrised by $\Phi : U \to \mathbb{R}^3$, ($U$ bounded) then we define:

$$\int_{\mathscr{C}} f dl = \int_a^b f(\gamma(t))\|\gamma'(t)\| \, dt, \qquad \int_{\mathscr{S}} f dS = \int_U f(\Phi(u, v))\sqrt{EG - F^2} du dv.$$

With these definitions, the length and the area of the surface amount to integrating the function 1. The reader should check, using the change of variable formula, that these are independent of parametrisation *and* orientation.

Remark 2.11.1. One might remember: $dl = \|\gamma'(t)\|dt$ and $dS = \sqrt{EG - F^2} du dv$.

Observe that we can rewrite line and surface integrals in these terms:

$$\int_{\mathscr{C}} \vec{F} \cdot d\vec{l} = \int_{\mathscr{C}} (\vec{F} \cdot \vec{t}) dl,$$

where $\vec{t}$ is the positive unit tangent vector to the oriented curve $\mathscr{C}$,

$$\int_{\mathscr{S}} \vec{F} \cdot d\vec{S} = \int_{\mathscr{S}} (\vec{F} \cdot \vec{n}) dS,$$

where $\vec{n}$ is the positive unit normal vector to the oriented surface $\mathscr{S}$.

Chapter 3

Fourier series

3.1 Introduction and motivating example

The final topic we cover in this course takes us away from geometry and into the realm of analysis, and more specifically methods of solving partial differential equations. The particular method we are going to study arose out of the study of two distinct physical problems: heat propagation in a finite medium, governed by the heat equation:

$$\frac{\partial u}{\partial t} = \Delta u,$$

and the movement of a vibrating string of length L, fixed at both ends, governed by the wave equation:

$$\frac{\partial^2 u}{\partial t^2} = \frac{T}{\mu}\frac{\partial^2 u}{\partial x^2}.$$

In the above, μ stands for the mass per unit length of the string and T is tension of the string. The quantity $c = \sqrt{\frac{T}{\mu}}$ has the units of speed and will the propagation speed of a wave.

Let us consider this second equation and try to find solutions, in a naïve fashion. Assuming some knowledge of basic ODEs, we might try to make some assumptions that bring us closer to ODE's, one such assumptions is that we can separate the variables:

$$u(x,t) = f(x)g(t).$$

Plugging this into the equation we find that:

$$g''(t)f(x) = c^2 f''(x)g(t),$$

dividing by u, we find then that:

$$\frac{1}{c^2}\frac{g''(t)}{g(t)} = \frac{f''(x)}{f(x)}.$$

Since the both sides of the equation depend only on different variables, we conclude that they must be equal to a constant λ and:

$$\begin{cases} g''(t) = c^2 \lambda g(t), \\ f''(x) = \lambda f(x). \end{cases}$$

Let us assume that $\lambda = k^2$, then the solutions are:

$$g(t) = c_1 \cos(ckt) + c_2 \sin(ckt), \quad f(x) = c_3 \cos(kx) + c_4 \sin(kx).$$

Some basic trigonometry can then be used to transform this into:

$$u(x,t) = g(t)f(x) = C_1 \cos(k(ct-x)) + C_2 \cos(k(ct+x)) + C_3 \sin(k(ct-x)) + C_4 \sin(k(ct+x)).$$

Now, our string is of length L and assumed fixed at both ends which means that:

$$\forall t \in \mathbb{R}, u(t,0) = 0 = u(t,L),$$

applying this for $t = 0$ we find that: $C_1 + C_2 = 0$, taking the derivative with respect to t and evaluating at $t = 0$ we find also that $C_3 + C_4 = 0$ and we can rewrite the solution to be of the form:

$$u(t,x) = (A \cos(kct) + B \sin(kct)) \sin(kx).$$

Using the same condition at L we conclude that:

$$A \sin(kL) = 0 \text{ and } B \sin(kL) = 0.$$

Therefore, in order for there to be a non-trivial solution, we must impose a condition on k:

$$k = \frac{\pi n}{L}, n \in \mathbb{Z}.$$

This method has therefore lead us to find an infinity of solutions to our problem. In fact, we can restrict to $n \in \mathbb{N}$, because by the parity properties of sin and cos any solution with $-n$ can be written in terms of n. Since the equation is linear, we can add solutions together to obtain new ones... In this way, we come to the idea that the series:

$$\sum_{n \in \mathbb{N}} (A_n \cos(\frac{n\pi c}{L} t) + B_n \sin(\frac{n\pi c}{L} t)) \sin(\frac{n\pi x}{L}),$$

is a solution.

J. Fourier's insight was that in fact there is a sense in which this is the general form of a solution and additionally any periodic function can be decomposed in this way.

