

Honors Mathematics III

Linear Algebra and Functions of Multiple Variables

Horst Hohberger

University of Michigan - Shanghai Jiao Tong University
Joint Institute

Summer Term 2018

Office Hours, Email, TAs

- ▶ Please read the Course profile, which has been uploaded to the Resources section on the Canvas course site.
- ▶ My office is Room 441C in the UM-SJTU Building.
- ▶ My email is horst@sjtu.edu.cn and I'll try to answer email queries within 24 hours.
- ▶ Office hours will be announced on Canvas.
- ▶ Please also make use of the chat room on Canvas for asking questions, making comments or giving feedback on the course.
- ▶ The recitation class schedule and the TA office hours and contact details will be announced on Canvas.
- ▶ Of course, you can also contact the TA with questions outside of office hours, but please observe basic rules of politeness. It is not appropriate to phone your TA at 10 pm in the evening and request help for the exercise set due the next day. Your TA is there to help you, but is not a 24/7 help line.

Course Project

This course requires you to complete a group project.

- ▶ The project accounts for 10% of the course grade.
- ▶ You may work together in groups of 4-5 students.
- ▶ There are three alternative projects; you should **choose one** of these.
- ▶ The mathematical background for the projects involves mainly material from Vv186 and the theory of curves (see Part II of the present course).
- ▶ The course project is a lot of work; **Do not leave it to the last minute but get started as soon as possible!**
- ▶ Further details will be published on Canvas.

Course Grade

The course contains four grade components:

- ▶ Three examinations,
- ▶ A term project.

The course grade will be calculated from these components using the following weighting:

- ▶ First midterm exam: 30%
- ▶ Second midterm exam: 30%
- ▶ Final exam: 30%
- ▶ Course work: 0%
- ▶ Term project: 10%

The course work (weekly assignments) will not contribute to the course grade, but each student needs to obtain at least 60% of the total marks of the assignments in order to receive a passing grade for the course. The course work will be completed by groups of 3 students which will remain unchanged throughout the term.

Honor Code Policy

Students should familiarize themselves with JI's Honor Code, found at
[http://umji.sjtu.edu.cn/academics/academic-integrity/honor-code/.](http://umji.sjtu.edu.cn/academics/academic-integrity/honor-code/)

The standard rules for examinations apply.

Furthermore, in group work (both projects and the course work) Section 5 of the Honor Code is fully enforced: any violation of the Honor Code by a group will cause all group members to be sanctioned equally.

Finally, while communication between members of a group is completely unrestricted, communication between groups (even oral communication) is strictly prohibited.

The Teaching Assistants will be happy to answer any questions regarding the application of the Honor Code.

Class Attendance and Absence for Medical Reasons

I do not formally require that you attend every class. However, if you are unable to attend a significant number of lectures, you should notify me.

The following rules have been laid down by the Academic Office:

- ▶ A student who has been absent from studies for more than one week because of illness or other emergency should consult the program advisor. **[and also talk to me!]**
- ▶ Absence for illness should be supported by a hospital/doctor's certificate. A note that a student visited a medical facility is **not sufficient** excuse for missing an assignment or an exam. The note must specifically indicate that the student was incapable of completing an assignment or taking the exam due to medical problems.

Class Attendance and Absence for Medical Reasons

- ▶ **Late** medical excuses must satisfy the following criteria to be valid:
 - (i) The problem must be confirmed by the doctor to be so severe that the student could not participate in the exam.
 - (ii) The problem must have occurred so suddenly that it was impractical to contact me in advance.
 - (iii) The student must be in contact with me immediately after the exam with the required documentation.

Course Focus

There will be frequent references to results and theorems from the previous term; to allow cross-referencing, a current version of last term's lecture slides has been placed on the Canvas site. All theorems referenced from this course will be prefixed by "186," e.g., 186 Theorem 1.2.1 refers to Theorem 1.2.1 in last term's lecture.

The course is essentially divided into three equal parts:

Part 1: Linear Algebra

- ▶ Systems of Linear Equations
- ▶ Finite-Dimensional Vector Spaces
- ▶ Inner Product Spaces
- ▶ Matrices and Linear Maps
- ▶ Determinants

Course Contents

Part 2: Differential Calculus

- ▶ Open and Closed Sets in Normed Vector Spaces
- ▶ Functions and Derivatives
- ▶ Curves, Potentials and Vector Fields
- ▶ Extrema of Real Functions

Part 3: Integral Calculus in \mathbb{R}^2 and \mathbb{R}^3

- ▶ The Riemann Integral and Measurable Sets
- ▶ Integration in Practice
- ▶ Surfaces and Surface Integrals
- ▶ Divergence and Rotation
- ▶ The Classical Fundamental Theorems of Integration and Differentiation

Mathematica

JI has obtained an unlimited student license for a computer algebra software called **Mathematica**, developed by Wolfram Research. You will be required to make use of the software in your homework assignments, so you should obtain a copy as follows:

(i) Visit

<https://user.wolfram.com/portal/registration.html>

and create a Wolfram ID. You must use an @sjtu.edu.cn email address and give your first and last names in pinyin (example: Xu Baishen enters last name: Xu and first name: Baishen).

(ii) Next, visit

<https://user.wolfram.com/portal/requestAK/c51e79e5334a3600a4f740a2b3720961216dbc17>

and request an Activation Number.

Mathematica

Make a note of the activation number. You will be directed to a page where you can download the installation binaries for the most current version of Mathematica (10.x). (You must select whether you want those for Windows, Linux or OS X.) The software binaries are about 1 GB in size; it may perhaps be possible to share them amongst yourselves to save download time. Try it out and let me know.

- (iii) After downloading, you can install the software. You will be asked to enter the Activation Number you noted above and you will need internet access. Mathematica will then run on a temporary two-week license. Your name will be checked against a list, and if successful, the license will automatically be extended for one year. Therefore, it is very important that you enter your name properly when you request the Wolfram ID.

Part I

Elements of Linear Algebra

Systems of Linear Equations

Finite-Dimensional Vector Spaces

Inner Product Spaces

Linear Maps

Matrices

Theory of Systems of Linear Equations

Determinants

Systems of Linear Equations

Finite-Dimensional Vector Spaces

Inner Product Spaces

Linear Maps

Matrices

Theory of Systems of Linear Equations

Determinants

Linear Systems of Equations

Throughout this course, we use the letter V to denote a (real or complex) vector space. Whenever necessary, we use the letter \mathbb{F} to stand for either \mathbb{R} or \mathbb{C} , depending on the context. ($\mathbb{F} = \mathbb{R}$ in the context of a real vector space, $\mathbb{F} = \mathbb{C}$ in the context of a complex vector space.)

A **linear system** of m (algebraic) equations in n unknowns $x_1, \dots, x_n \in V$ is a set of equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{1.1.1}$$

where $b_1, \dots, b_m \in V$ and $a_{ij} \in \mathbb{F}$, $i = 1, \dots, m$, $j = 1, \dots, n$.

If $b_1 = b_2 = \cdots = b_m = 0$, then (1.1.1) is called a **homogeneous system**. Otherwise, it is called an **inhomogeneous system**.

Linear Systems of Equations

1.1.1. Examples.

1. This is an inhomogeneous system of equations in \mathbb{R} :

$$x_1 + 3x_2 - x_3 = 1$$

$$x_1 - 2x_2 = 2$$

$$10x_2 + x_3 = 1$$

2. This is a homogeneous system of equations in \mathbb{R} :

$$x_1 + 3x_2 - x_3 = 0$$

$$x_1 - 2x_2 = 0$$

$$4x_1 + 7x_2 - 3x_3 = 0$$

3. This is an inhomogeneous system of equations in \mathbb{R}^2 :

$$2x_1 + x_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad x_1 - x_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Linear Systems of Equations

In these examples, the number m of equations is equal to the number n of variables. This is of course not always the case. If $m < n$ we say that the system is **underdetermined**, if $m > n$ it is called **overdetermined**.

A **solution** of a linear system of equations (1.1.1) is a tuple of elements $(y_1, \dots, y_n) \in V^n$ such that the predicate (1.1.1) becomes a true statement.

We will prove later that an inhomogeneous system of equations may have either

- ▶ a unique solution or
- ▶ no solution or
- ▶ an infinite number of solutions.

A homogeneous system evidently always has the **trivial solution**

$$x_1 = x_2 = \cdots = x_n = 0.$$

It further either has

- ▶ no non-trivial solution or
- ▶ an infinite number of non-trivial solutions.

Solving Linear Systems

We will later discuss the theory of existence and uniqueness of solutions for linear systems of equations more extensively. For now, we want to discuss a practical method for actually finding solutions.

In school, you have probably learned that there are some basic strategies for solving systems of equations:

- ▶ solving one of the equations for a variable, and then substituting into the other equations, thereby reducing the number of variables.
- ▶ manipulating two equations until they have identical expressions on one side, then setting them equal.
- ▶ adding and subtracting multiples of one equation to another equation.

Perhaps you have encountered other strategies, but we will look at the last of the three given here. We want to develop a method of systematically solving systems of equations. If we employ a good strategy of adding equations to each other, we will be able to determine the unknowns efficiently and systematically.

Solving Linear Systems

1.1.2. Example. Consider the system,

$$x_1 + 3x_2 - x_3 = 1$$

$$x_1 - 2x_2 = 2$$

$$10x_2 + x_3 = 1$$

Let us subtract the first equation from the second equation:

$$x_1 + 3x_2 - x_3 = 1$$

$$-5x_2 + x_3 = 1$$

$$10x_2 + x_3 = 1$$

Next, we add twice the second equation to the third equation:

$$x_1 + 3x_2 - x_3 = 1$$

$$-5x_2 + x_3 = 1$$

$$3x_3 = 3$$

Solving Linear Systems

We read off from

$$x_1 + 3x_2 - x_3 = 1$$

$$-5x_2 + x_3 = 1$$

$$3x_3 = 3$$

(starting from the last equation and proceeding upwards) that $x_3 = 1$, $x_2 = 0$ and $x_1 = 2$.

Instead of reading off the solution, we could have proceeded more systematically: we divide the last equation by three and the second equation by -5 :

$$x_1 + 3x_2 - x_3 = 1$$

$$x_2 - \frac{1}{5}x_3 = -\frac{1}{5}$$

$$x_3 = 1$$

Solving Linear Systems

We then add $1/5$ times the last equation to the second equation, and the simple last equation to the first equation:

$$\begin{aligned}x_1 + 3x_2 &= 1 \\x_2 &= 0 \\x_3 &= 1\end{aligned}$$

Lastly, we subtract thrice the second equation from the first equation:

$$\begin{aligned}x_1 &= 1 \\x_2 &= 0 \\x_3 &= 1\end{aligned}$$

This gives us the solution directly.

Equivalence of Linear Systems

By adding and subtracting one equation from another, we have been effectively changing the system of equations. Formally, it may be useful to understand the validity of this procedure using the notion of equivalence.

We say that two systems of linear equations are **equivalent** if any solution of the first system is also a solution of the second system and vice-versa.

Thus the systems

$$x_1 + 3x_2 - x_3 = 1 \qquad \qquad \qquad x_1 = 1$$

$$-5x_2 + x_3 = 1 \qquad \qquad \text{and} \qquad \qquad x_2 = 0$$

$$10x_2 + x_3 = 1 \qquad \qquad \qquad x_3 = 1$$

are equivalent.

Simplifying Notation

Listing the variables and the equality sign is essentially a waste of space.
Instead of saying that we transform

$$x_1 + 3x_2 - x_3 = 1$$

$$-5x_2 + x_3 = 1$$

$$10x_2 + x_3 = 1$$

to

$$x_1 + 3x_2 - x_3 = 1$$

$$-5x_2 + x_3 = 1$$

$$3x_3 = 3$$

by adding twice the second equation to the third equation, it would be more efficient to write

$$\left| \begin{array}{ccc|c} 1 & 3 & -1 & 1 \\ 0 & -5 & 1 & 1 \\ 0 & 10 & 1 & 1 \end{array} \right| \xrightarrow{\cdot 2} \sim \left| \begin{array}{ccc|c} 1 & 3 & -1 & 1 \\ 0 & -5 & 1 & 1 \\ 0 & 0 & 3 & 3 \end{array} \right|$$

We will use this notation forthwith.

The Gauß – Jordan Algorithm

The goal of the **Gauß-Jordan algorithm** (also called **Gaußian elimination**) is to transform a system

$$\begin{array}{ccc|c} * & * & * & \diamond \\ * & * & * & \diamond \\ * & * & * & \diamond \end{array}, \quad * \in \mathbb{R} \text{ or } \mathbb{C}, \quad \diamond \in V$$

first into the form

$$\left[\begin{array}{ccc|c} 1 & * & * & \diamond \\ 0 & 1 & * & \diamond \\ 0 & 0 & 1 & \diamond \end{array} \right] \quad (1.1.2)$$

and subsequently into

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & \diamond \\ 0 & 1 & 0 & \diamond \\ 0 & 0 & 1 & \diamond \end{array} \right]. \quad (1.1.3)$$

(Ideally; it may not always be possible to achieve the form (1.1.2))

The Gauß – Jordan Algorithm

We are allowed to achieve this using ***elementary row manipulations***.

These are

1. Swapping (interchanging) two rows,
2. Multiplying each element in a row with a number,
3. Adding a multiple of one row to another row.

Of course, each “row” represents an equation, so we are simply manipulating equations. It is obvious that these manipulations will transform a system into an equivalent system.

A system in the form (1.1.2) is said to be in ***upper triangular form***; a system in the form (1.1.3) is said to be in ***diagonal form***.

The procedure for transforming a system into upper triangular form is called ***forward elimination***; the subsequent procedure for achieving diagonal form is called ***backward substitution***.

The Gauß – Jordan Algorithm

1.1.3. Example. Consider the system

$$2x_1 + x_2 + x_3 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad x_1 - x_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad x_1 + x_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

We rewrite this as

$$\left| \begin{array}{ccc|c} 2 & 1 & 1 & \begin{pmatrix} 2 \\ 1 \end{pmatrix} \\ 1 & -1 & 0 & \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ 1 & 0 & 1 & \begin{pmatrix} 1 \\ 1 \end{pmatrix} \end{array} \right| \quad (1.1.4)$$

We now proceed with forward elimination to achieve upper diagonal form.

The Gauß – Jordan Algorithm (Forward Elimination)

Step 1a: Ensure that the top left hand element is equal to 1:

$$\left| \begin{array}{ccc|c} 2 & 1 & 1 & \left(\begin{smallmatrix} 2 \\ 1 \end{smallmatrix} \right) \\ 1 & -1 & 0 & \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix} \right) \\ 1 & 0 & 1 & \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix} \right) \end{array} \right| \sim \left| \begin{array}{ccc|c} 1 & -1 & 0 & \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix} \right) \\ 2 & 1 & 1 & \left(\begin{smallmatrix} 2 \\ 1 \end{smallmatrix} \right) \\ 1 & 0 & 1 & \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix} \right) \end{array} \right|$$

Step 1b: Eliminate (transform to zero) all lower entries in the first column:

$$\left| \begin{array}{ccc|c} 1 & -1 & 0 & \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix} \right) \\ 2 & 1 & 1 & \left(\begin{smallmatrix} 2 \\ 1 \end{smallmatrix} \right) \\ 1 & 0 & 1 & \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix} \right) \end{array} \right| \xrightarrow{\cdot(-2)} \left| \begin{array}{ccc|c} 1 & -1 & 0 & \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix} \right) \\ 0 & 3 & 1 & \left(\begin{smallmatrix} 2 \\ -1 \end{smallmatrix} \right) \\ 1 & 0 & 1 & \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix} \right) \end{array} \right| \xrightarrow{+} \left| \begin{array}{ccc|c} 1 & -1 & 0 & \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix} \right) \\ 0 & 3 & 1 & \left(\begin{smallmatrix} 2 \\ -1 \end{smallmatrix} \right) \\ 0 & 1 & 1 & \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix} \right) \end{array} \right|$$

The Gauß – Jordan Algorithm (Forward Elimination)

Step 2a: Ensure that the entry in the second row and second column is equal to 1:

$$\left| \begin{array}{ccc|c} 1 & -1 & 0 & \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix} \right) \\ 0 & 3 & 1 & \left(\begin{smallmatrix} 2 \\ -1 \end{smallmatrix} \right) \\ 0 & 1 & 1 & \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix} \right) \end{array} \right| \sim \left| \begin{array}{ccc|c} 1 & -1 & 0 & \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix} \right) \\ 0 & \textcolor{red}{1} & 1 & \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix} \right) \\ 0 & 3 & 1 & \left(\begin{smallmatrix} 2 \\ -1 \end{smallmatrix} \right) \end{array} \right|$$

Step 2b: Eliminate (transform to zero) all entries in the second column below the second row:

$$\left| \begin{array}{ccc|c} 1 & -1 & 0 & \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix} \right) \\ 0 & 1 & 1 & \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix} \right) \\ 0 & 3 & 1 & \left(\begin{smallmatrix} 2 \\ -1 \end{smallmatrix} \right) \end{array} \right| \xrightarrow{\cdot(-3)} \left| \begin{array}{ccc|c} 1 & -1 & 0 & \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix} \right) \\ 0 & 1 & 1 & \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix} \right) \\ 0 & 0 & -2 & \left(\begin{smallmatrix} -1 \\ -1 \end{smallmatrix} \right) \end{array} \right|$$

The Gauß – Jordan Algorithm (Forward Elimination)

Step 3: Ensure that the entry in the third row and third column is equal to 1:

$$\left| \begin{array}{ccc|c} 1 & -1 & 0 & \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ 0 & 1 & 1 & \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ 0 & 0 & -2 & \begin{pmatrix} -1 \\ -1 \end{pmatrix} \end{array} \right| : (-2) \sim \left| \begin{array}{ccc|c} 1 & -1 & 0 & \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ 0 & 1 & 1 & \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ 0 & 0 & 1 & \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} \end{array} \right|$$

The system now has upper triangular form. We next commence the backward substitution.