3.2 Fourier coefficients

To simplify the presentation, we are going to consider functions with values in $\mathbb{C}$; this will allow us to work with the functions $x \mapsto \exp(inx)$.

All we need to know about $\mathbb{C}$, if you've never encountered it before, is that its elements admit a unique decomposition:

$$z = a + ib, a, b \in \mathbb{R},$$

where i is the imaginary unit and has the curious property:

$$i^2 = -1.$$

The rules of algebra remain the same as in $\mathbb{R}$, appending this rule for the number i.

We define the modulus of $z \in \mathbb{C}$ to be $|z| = \sqrt{a^2 + b^2}$ (observe that it is identical to the Euclidean norm) and the conjugate of $z = a + ib$ to be $\bar{z} = a - ib$. The reals a and b are called the real and imaginary parts of z respectively. If the reader is unfamiliar with $\mathbb{C}$ then it might be useful to check that:

$$z\bar{z} = |z|^2, \quad z^{-1} = \frac{\bar{z}}{|z|^2} \quad (z \neq 0).$$

The theory of series passes over to $\mathbb{C}$ without modification, in particular we can extend the exponential function to all complex numbers:

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!},$$

and we have: $e^{z+z'} = e^z e^{z'}$.

The reader can check that:

$$e^{ix} = \cos(x) + i\sin(x).$$

In particular:

$$e^{i2\pi n} = 1, \qquad e^{i\pi} = -1, \qquad e^{in\pi} = (-1)^n, \qquad e^{i\frac{\pi}{2}} = i$$

The set of all complex valued continuous functions 2π-periodic functions on $\mathbb{R}$, $C^0_{2\pi}(\mathbb{R})$, is easily endowed with a vector space structure:

$$\begin{cases} (f+g)(x) = f(x) + g(x), \\ (\lambda f)(x) = \lambda f(x). \end{cases}$$

Remark 3.2.1. We can in fact identify it with the set of continuous functions defined on a circle. This point of view is useful from a theoretical perspective (a circle is compact), but we will not insist too much upon it.

It can also be endowed with a scalar product:

$$\langle f, g\rangle = \frac{1}{2\pi}\int_0^{2\pi} f(t)\bar{g}(t)dt,$$

where $\bar{g}(t)$ denotes the complex conjugate of $g(t)$.

$$z = a + ib \Rightarrow \bar{z} = a - ib.$$

Let us compute this on some examples; for instance:

$$f(t) = e^{int}, \quad g(t) = e^{imt}.$$

where $n, m \in \mathbb{Z}$. Then:

$$\frac{1}{2\pi}\int_0^{2\pi} f(t)\bar{g}(t)dt = \frac{1}{2\pi}\int_0^{2\pi} e^{i(n-m)t}dt.$$

If $n = m$ then the integral is equal to 1, if not, we can integrate:

$$\frac{1}{2\pi}\int_0^{2\pi} e^{i(n-m)t}dt = \frac{1}{2\pi i(n-m)}\left[e^{i(n-m)t}\right]_{t=0}^{t=2\pi} = 0.$$

So overall:

$$\langle e^{int}, e^{imt}\rangle = \begin{cases} 1 & \text{if } n = m \\ 0 & \text{otherwise.} \end{cases}$$

It looks just like an orthonormal basis in $\mathbb{R}^n$. In fact, it is also true here, the precise statement is as follows.

Theorem 3.1

Let $L^2_{2\pi}(\mathbb{R})$ denote the completion of $C^0_{2\pi}(\mathbb{R})$ for the norm $\|f\|_2 = \sqrt{\langle f, f\rangle}$, then $(e^{inx})_{n\in\mathbb{Z}}$ is an orthonormal basis in the sense that any $f \in L^2_{2\pi}(\mathbb{R})$ can be written as a series:

$$\sum_{n\in\mathbb{Z}} c_n e^{inx},$$

which converges *in the sense of the norm* $\|\cdot\|_2$. The coefficients c_n are precisely:

$$c_n = \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-int}dt,$$

and called the *Fourier coefficients* of f.

Remark 3.2.2. The proof of this result relies on a deep theorem known as the Stone-Weierstrass theorem. The abstract space $L^2_{2\pi}(\mathbb{R})$ contains more than just continuous functions. In particular, we can calculate the Fourier coefficients of piecewise continuous functions too !