The Gauß – Jordan Algorithm (Backward Substitution)

Step 1: Eliminate all entries in the third column above the third row:

$$\begin{array}{ccc|c} 1 & -1 & 0 & \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix} \right) \\ 0 & 1 & 1 & \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix} \right) \\ 0 & 0 & 1 & \left(\begin{smallmatrix} 1/2 \\ 1/2 \end{smallmatrix} \right) \end{array} \xrightarrow{\quad \left[\begin{smallmatrix} & & 1 \\ & & 0 \end{smallmatrix} \right] \cdot (-1) \quad} \sim \begin{array}{ccc|c} 1 & -1 & 0 & \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix} \right) \\ 0 & 1 & 0 & \left(\begin{smallmatrix} 1/2 \\ -1/2 \end{smallmatrix} \right) \\ 0 & 0 & 1 & \left(\begin{smallmatrix} 1/2 \\ 1/2 \end{smallmatrix} \right) \end{array}$$

Step 2: Eliminate all entries in the second column above the second row:

$$\begin{array}{ccc|c} 1 & -1 & 0 & \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix} \right) \\ 0 & 1 & 0 & \left(\begin{smallmatrix} 1/2 \\ -1/2 \end{smallmatrix} \right) \\ 0 & 0 & 1 & \left(\begin{smallmatrix} 1/2 \\ 1/2 \end{smallmatrix} \right) \end{array} \xrightarrow{\quad \left[\begin{smallmatrix} & 1 & \\ & 0 & 1 \end{smallmatrix} \right] \quad} \sim \begin{array}{ccc|c} 1 & 0 & 0 & \left(\begin{smallmatrix} 1/2 \\ 1/2 \end{smallmatrix} \right) \\ 0 & 1 & 0 & \left(\begin{smallmatrix} 1/2 \\ -1/2 \end{smallmatrix} \right) \\ 0 & 0 & 1 & \left(\begin{smallmatrix} 1/2 \\ 1/2 \end{smallmatrix} \right) \end{array}$$

Our system now has diagonal form, and we may directly read of the solution.

The Gauß – Jordan Algorithm

We see that the system

$$2x_1 + x_2 + x_3 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad x_1 - x_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad x_1 + x_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

is solved by

$$x_1 = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}, \quad x_2 = \begin{pmatrix} 1/2 \\ -1/2 \end{pmatrix}, \quad x_3 = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}.$$

We notice that instead of solving a single system in \mathbb{R}^2 , we could have solved two systems in \mathbb{R} , determining the components of x_1, x_2, x_3 separately from

$$2x_{11} + x_{21} + x_{31} = 2, \quad x_{11} - x_{21} = 0, \quad x_{11} + x_{31} = 1$$

and

$$2x_{12} + x_{22} + x_{32} = 1, \quad x_{12} - x_{22} = 1, \quad x_{12} + x_{32} = 1.$$

Existence and Uniqueness of Solutions

1.1.4. Remark. A system of m equations with n unknowns will have a unique solution if and only if it is **diagonizable**, i.e., if it can be transformed into diagonal form (1.1.3). Since backward substitution will always work, we see that a unique solution exists if and only if the system can be transformed into an upper triangular form, such as

$$\left| \begin{array}{ccc|c} 1 & * & * & \diamond \\ 0 & 1 & * & \diamond \\ 0 & 0 & 1 & \diamond \end{array} \right| \quad (m = n = 3)$$

or

$$\left| \begin{array}{ccc|c} 1 & * & * & \diamond \\ 0 & 1 & * & \diamond \\ 0 & 0 & 1 & \diamond \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right| \quad (m = 5, n = 3)$$

Existence and Uniqueness of Solutions

Thus $m \geq n$ is a necessary condition for the existence of a unique solution.

A system has no solution if one of the rows has the form

$$\begin{array}{cccc|c} 0 & \dots & 0 & 0 & \diamond \\ \end{array} \quad \diamond \neq 0,$$

which represents the false statement $0 = \diamond$.

If a system has more than one solution, it can be transformed into a so-called **echelon** form, e.g.,

$$\left[\begin{array}{ccccc|c} 1 & * & * & * & * & \diamond \\ 0 & 1 & * & * & * & \diamond \\ 0 & 0 & 0 & 1 & * & \diamond \\ 0 & 0 & 0 & 0 & 1 & \diamond \end{array} \right] \quad (1.1.5)$$

In this case, one of the unknowns acts as a parameter.

The Solution Set

1.1.5. Definition. The **solution set** S of a system of equations (1.1.1) is the set of all n -tuples of numbers x_1, \dots, x_n that satisfy (1.1.1).

- ▶ If a linear system has a unique solution, the set S contains a single point.
- ▶ If there is no solution, $S = \emptyset$.
- ▶ If there is more than one solutions, S is an infinite set.

1.1.6. Example. Consider the real system given by

$$\begin{aligned}x_1 + 2x_2 + 3x_3 &= 0, \\4x_1 + 5x_2 + 6x_3 &= 0, \\7x_1 + 8x_2 + 9x_3 &= 0.\end{aligned}$$

A Homogeneous System

Applying our algorithm,

$$\left| \begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 4 & 5 & 6 & 0 \\ 7 & 8 & 9 & 0 \end{array} \right| \xrightarrow{\begin{array}{l} \cdot(-4) \\ + \\ + \end{array}} \left| \begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 0 & -3 & -6 & 0 \\ 0 & -6 & -12 & 0 \end{array} \right|$$

$$\sim \left| \begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 0 & -3 & -6 & 0 \\ 0 & -6 & -12 & 0 \end{array} \right| \xrightarrow{\begin{array}{l} :(-3) \\ \cdot 6 \\ + \end{array}} \left| \begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right|$$

$$\sim \left| \begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right| \xrightarrow{\begin{array}{l} + \\ \cdot(-2) \end{array}} \left| \begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right|$$

$$\sim \left| \begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right|$$

Writing this system out explicitly,

$$x_1 = x_3,$$

$$x_2 = -2x_3$$

where $x_3 \in \mathbb{R}$ is arbitrary. It is often convenient to introduce a parameter:

$$x_1 = \alpha,$$

$$x_2 = -2\alpha,$$

$$x_3 = \alpha,$$

$$\alpha \in \mathbb{R}.$$

A Homogeneous System

In vector notation, the solution is

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \alpha \\ -2\alpha \\ \alpha \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \quad \alpha \in \mathbb{R}.$$

The solution set is

$$S = \left\{ x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 : x = \alpha \cdot \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \alpha \in \mathbb{R} \right\}.$$

Geometrically, S corresponds to a straight line through the origin. We will return to discuss the geometric properties of solutions to systems of equations (they turn out to be **affine spaces**) later.

 Gauß – Jordan with Mathematica

1.1.7. Example. Consider the system of equations

$$x_1 - 2x_2 + 3x_3 + 4x_4 = 2,$$

$$x_1 - 2x_2 + 5x_3 + 5x_4 = 3,$$

$$-x_1 + 2x_2 - x_3 - 4x_4 = 2.$$

In our array notation, this is

$$\begin{array}{cccc|c} 1 & -2 & 3 & 4 & 2 \\ 1 & -2 & 5 & 5 & 3 \\ -1 & 2 & -1 & -4 & 2 \end{array}$$

Gauß – Jordan with Mathematica

To enter a table/array/matrix in Mathematica (these are all represented in the same way), use the following command structure:

```
A = {{1, -2, 3, 4, 2}, {1, -2, 5, 5, 3}, {-1, 2, -1, -4, 2}}  
{{1, -2, 3, 4, 2}, {1, -2, 5, 5, 3}, {-1, 2, -1, -4, 2}}
```

For convenience, we have here given our array a name, “**A**”. We can retrieve the array by referring to **A**:

A

```
{ {1, -2, 3, 4, 2}, {1, -2, 5, 5, 3}, {-1, 2, -1, -4, 2} }
```



Gauß – Jordan with Mathematica

A more easily readable form is obtained using the **TableForm** command:

```
TableForm[A]
```

| | | | | |
|-----|-----|-----|-----|---|
| 1 | - 2 | 3 | 4 | 2 |
| 1 | - 2 | 5 | 5 | 3 |
| - 1 | 2 | - 1 | - 4 | 2 |

The **RowReduce** command implements the Gauß-Jordan algorithm, returning the echelon form:

```
TableForm[RowReduce[A]]
```

| | | | | |
|---|-----|---|---|-----|
| 1 | - 2 | 0 | 0 | 8 |
| 0 | 0 | 1 | 0 | 2 |
| 0 | 0 | 0 | 1 | - 3 |

Fundamental Lemma for Homogeneous Equations

We will discuss the general theory of uniqueness and existence of solutions to linear equations, after we have studied vector spaces a little more closely.

However, the following **fundamental lemma** requires no additional theory:

1.1.8. Lemma. The homogeneous system

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0$$

$$\vdots$$

$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0$$

of m equations in n real or complex unknowns x_1, \dots, x_n has a non-trivial solution if $n > m$.

Fundamental Lemma for Homogeneous Equations

Proof.

We proceed by induction in m , the number of equations. This means that for any $m \in \mathbb{N} \setminus \{0\}$ we will establish that the system has a non-trivial solution if $n > m$.

We first prove the statement of the lemma for $m = 1$, i.e., we show that

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0, \quad a_{1k} \neq 0, \quad k = 1, \dots, n \quad (1.1.6)$$

has a non-trivial solution whenever $n > 1$.

Proof by induction: For $n = 2$, $a_{11}x_1 + a_{12}x_2 = 0$ has the solution $x_2 = 1$, $x_1 = -a_{12}/a_{11}$. If (1.1.6) has a non-trivial solution (x_1, \dots, x_n) , then

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + a_{1(n+1)}x_{n+1} = 0$$

has the non-trivial solution $(x_1, \dots, x_n, 0)$. □

Fundamental Lemma for Homogeneous Equations

Proof (continued).

We assume that in the system

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0$$

⋮

$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0$$

at least one $a_{ij} \neq 0$. By reordering the equations and renumbering the indices, we can ensure that $a_{11} \neq 0$. We write this system as

$$\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 \\ a_{31} & a_{32} & \dots & a_{3n} & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & 0 \end{array}$$

Fundamental Lemma for Homogeneous Equations

Proof (continued).

Then

$$\begin{array}{cccc|c}
 a_{11} & a_{12} & \dots & a_{1n} & 0 \xrightarrow{\cdot(-\frac{a_{21}}{a_{11}})} \xrightarrow{\cdot(-\frac{a_{31}}{a_{11}})} \xrightarrow{\cdot(-\frac{a_{m1}}{a_{11}})} \\
 a_{21} & a_{22} & \dots & a_{2n} & 0 \leftarrow + \\
 a_{31} & a_{32} & \dots & a_{3n} & 0 \leftarrow + \\
 \vdots & \vdots & \vdots & \vdots & \\
 a_{m1} & a_{m2} & \dots & a_{mn} & 0 \leftarrow +
 \end{array}$$

$$\sim \begin{array}{cccc|c}
 a_{11} & a_{12} & \dots & a_{1n} & 0 \\
 0 & a_{22} - \frac{a_{21}a_{12}}{a_{11}} & \dots & a_{2n} - \frac{a_{21}a_{1n}}{a_{11}} & 0 \\
 0 & a_{32} - \frac{a_{31}a_{12}}{a_{11}} & \dots & a_{3n} - \frac{a_{31}a_{1n}}{a_{11}} & 0 \\
 \vdots & \vdots & & \vdots & \vdots \\
 0 & a_{m2} - \frac{a_{m1}a_{12}}{a_{11}} & \dots & a_{mn} - \frac{a_{m1}a_{1n}}{a_{11}} & 0
 \end{array} \tag{1.1.7}$$

Fundamental Lemma for Homogeneous Equations

Proof (continued).

The boxed area represents a homogeneous system of $m - 1$ equations in $n - 1$ unknowns x_2, \dots, x_n .

We continue with our proof by induction. The case $m = 1$ has been established. Now assume that for $m - 1$ there exists a non-trivial solution whenever the number of unknowns is greater than $m - 1$. A system with m equations and $n > m$ unknowns may be transformed into the form (1.1.7). The subsystem indicated by the boxed area in (1.1.7) by assumption has a non-trivial solution x_2, \dots, x_n . Then the system of m equations in n unknowns has the solution

$$x = \left(-\frac{1}{a_{11}}(a_{12}x_2 + \cdots + a_{1n}x_n), x_2, \dots, x_n \right),$$

which is also non-trivial.



Systems of Linear Equations

Finite-Dimensional Vector Spaces

Inner Product Spaces

Linear Maps

Matrices

Theory of Systems of Linear Equations

Determinants

Linear Independence

We assume throughout that V is a real or complex vector space. As usual, we will use the letter \mathbb{F} to denote either \mathbb{R} (for real vector spaces) or \mathbb{C} (for complex vector spaces).

We want to distinguish elements of vector spaces that are not simply multiples of each other. For example, the vectors $u, v \in \mathbb{R}^2$,

$$u = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad v = \begin{pmatrix} -2 \\ -4 \end{pmatrix}$$

are multiples of each other, because $v = -2u$. In general, we say that $u, v \in V$ are multiples of each other if

$$\exists_{\lambda \in \mathbb{F}} : u = \lambda v \quad \text{or} \quad \begin{array}{c} \exists_{\lambda_1, \lambda_2 \in \mathbb{F}} : \lambda_1 u + \lambda_2 v = 0 \\ |\lambda_1| + |\lambda_2| \neq 0 \end{array}$$

Linear Independence

If u and v are not multiples of each other, we say that they are (**linearly independent**). This means that

$$\neg \left(\exists_{\substack{\lambda_1, \lambda_2 \in \mathbb{F} \\ |\lambda_1| + |\lambda_2| \neq 0}} : \lambda_1 u + \lambda_2 v = 0 \right)$$

or

$$\forall_{\lambda_1, \lambda_2 \in \mathbb{F}} \quad \lambda_1 u + \lambda_2 v = 0 \quad \Rightarrow \quad \lambda_1 = \lambda_2 = 0.$$

1.2.1. Definition. Let V be a real or complex vector space and $v_1, \dots, v_n \in V$. Then the vectors v_1, \dots, v_n are said to be **independent** if for all $\lambda_1, \dots, \lambda_n \in \mathbb{F}$

$$\sum_{k=1}^n \lambda_k v_k = 0 \quad \Rightarrow \quad \lambda_1 = \lambda_2 = \dots = \lambda_n = 0.$$

A finite set $M \subset V$ is called an **independent set** if its elements are independent.

Linear Independence

1.2.2. Example. The vectors

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

are independent (and $M = \{v_1, v_2\}$ is an independent set), because

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0 = \lambda_1 v_1 + \lambda_2 v_2 = \begin{pmatrix} \lambda_1 \\ 2\lambda_2 \end{pmatrix}$$

is equivalent to the system of equations

$$0 = \lambda_1,$$

$$0 = 2\lambda_2,$$

which has the unique solution $\lambda_1 = 0$ and $\lambda_2 = 0$.

Linear Independence

1.2.3. Example. The vectors

$$v_1 = \begin{pmatrix} 1 \\ 4 \\ 7 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 2 \\ 5 \\ 8 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 3 \\ 6 \\ 9 \end{pmatrix}$$

are not independent, because

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = 0 = \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 = \begin{pmatrix} \lambda_1 + 2\lambda_2 + 3\lambda_3 \\ 4\lambda_1 + 5\lambda_2 + 6\lambda_3 \\ 7\lambda_1 + 8\lambda_2 + 9\lambda_3 \end{pmatrix}$$

has a non-trivial solution, as we have seen in Example 1.1.6. For example, we can take $\lambda_1 = 1$, $\lambda_2 = -2$, $\lambda_3 = 1$. Hence,

$$\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 = 0 \quad \not\Rightarrow \quad \lambda_1 = \lambda_2 = \lambda_3 = 0,$$

and the vectors are not independent.

Linear Combinations and Span

1.2.4. Definition. Let $v_1, \dots, v_n \in V$ and $\lambda_1, \dots, \lambda_n \in \mathbb{F}$. Then the expression

$$\sum_{k=1}^n \lambda_k v_k = \lambda_1 v_1 + \cdots + \lambda_n v_n$$

is called a **linear combination** of the vectors v_1, \dots, v_n .

The set

$$\text{span}\{v_1, \dots, v_n\} = \left\{ y \in V : y = \sum_{k=1}^n \lambda_k v_k, \lambda_1, \dots, \lambda_n \in \mathbb{F} \right\}$$

is called the (**linear**) **span** or the **linear hull** of the vectors v_1, \dots, v_n .