If a function f is *real-valued*, then $\bar{f} = f$, but by linearity, $\overline{c_n} = c_{-n}$. Hence, the series can be rewritten in the form:

$$\sum_{n\in\mathbb{Z}} c_n e^{inx} = c_0 + \sum_{n\in\mathbb{N}^*} \left(c_n e^{inx} + c_{-n} e^{-inx}\right) = c_0 + \sum_{n\in\mathbb{N}^*} \left(c_n e^{inx} + \overline{c_n} e^{-inx}\right).$$

Setting $c_n = \frac{1}{2}a_n - i\frac{1}{2}b_n$ we have:

$$\sum_{n\in\mathbb{N}^*} \left(a_n \cos(nx) + b_n \sin(nx)\right),$$

and we obtain for $n \geqslant 1$

$$a_n = \frac{1}{\pi}\int_0^{2\pi} f(t)\cos(nt)dt, \qquad b_n = \frac{1}{\pi}\int_0^{2\pi} f(t)\sin(nt)dt.$$

These are also used called the Fourier coefficients of f. Computationally, c_n is often easier to work with which is why I have preferred it to a_n and b_n.

Example 3.1. Let us check that things are coherent by considering the case of $\cos(x) = \frac{1}{2}e^{ix} + \frac{1}{2}e^{-ix}$. We have:

$$c_n = \frac{1}{4\pi}\int_0^{2\pi} e^{i(1-n)x}dx + \frac{1}{4\pi}\int_0^{2\pi} e^{-(n+1)x}dx.$$

Using our above computation for $x \mapsto \exp inx$ we find that this is non vanishing and equal to $\frac{1}{2}$ if and only if $n = \pm 1$ and 0 otherwise. ◇

Example 3.2. Calculate the Fourier coefficients of the 2π-periodic function such that: $f(x) = x$ for $x \in [0, 2\pi)$.

$$c_n = \frac{1}{2\pi}\int_0^{2\pi} te^{-int}dt.$$

If $n = 0$, $c_0 = \pi$. When $n \neq 0$, we integrate by parts:

$$\int_0^{2\pi} te^{-int}dt = \left[\frac{1}{-in}te^{-int}\right]_0^{2\pi} + \frac{1}{in}\int_0^{2\pi} e^{int}dt = \frac{2\pi i}{n},$$

hence:

$$\begin{cases} c_0 = \pi, \\ c_n = \frac{i}{n}. \end{cases}$$

The Fourier series is therefore:

$$\pi + \sum_{n\in\mathbb{N}^*} \frac{i}{n} e^{inx} = \pi - \sum_{n\geqslant 1} 2\frac{\sin(nx)}{n}.$$

Observe that *if* the series converges pointwise (i.e. for a fixed x) then taking $x = \frac{\pi}{2}$ we could calculate the series:

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1}.$$

In fact, we can. ◇

3.3 Pointwise convergence

There is a subtlety in the above discussion... the Fourier series converges in the sense of the norm $\|\cdot\|_2$, but a priori nothing guarantees that we have the *numerical* convergence:

$$f(x) = \sum_{n=-\infty}^{+\infty} c_n e^{inx} \,!$$

So, although economical and efficient, our current exposition apparently has a major pitfall!

It turns out that the answer to this question is less satisfying than one might have hoped.

There is a continuous function whose Fourier series diverges for at one point (at least).

On the other hand, a hard theorem, known as Carleson's theorem, establishes that:

Theorem 3.2: Carleson

The Fourier series of a continuous periodic function f converges pointwise to the function f *almost everwhere*. This means that there is a set $\mathcal{N}$, of Lebesgue measure 0, such that:

$$\forall x \in \mathbb{R} \setminus \mathcal{N}, \quad f(x) = \sum_{n\in\mathbb{Z}} c_n e^{inx},$$

Strengthening the regularity assumptions on the function f leads to nicer results, the following result is a special case of the Dirichlet-Dini criterion:

Theorem 3.3: Simplified Dirichlet

Let f be a piecewise differentiable 2π-periodic function with finite left and right limits at every point in a period, then:

$$\frac{f(x^+) + f(x^-)}{2} = \sum_{n\in\mathbb{Z}} c_n e^{inx}.$$

In particular, when f is continuous at x, it converges to $f(x) = f(x^+) = f(x^-)$.

Remark 3.3.1.

$$f(x^+) = \lim_{\substack{y\to x\\ y>x}} f(x), \qquad f(x^-) = \lim_{\substack{y\to x\\ y<x}} f(x_-).$$

3.4 Parseval's theorem

In $\mathbb{R}^n$ (or any n-dimensional Euclidean space), when we decompose a vector $x = (x_1, \dots, x_n)$ onto an orthonormal basis, we have the following relationship:

$$(x|x)^2 = \sum_{i=1}^{n} x_i^2.$$

This has a direct generalisation to Hilbert spaces (of which $L^2_{2\pi}(\mathbb{R})$ is an example), and is known as Parseval's identity:

Theorem 3.4: Parseval's identity

$$\langle f, f\rangle = \frac{1}{2\pi}\int_0^{2\pi} |f(x)|^2 dx = \sum_{n\in\mathbb{Z}} |c_n|^2.$$