Linear Combinations and Span

1.2.5. Example. $\text{span} \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\} = \mathbb{R}^2$.

We need to show that every $x \in \mathbb{R}^2$ can be written as

$$x = \lambda_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

for some $\lambda_1, \lambda_2 \in \mathbb{R}$. This means we need to solve

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 + 2\lambda_2 \\ \lambda_2 \end{pmatrix}$$

This is easily done, and we obtain $\lambda_2 = x_2$ and $\lambda_1 = x_1 - 2x_2$. Thus for any $x \in \mathbb{R}^2$ we have $x \in \text{span} \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\}$. Since $\text{span} \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\} \subset \mathbb{R}^2$ by definition, we are finished.

Linear Combinations and Span

1.2.6. Lemma. The vectors $v_1, \dots, v_n \in V$ are independent if and only if none of them is contained in the span of all the others.

Proof.

We prove the contraposition of the statement:

$$\exists_{k \in \{1, \dots, n\}} v_k \in \text{span}\{v_1, \dots, v_{k-1}, v_{k+1}, v_n\}$$

$$\Leftrightarrow \exists_{k \in \{1, \dots, n\}} \exists_{\substack{\lambda_i \in \mathbb{F} \\ i \in \{1, \dots, n\} \setminus \{k\} \\ \sum |\lambda_i| \neq 0}} v_k = \sum_i \lambda_i v_i$$

$$\Leftrightarrow \exists_{\substack{\lambda_i \in \mathbb{F} \\ i \in \{1, \dots, n\} \\ \sum |\lambda_i| \neq 0}} \sum_i \lambda_i v_i = 0$$

□

Span of Subsets

More generally, if V is a vector space and M is some subset of V , then we can define the **span of M** as the set containing all (finite) linear combinations of elements of M , i.e.,

$$\text{span } M := \left\{ v \in V : \exists_{n \in \mathbb{N}} \exists_{\lambda_1, \dots, \lambda_n \in \mathbb{F}} \exists_{m_1, \dots, m_n \in M} : v = \sum_{i=1}^n \lambda_i m_i \right\}.$$

Note that this definition does not presume that M is a subspace, just an arbitrary subset of V . Furthermore, although only finite linear combinations are considered, the set M may well be infinite in size.

Moreover, even though M is just any set, $\text{span } M$ will be a subspace of V .

1.2.7. Example. Let $M = \{f \in C(\mathbb{R}) : f(x) = x^n, n \in \mathbb{N}\}$ denote the set of all monomials in the space of continuous functions on \mathbb{R} . Then $\mathcal{P}(\mathbb{R}) := \text{span } M$ is the space of all polynomials (of any degree) in $C(\mathbb{R})$.

Basis

1.2.8. Definition. Let V be a real or complex vector space. An n -tuple $\mathcal{B} = (b_1, \dots, b_n) \in V^n$ is called an **(ordered and finite) basis** of V if every vector v has a unique representation

$$v = \sum_{i=1}^n \lambda_i b_i, \quad \lambda_i \in \mathbb{F}.$$

The numbers λ_i are called the **coordinates** of v with respect to \mathcal{B} .

1.2.9. Example. The tuple of vectors (e_1, \dots, e_n) , $e_i \in \mathbb{R}^n$,

$$e_i = (0, \dots, 0, \underset{\substack{\uparrow \\ i\text{th}}}{1}, 0, \dots, 0), \quad i = 1, \dots, n,$$

is called the **standard basis** or **canonical basis** of \mathbb{R}^n .

Characterization of Bases

Sometimes we are not interested in the order of the elements of a basis, and write $\mathcal{B} = \{b_1, \dots, b_n\}$, replacing the tuple by a set. This is known as an **unordered basis**.

1.2.10. Theorem. Let V be a real or complex vector space. An n -tuple $\mathcal{B} = (b_1, \dots, b_n) \in V^n$ is a basis of V if and only if

- (i) the vectors b_1, \dots, b_n are linearly independent, i.e., \mathcal{B} is an independent set, and
- (ii) $V = \text{span } \mathcal{B}$.

Characterization of Bases

Proof.

(\Rightarrow) Suppose that \mathcal{B} is a basis of V . Then every $v \in V$ can be expressed as

$$v = \sum_{i=1}^n \lambda_i b_i$$

for some coefficients $\lambda_i \in \mathbb{F}$. Hence, $V \subset \text{span } \mathcal{B}$. From $\mathcal{B} \subset V$ it is clear that $\text{span } \mathcal{B} \subset V$, so we deduce $V = \text{span } \mathcal{B}$. The zero vector $0 \in V$ has the representation

$$0 = 0 \cdot b_1 + \cdots + 0 \cdot b_n.$$

Since \mathcal{B} is a basis, this representation is unique, i.e.,

$$\sum_{i=1}^n \lambda_i b_i = 0 \quad \Rightarrow \quad \lambda_1 = \cdots = \lambda_n = 0.$$

Characterization of Bases

Proof (continued).

It follows that \mathcal{B} is an independent set.

(\Leftarrow) Suppose that $\mathcal{B} \subset V$ satisfies $\text{span } \mathcal{B} = V$. Then every $v \in V$ is an element of the span of \mathcal{B} , so

$$v = \sum_{i=1}^n \lambda_i b_i$$

for some coefficients $\lambda_i \in \mathbb{F}$. It remains to show that this representation is unique. Suppose that

$$v = \sum_{i=1}^n \lambda_i b_i = \sum_{i=1}^n \mu_i b_i, \quad \lambda_i, \mu_i \in \mathbb{F}.$$

Characterization of Bases

Proof (continued).

Then

$$0 = \sum_{i=1}^n (\lambda_i - \mu_i) b_i.$$

Since the b_i are all independent, this implies

$$\lambda_i - \mu_i = 0, \quad i = 1, \dots, n,$$

so the representation is unique. □

Finite- and Infinite-Dimensional Spaces

1.2.11. Definition. Let V be a real or complex vector space. Then V is called ***finite-dimensional*** if either

- ▶ $V = \{0\}$ or
- ▶ V possesses a finite basis.

If V is not finite-dimensional, we say that it is ***infinite-dimensional***.

1.2.12. Example.

1. The space of polynomials of degree at most n ,

$$\mathcal{P}_n = \{f \in C(\mathbb{R}) : f(x) = \sum_{k=0}^n a_k x^k, a_0, a_1, \dots, a_n \in \mathbb{R}\}$$

is finite-dimensional, because it has the basis $\mathcal{B} = (1, x, x^2, \dots, x^n)$.

2. The space of real polynomials of any degree, $\mathcal{P}(\mathbb{R})$, is infinite-dimensional. (See Example 1.2.7.)

Length of Bases

1.2.13. Theorem. Let V be a real or complex finite-dimensional vector space, $V \neq \{0\}$. Then any basis of V has the same length (number of elements) n .

Proof.

Let $\mathcal{A} = (a_1, \dots, a_n)$ be a basis of V . We will show that no tuple $\mathcal{B} = (b_1, \dots, b_m)$ with $m > n$ can be a basis of V . (We do not need to consider the case $m < n$, because we could just switch the role of \mathcal{A} and \mathcal{B} .) Thus, suppose that \mathcal{A} and \mathcal{B} given as above are both bases. Then for every $j = 1, \dots, m$ there exist uniquely determined numbers $c_{ij} \in \mathbb{F}$, $i = 1, \dots, n$, such that

$$b_j = \sum_{i=1}^n c_{ij} a_i.$$

Length of Bases

Proof (continued).

Now let $\lambda_1, \dots, \lambda_m \in \mathbb{F}$ and consider the linear combination

$$\sum_{j=1}^m \lambda_j b_j = \sum_{j=1}^m \sum_{i=1}^n c_{ij} \lambda_j a_i = \sum_{i=1}^n \underbrace{\left(\sum_{j=1}^m c_{ij} \lambda_j \right) a_i}_{=: \mu_i}$$

Now, since \mathcal{A} and \mathcal{B} are bases and as we know that $0 \in V$ has a unique representation in terms of basis vectors, we have

$$\begin{aligned} \lambda_1 = \lambda_2 = \cdots = \lambda_m = 0 &\Leftrightarrow \sum_{j=1}^m \lambda_j b_j = 0 \Leftrightarrow \sum_{i=1}^n \mu_i a_i = 0 \\ &\Leftrightarrow \mu_1 = \mu_2 = \cdots = \mu_n = 0 \\ &\Leftrightarrow \forall i=1, \dots, n \quad \sum_{j=1}^m c_{ij} \lambda_j = 0 \end{aligned}$$

Length of Bases

Proof (continued).

This means that the homogeneous system of equations

$$\begin{aligned} c_{11}\lambda_1 + c_{12}\lambda_2 + \cdots + c_{1m}\lambda_m &= 0 \\ &\vdots \\ c_{n1}\lambda_1 + c_{n2}\lambda_2 + \cdots + c_{nm}\lambda_m &= 0 \end{aligned} \tag{1.2.1}$$

has only the trivial solution

$$\lambda_1 = \lambda_2 = \cdots = \lambda_m = 0.$$

However, we have assumed that $m > n$, i.e., there are more unknowns than equations. By the Fundamental Lemma 1.1.8, there must exist a non-trivial solution. Thus we have a contradiction, so \mathcal{B} can not be a basis if \mathcal{A} is. □

Dimension

1.2.14. Definition. Let V be a finite-dimensional real or complex vector space. We define the **dimension** of V , denoted $\dim V$, as follows:

- (i) If $V = \{0\}$, $\dim V = 0$.
- (ii) If $V \neq \{0\}$, $\dim V = n$, where n is the length of any basis of V .

If V is an infinite-dimensional vector space we write $\dim V = \infty$.

1.2.15. Examples.

- (i) $\dim \mathbb{R}^n = n$
- (ii) $\dim \mathcal{P}_n = n + 1$
- (iii) $\dim C(\mathbb{R}) = \infty$
- (iv) $\dim \{(x_1, x_2) \in \mathbb{R}^2 : x_2 = 3x_1\} = 1$

Characterization of Bases

1.2.16. Remark. In an n -dimensional vector space V a basis is an independent set with n elements that spans V . A few questions arise naturally:

1. Is any independent set with n elements in an n -dimensional space a basis?
2. If not, is it possible to find independent sets with more than n elements in an n -dimensional space?

Maximal Subsets

In order to answer these and similar questions, we will work towards a fundamentally important result called the **basis extension theorem**. First, we need a lemma:

1.2.17. Lemma. Let $a_1, \dots, a_{n+1} \in V$ and assume that a_1, \dots, a_n are independent and that a_1, \dots, a_{n+1} are dependent. Then a_{n+1} is a linear combination of (some of) a_1, \dots, a_n .

The proof is quite easy and left to you as an exercise.

1.2.18. Definition. Let V be a real or complex vector space and $A \subset V$ a **finite** set. An independent subset $F \subset A$ is called **maximal** if every $x \in A$ is a linear combination of elements of F .

If A is finite and $F \subset A$ is maximal, then $\text{span } F = \text{span } A$. A maximal subset is of course not defined uniquely.

Maximal Subsets

1.2.19. Example. Let $V = \mathbb{R}^3$ and

$$A = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right\}.$$

Then

$$F_1 = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}, \quad F_2 = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right\}, \quad F_3 = \left\{ \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right\}$$

are all maximal independent subsets of A . Furthermore,

$$\text{span } A = \text{span } F_1 = \text{span } F_2 = \text{span } F_3.$$

Maximal Subsets

1.2.20. Theorem. Let V be a vector space and $A \subset V$ a finite set. Then every independent subset $A' \subset A$ lies in some maximal subset $F \subset A$.

Proof.

We proceed algorithmically. We ask: Does there exist a vector $x \in A \setminus A'$ such that $x \notin \text{span } A'$?

- ▶ If **no**, we are finished, because A' is maximal.
- ▶ If **yes**, we take this x and define $A'' = A' \cup \{x\}$.

By Lemma 1.2.17, A'' is independent (otherwise $x \in \text{span } A'$ and we have a contradiction) and we can repeat the procedure, substituting A'' for A' .

Since A is finite, the loop will terminate at some point and we obtain a maximal independent subset of A . □

Basis Extension Theorem

1.2.21. Basis Extension Theorem. Let V be a finite-dimensional vector space and $A' \subset V$ an independent set. Then there exists a basis of V containing A' .

Proof.

Write $A' = \{a_1, \dots, a_m\}$ and choose a basis $\mathcal{A} = \{a_{m+1}, \dots, a_{m+n}\}$ of V , $\dim V = n$. We now define

$$A = \{a_1, \dots, a_{m+n}\} \supset A'.$$

By Theorem 1.2.20 there exists a maximal independent subset F of A containing A' .

Since \mathcal{A} is a basis, $V = \text{span } \mathcal{A} = \text{span } A$. Furthermore, $\text{span } F = \text{span } A$, so $\text{span } F = V$. Thus F is a basis. □

Basis Extension Theorem

1.2.22. Corollary. Let V be an n -dimensional vector space, $n \in \mathbb{N}$. Then any independent set A with n elements is a basis of V .

Proof.

By the basis extension theorem there is a basis containing A . Since this basis will have n elements, A itself is this basis. \square

1.2.23. Corollary. Let V be an n -dimensional vector space, $n \in \mathbb{N}$. Then an independent set A may have at most n elements.

Proof.

By the basis extension theorem there is a basis containing A . Since this basis will have n elements, A may not have more elements than this. \square

Sums of Vector Spaces

1.2.24. Definition. Let V be a real or complex vector space and U, W be sets in V .

- (i) We define the **sum of U and W** by

$$U + W := \left\{ v \in V : \exists_{u \in U} \exists_{w \in W} : v = u + w \right\}.$$

- (ii) If U and W are subspaces of V with $U \cap W = \{0\}$, the sum $U + W$ is called **direct**, and we denote it by $U \oplus W$.

1.2.25. Remark. It is easy to see that if U, W are subspaces of V , then $U + W$ (or $U \oplus W$) is a subspace of V ; check this for yourself!

Sums of Vector Spaces

1.2.26. Examples.

- (i) Let $U, W \subset \mathbb{R}^2$ be given by

$$U = \text{span} \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}, \quad W = \text{span} \left\{ \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\}.$$

Then every $x \in \mathbb{R}^2$ has a representation in the form $x = u + w$, where $u \in U$ and $w \in W$. (Why? See also Example 1.2.5) Therefore, $\mathbb{R}^2 = U + W$.

Furthermore, $U \cap W = \{0\}$ since $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ are independent (See Lemma 1.2.6.) Hence, we can write

$$\mathbb{R}^2 = U \oplus W.$$

Sums of Vector Spaces

(ii) Let $U, W \subset \mathbb{R}^3$ be given by

$$U = \text{span} \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\}, \quad W = \text{span} \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

Then $U + W = \mathbb{R}^3$, but the sum is not direct, because
 $(0, 1, 0) \in U \cap W$

(iii) We could write $V = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_3 = 0\}$ as

$$V = \text{span} \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\} \oplus \text{span} \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

Sums of Vector Spaces

1.2.27. Lemma. The sum $U + W$ of vector spaces U, W is direct if and only if all $x \in U + W$, $x \neq 0$, have a **unique** representation $x = u + w$, $u \in U$, $w \in W$.

Proof.

(\Rightarrow) We show the contraposition: if the representation is not unique for all $x \in U + W$, then the sum is not direct. Let $x = u + w = u' + w'$ with $u, u' \in U$, $w, w' \in W$. Then $u - u' = w' - w$, so $u - u' \in U$ and $u - u' \in W$. Thus $U \cap W \neq \{0\}$.

Sums of Vector Spaces

Proof (continued).

(\Leftarrow) We again show the contraposition: if the sum is not direct, then there exists some $x \in U + W$ with a non-unique representation. This is obvious, because if $0 \neq x \in U \cap W$, then we may write

$$x = \underbrace{x}_{\in U} + \underbrace{0}_{\in W} = \underbrace{\frac{1}{2}x}_{\in U} + \underbrace{\frac{1}{2}x}_{\in W},$$

so this x has more than one representation. □

Sums of Vector Spaces

1.2.28. Theorem. Let V be a vector space and $U, W \subset V$ be finite-dimensional subspaces of V . Then

$$\dim(U + W) + \dim(U \cap W) = \dim U + \dim W.$$

The proof will be discussed in recitation class.

Systems of Linear Equations

Finite-Dimensional Vector Spaces

Inner Product Spaces

Linear Maps

Matrices

Theory of Systems of Linear Equations

Determinants

Inner Product Spaces

1.3.1. Definition. Let V be a real or complex vector space. Then a map $\langle \cdot, \cdot \rangle: V \times V \rightarrow \mathbb{F}$ is called a **scalar product** or **inner product** if for all $u, v, w \in V$ and all $\lambda \in \mathbb{F}$

- (i) $\langle v, v \rangle \geq 0$ and $\langle v, v \rangle = 0$ if and only if $v = 0$,
- (ii) $\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle$,
- (iii) $\langle u, \lambda v \rangle = \lambda \langle u, v \rangle$,
- (iv) $\langle u, v \rangle = \overline{\langle v, u \rangle}$.

The pair $(V, \langle \cdot, \cdot \rangle)$ is called an **inner product space**.

1.3.2. Remark. Properties (iii) and (iv) imply that

$$\langle \lambda u, v \rangle = \overline{\langle v, \lambda u \rangle} = \overline{\lambda} \overline{\langle v, u \rangle} = \overline{\lambda} \langle u, v \rangle.$$

We say that the inner product is linear in the second component and anti-linear in the first component.