When f is real-valued, we can express this in terms of a_n and b_n:

$$\frac{1}{2\pi}\int_0^{2\pi} |f(x)|^2 dx = c_0 + \frac{1}{2}\sum_{n=1}^{\infty}(a_n^2 + b_n^2).$$

Example 3.3. The Fourier coefficients of the 2π-periodic function defined by $f(x) = x, x \in [0, 2\pi)$ are given by:

$$\begin{cases} c_0 = \pi \\ c_n = \frac{i}{n}. \end{cases}$$

By Parseval's theorem we have:

$$\frac{4\pi^2}{3} = \frac{1}{2\pi}\int_0^{2\pi} x^2 dx = \pi^2 + 2\sum_{n=1}^{\infty}\frac{1}{n^2},$$

from which we deduce that:

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

◇

3.5 Fourier series and derivation

If $f : \mathbb{R} \to \mathbb{C}$ is 2π-periodic and if class $\mathcal{C}^k$ ($k \geqslant 1$), the chain rule shows that all of its derivatives are also periodic functions. Let us denote by $c_n(f)$ the Fourier coefficients of a function f then by, integration by parts:

$$c_n(f') = \frac{1}{2\pi} \int_0^{2\pi} f'(x) e^{-int} = inc_n(f).$$

By induction, we conclude that:

$$c_n(f^{(k)}) = (in)^k c_n(f).$$

This shows that there is link between regularity of the function f, and the decreasing behaviour (which improves) with regularity.

It also shows that representing f as a Fourier series, the derivation operation is replaced by a multiplication operator on each coefficient. In the study of PDEs, this can be a useful tool to "separate variables" and replace the PDE with an infinite amount of ODEs (on each coefficient). As an example, let us consider the heat equation on a circle; for this we can consider that u is a 2π-periodic function in the variable x.

$$\partial_t u(t,x) = \frac{d^2}{dx^2} u(t,x).$$

Expanding (formally) into Fourier series in the x variable:

$$u(t,x) = \sum_{n \in \mathbb{Z}} c_n(t) e^{-inx} dt,$$

the PDE becomes:

$$\sum_{n \in \mathbb{Z}} c_n'(t) + n^2 c_n(t) = 0.$$

The only $L^2_{2\pi}(\mathbb{R})$- function with vanishing Fourier coefficients is the function 0, so this equation will amount to:

$$\forall n \in \mathbb{Z}, \qquad c_n'(t) + n^2 c_n(t) = 0,$$

therefore:

$$c_n(t) = \gamma_n e^{-n^2 t}.$$

For some constant γ and a solution may be represented as follows, $t \geqslant 0$

$$u(t,x) = \sum_{n\in\mathbb{Z}} \gamma_n e^{-n^2 t} e^{inx}.$$

If we impose an initial condition: $u(0,x) = u_0(x)$ where u_0 is a 2π-periodic function with Fourier coefficients $c_n(f_0)$, then:

$$u(t,x) = \sum_{n\in\mathbb{Z}} c_n(f_0) e^{-n^2 t} e^{inx}.$$

This a priori formal object is in fact an "honest" solution to the equation for $t \geqslant 0$.

3.6 Some additional topics

Appendix A

Appendix

A.1 The difference between f and $f(x)$

There is a conceptual difference between a *function* f and the *value* of a function at a point x which we denote by $f(x)$. When people say things like: "let us consider the *function* $f(x)$", you should throw your shoes at them or/and any other object at your immediate disposal!

Not distinguishing between these concepts can cause mountains of confusion when doing mathematics, even if you are just trying to work out what other people mean (especially physicists who are very economical with notation).

You can think of f as a *rule* that tells you, given some x, how to calculate the value $f(x)$. We write this $f : x \mapsto f(x)$. Here, x is some element of the domain and $f(x)$ is an element of the target domain.

So when we say that a matrix A determines a linear map u, we mean that given the matrix A we can work out the rule to calculate $u(X)$ for any $X \in \mathbb{R}^n$, in this case the rule (i.e. the function u), amounts to:

$$u = [X \mapsto AX].$$

A.2 A tool we've been missing that is required for many proofs

Proposition A.1: Mean value theorem

Let $f, g : [a, b] \to (E, \|\cdot\|_E)$ be continuous functions on $[a, b]$ and differentiable on (a, b), and suppose that:

$$\forall t \in (a, b), \left\|f'(t)\right\|_E \leqslant g'(t),$$

then:

$$\|f(b) - f(a)\|_E \leqslant g(b) - g(a).$$

Bibliography

[1] Jerrold E. Marsden and Anthony Tromba, *Vector Calculus*, 6th ed.

[2] Walter Rudin, *Principles of Mathematical Analysis*

[3] Donald L. Cohn, *Measure Theory*, 2nd ed.

[4] Walter Rudin, *Real and Complex analysis*