The Induced Norm

1.3.3. Examples.

- In \mathbb{R}^n we define the **canonical** or **standard scalar product**

$$\langle x, y \rangle := \sum_{i=1}^n x_i y_i, \quad x, y \in \mathbb{R}^n. \quad (1.3.1)$$

- In \mathbb{C}^n we can define the inner product

$$\langle x, y \rangle := \sum_{i=1}^n \overline{x_i} y_i, \quad x, y \in \mathbb{C}^n.$$

- In $C([a, b])$, the space of complex-valued, continuous functions on the interval $[a, b]$, we can define an inner product by

$$\langle f, g \rangle := \int_a^b \overline{f(x)} g(x) dx, \quad f, g \in C([a, b]).$$

The Induced Norm

1.3.4. Definition. Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space. The map

$$\|\cdot\|: V \rightarrow \mathbb{R}, \quad \|v\| = \sqrt{\langle v, v \rangle}$$

is called the **induced norm** on V .

1.3.5. Examples.

- The induced norm in \mathbb{R}^n and \mathbb{C}^n is given by

$$\|x\| = \sqrt{\langle x, x \rangle} = \sqrt{\sum_{i=1}^n |x_i|^2} = \|x\|_2, \quad (1.3.2)$$

which is the usual euclidean norm.

- The induced norm on $C([a, b])$ is

$$\|f\| = \sqrt{\langle f, f \rangle} = \sqrt{\int_a^b |f(x)|^2 dx} = \|f\|_2$$

which is just the 2-norm.

The Induced Norm

1.3.6. Cauchy-Schwarz Inequality. Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product vector space. Then

$$|\langle u, v \rangle| \leq \|u\| \cdot \|v\| \quad \text{for all } u, v \in V$$

where $\|\cdot\|$ is the induced norm.

Proof.

Let $e := v/\|v\|$. Then $\langle e, e \rangle = \langle v, v \rangle / \|v\|^2 = 1$ and

$$\begin{aligned} 0 &\leq \|u - \langle e, u \rangle e\|^2 = \langle u - \langle e, u \rangle e, u - \langle e, u \rangle e \rangle \\ &= \|u\|^2 - |\langle e, u \rangle|^2 \end{aligned}$$

It follows that

$$|\langle u, v \rangle|^2 = \|v\|^2 \cdot |\langle u, e \rangle|^2 \leq \|u\|^2 \cdot \|v\|^2.$$

□

The Induced Norm

1.3.7. Corollary. The induced norm is actually a norm, i.e., it satisfies

- (i) $\|v\| \geq 0$, $\|v\| = 0 \Leftrightarrow v = 0$,
- (ii) $\|\lambda v\| = |\lambda| \cdot \|v\|$,
- (iii) $\|u + v\| \leq \|u\| + \|v\|$

for all $u, v \in V$ and $\lambda \in \mathbb{F}$.

Proof.

All properties except for the triangle inequality are easily checked. By the Cauchy-Schwarz inequality, we have

$$\begin{aligned}\|u + v\|^2 &= \|u\|^2 + \|v\|^2 + 2 \operatorname{Re} \langle u, v \rangle \\ &\leq \|u\|^2 + \|v\|^2 + 2|\langle u, v \rangle| \\ &\leq \|u\|^2 + \|v\|^2 + 2\|u\|\|v\| \\ &= (\|u\| + \|v\|)^2.\end{aligned}$$



Angle Between Vectors

1.3.8. Remark. Every inner product space is also a normed vector space and by extension a metric space.

1.3.9. Definition. Let V be a real inner product space and $u, v \in V$. We define the **angle** $\alpha(u, v) \in [0, \pi]$ **between** u **and** v by

$$\cos \alpha(u, v) = \frac{\langle u, v \rangle}{\|u\| \|v\|}. \quad (1.3.3)$$

This definition makes sense, since by the Cauchy-Schwarz inequality

$$\left| \frac{\langle u, v \rangle}{\|u\| \|v\|} \right| = \frac{|\langle u, v \rangle|}{\|u\| \|v\|} \leq 1.$$

In \mathbb{R}^2 and \mathbb{R}^3 the expression (1.3.3) of course corresponds to our geometric notion of the (cosine of the) angle between two vectors.

Angle Between Vectors

1.3.10. Example. For $x, y \in \mathbb{R}^2$ we have $\sphericalangle(x, y) = \alpha(x, y)$.

We may assume that $\|x\| = \|y\| = 1$ and we consider the case

$$x = \begin{pmatrix} \cos \varphi_1 \\ \sin \varphi_1 \end{pmatrix}, \quad y = \begin{pmatrix} \cos \varphi_2 \\ \sin \varphi_2 \end{pmatrix}, \quad 0 < \varphi_1 < \varphi_2 < \pi.$$

(Cf. the section on polar coordinates in last term's lecture.)

Then $\sphericalangle(x, y) = \varphi_2 - \varphi_1$ and

$$\begin{aligned} \cos \sphericalangle(x, y) &= \cos(\varphi_2 - \varphi_1) = \cos \varphi_2 \cos \varphi_1 + \sin \varphi_2 \sin \varphi_1 \\ &= \langle x, y \rangle = \cos \alpha(x, y) \end{aligned}$$

In a similar manner, one can prove that $\sphericalangle(x, y) = \alpha(x, y)$ for $x, y \in \mathbb{R}^3$.

Vectors, Norms and Inner Products

We can use a **Table** command to create a general vector:

```
X = Table[xi, {i, 3}]
```

```
{x1, x2, x3}
```

The standard inner product (1.3.1) is implemented by a simple dot:

```
Y = Table[yi, {i, 3}];
```

```
X.Y
```

$$x_1 Y_1 + x_2 Y_2 + x_3 Y_3$$

The induced norm (1.3.2) is given by the **Norm** command:

```
Norm[X]
```

$$\sqrt{\text{Abs}[x_1]^2 + \text{Abs}[x_2]^2 + \text{Abs}[x_3]^2}$$

Orthogonality

1.3.11. Definition. Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product vector space.

- (i) Two vectors $u, v \in V$ are called **orthogonal** or **perpendicular** if $\langle u, v \rangle = 0$. We then write $u \perp v$.
- (ii) We call

$$M^\perp := \left\{ v \in V : \forall_{m \in M} \langle m, v \rangle = 0 \right\}$$

the **orthogonal complement** of a set $M \subset V$.

For short, we sometimes write $v \perp M$ instead of $v \in M^\perp$ or $v \perp m$ for all $m \in M$.

1.3.12. Lemma. The orthogonal complement M^\perp is a subspace of V .

Proof.

If $v_1, v_2 \in M^\perp$, then $\langle v_1 + v_2, m \rangle = \langle v_1, m \rangle + \langle v_2, m \rangle = 0 + 0 = 0$ for all $m \in M$, so $v_1 + v_2 \in M^\perp$. Similarly, if $v \in M^\perp$ and $\lambda \in \mathbb{F}$, then $\langle \lambda v, m \rangle = \bar{\lambda} \langle v, m \rangle = 0$, so $\lambda v \in M^\perp$. Thus M^\perp is a subspace of V . □

Orthogonality

1.3.13. Pythagoras's Theorem. Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space and M some subset of V . Let $z = x + y$, where $x \in M$ and $y \in M^\perp$. Then

$$\|z\|^2 = \|x\|^2 + \|y\|^2.$$

Proof.

We see directly that

$$\begin{aligned}\|z\|^2 &= \langle z, z \rangle = \langle x + y, x + y \rangle \\ &= \langle x, x \rangle + \underbrace{\langle x, y \rangle}_{=0} + \underbrace{\langle y, x \rangle}_{=0} + \langle y, y \rangle \\ &= \|x\|^2 + \|y\|^2.\end{aligned}$$

□

Orthonormal Systems

1.3.14. Definition. Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product vector space. A tuple of vectors $(v_1, \dots, v_r) \subset V$ is called a **(finite) orthonormal system** if

$$\langle v_j, v_k \rangle = \delta_{jk} := \begin{cases} 1 & \text{for } j = k, \\ 0 & \text{for } j \neq k, \end{cases} \quad j, k = 1, \dots, r,$$

i.e., if $\|v_k\| = 1$ and $v_j \perp v_k$ for $j \neq k$.

1.3.15. Example. The standard basis vectors in \mathbb{R}^3 ,

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

form an orthonormal system (e_1, e_2, e_3) with respect to the standard scalar product.

Orthonormal Systems

1.3.16. Lemma. Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product vector space and $\mathcal{F} = (v_1, \dots, v_r) \subset V$ an orthonormal system. Then the elements of \mathcal{F} are linearly independent.

Proof.

We want to prove that for any $\lambda_1, \dots, \lambda_r \in \mathbb{F}$

$$\sum_{i=1}^r \lambda_i v_i = 0 \quad (1.3.4)$$

implies $\lambda_1 = \dots = \lambda_r = 0$. We take the scalar product of (1.3.4) with v_1 :

$$\begin{aligned} 0 &= \langle v_1, 0 \rangle = \langle v_1, \lambda_1 v_1 + \dots + \lambda_r v_r \rangle \\ &= \lambda_1 \underbrace{\langle v_1, v_1 \rangle}_{=1} + \lambda_2 \underbrace{\langle v_1, v_2 \rangle}_{=0} + \dots + \lambda_r \underbrace{\langle v_1, v_r \rangle}_{=0} = \lambda_1, \end{aligned}$$

so (1.3.4) implies $\lambda_1 = 0$. Similarly, we obtain $\lambda_2 = \dots = \lambda_r = 0$. □

Orthonormal Bases

1.3.17. Definition. Let $(V, \langle \cdot, \cdot \rangle)$ be a finite-dimensional inner product vector space and $\mathcal{B} = (e_1, \dots, e_n)$ a basis of V . If \mathcal{B} is also an orthonormal system, we say that \mathcal{B} is an **orthonormal basis** (ONB).

1.3.18. Theorem. Let $(V, \langle \cdot, \cdot \rangle)$ be a finite-dimensional inner product vector space and $\mathcal{B} = (e_1, \dots, e_n)$ an orthonormal basis of V . Then every $v \in V$ has the basis representation

$$v = \sum_{j=1}^n \langle e_j, v \rangle e_j.$$

1.3.19. Definition. The numbers $\langle e_j, v \rangle$ are called **Fourier coefficients** of v with respect to the basis \mathcal{B} . The vector

$$\pi_{e_i} v := \langle e_i, v \rangle e_i$$

is called the **projection of v onto e_i**

Orthonormal Bases

Proof of Theorem 1.3.18.

Since \mathcal{B} is a basis, for every $v \in V$ there exist coefficients $\lambda_1, \dots, \lambda_n \in \mathbb{F}$ such that

$$v = \sum_{j=1}^n \lambda_j e_j.$$

Now for any $k = 1, \dots, n$, we have

$$\langle e_k, v \rangle = \sum_{j=1}^n \lambda_j \langle e_k, e_j \rangle = \sum_{j=1}^n \lambda_j \delta_{kj} = \lambda_k,$$

so it follows that

$$v = \sum_{j=1}^n \lambda_j e_j = \sum_{j=1}^n \langle e_j, v \rangle e_j.$$



Orthonormal Bases

The following result, which follows directly from Theorem 1.3.18, generalizes Pythagoras's Theorem 1.3.13:

1.3.20. Parseval's Theorem. Let $(V, \langle \cdot, \cdot \rangle)$ be a finite-dimensional inner product vector space and $\mathcal{B} = \{e_1, \dots, e_n\}$ an orthonormal basis of V . Then

$$\|v\|^2 = \sum_{i=1}^n |\langle v, e_i \rangle|^2$$

for any $v \in V$.

We have generalized the concepts of angle and orthogonality to vector spaces and thereby obtained Pythagoras's Theorem and now Parseval's inequality. For understanding the geometry of vector spaces (and thereby extending the "elementary" geometry of \mathbb{R}^3 , the projection of a vector onto subspaces is of fundamental importance. The following theorem develops this concept a little further.

The Projection Theorem

1.3.21. **Projection Theorem.** Let $(V, \langle \cdot, \cdot \rangle)$ be a (possibly infinite-dimensional) inner product vector space and (e_1, \dots, e_r) , $r \in \mathbb{N}$, be an orthonormal system in V . Denote $U := \text{span}\{e_1, \dots, e_r\}$.

Then for every $v \in V$ there exists a unique representation

$$v = u + w \quad \text{where } u \in U \text{ and } w \in U^\perp$$

and $u = \sum_{i=1}^r \langle e_i, v \rangle e_i$, $w := v - u$.

1.3.22. **Definition.** The vector

$$\pi_U v := \sum_{i=1}^r \langle e_i, v \rangle e_i$$

is called the **orthogonal projection of v onto U** . The projection theorem essentially states that $\pi_U v$ always exists and is independent of the choice of the orthonormal system (it depends only on the span U of the system).

The Projection Theorem

1.3.23. Example. Consider the subspace $U = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_3 = 0\}$ of \mathbb{R}^3 . An orthonormal basis of U is given by $\mathcal{B} = \{e_1, e_2\}$, where

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Then the projection of a vector $y = (y_1, y_2, y_3)$ onto U is given by

$$\begin{aligned}\pi_U y &= \langle e_1, y \rangle e_1 + \langle e_2, y \rangle e_2 \\ &= \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \right\rangle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \left\langle \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \right\rangle \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \\ &= y_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + y_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ 0 \end{pmatrix}\end{aligned}$$

The Projection Theorem

Proof of the Projection Theorem.

We first show the uniqueness of the decomposition: Assume $v = u + w = u' + w'$. Then by Pythagoras's theorem,

$$0 = \|u - u' + (w - w')\|^2 = \|u - u'\|^2 + \|w - w'\|^2,$$

so $\|u - u'\| = \|w - w'\| = 0$. Thus $u = u'$ and $w = w'$.

Regarding the existence of such a decomposition, it is clear that

$$u = \sum_{i=1}^r \langle e_i, v \rangle e_i$$

lies in U . We need to show that $w \in U^\perp$, i.e., $u \perp v - u$.

The Projection Theorem

Proof of the Projection Theorem (continued).

Note first that

$$\begin{aligned}\|u\|^2 &= \langle u, u \rangle = \left\langle \sum_{i=1}^r \langle e_i, v \rangle e_i, \sum_{j=1}^r \langle e_j, v \rangle e_j \right\rangle \\ &= \sum_{i=1}^r \sum_{j=1}^r \overline{\langle e_i, v \rangle} \langle e_j, v \rangle \underbrace{\langle e_i, e_j \rangle}_{=\delta_{ij}} = \sum_{i=1}^r |\langle e_i, v \rangle|^2.\end{aligned}$$

It then follows that

$$\begin{aligned}\langle v - u, u \rangle &= \langle v, u \rangle - \|u\|^2 = \left\langle v, \sum_{i=1}^r \langle e_i, v \rangle e_i \right\rangle - \|u\|^2 \\ &= \sum_{i=1}^r \langle e_i, v \rangle \overline{\langle e_i, v \rangle} - \sum_{i=1}^r |\langle e_i, v \rangle|^2 = 0.\end{aligned}$$

□

Orthogonal Subspaces

An immediate consequence of the Projection Theorem is as follows:

1.3.24. Corollary. Let $(V, \langle \cdot, \cdot \rangle)$ be a (possibly infinite-dimensional) inner product vector space and let $U \subset V$ be a finite-dimensional subspace.

Then

$$V = U \oplus U^\perp$$

If V is finite-dimensional, then

$$\dim V = \dim U + \dim U^\perp.$$

This follows directly from the Projection Theorem with Lemma 1.2.27 and Theorem 1.2.28.

Bessel's Inequality

As a consequence of the Projection Theorem 1.3.21 and Pythagoras's Theorem 1.3.13 we obtain the following important result:

1.3.25. Bessel Inequality. Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space and (e_1, \dots, e_n) an orthonormal system in V . Then, for any $v \in V$ and any $r \leq n$,

$$\sum_{k=1}^r |\langle e_k, v \rangle|^2 \leq \|v\|^2. \quad (1.3.5)$$

Proof.

By Pythagoras's Theorem 1.3.13 we then have $\|v - u\|^2 + \|u\|^2 = \|v\|^2$, so

$$0 \leq \|v - u\|^2 = \|v\|^2 - \|u\|^2 = \|v\|^2 - \sum_{i=1}^r |\langle e_i, v \rangle|^2. \quad \square$$

Best Approximation

Now suppose that we want to approximate an element $v \in V$ using a linear combination of the first r elements of an orthonormal system,

$$v \approx \sum_{i=1}^r \lambda_i e_i, \quad \lambda_1, \dots, \lambda_r \in \mathbb{F}. \quad (1.3.6)$$

The question is how to choose the coefficients $\lambda_1, \dots, \lambda_r$ to make the approximation “as good as possible”. We note that

$$\begin{aligned} \left\| v - \sum_{i=1}^r \lambda_i e_i \right\|^2 &= \|v\|^2 + \sum_{i=1}^r |\lambda_i|^2 - \sum_{i=1}^r \lambda_i \langle v, e_i \rangle - \sum_{i=1}^r \bar{\lambda}_i \langle e_i, v \rangle \\ &= \|v\|^2 + \sum_{i=1}^r |\langle e_i, v \rangle - \lambda_i|^2 - \sum_{i=1}^r |\langle e_i, v \rangle|^2 \end{aligned} \quad (1.3.7)$$

It is clear that (1.3.7) is minimal if $\lambda_i = \langle e_i, v \rangle$, i.e., the coefficients in (1.3.6) are just the Fourier coefficients.

Best Approximation

From (1.3.7) we can also see that

$$\left\| v - \sum_{i=1}^{r'} \langle e_i, v \rangle e_i \right\| \leq \left\| v - \sum_{i=1}^r \langle e_i, v \rangle e_i \right\| \quad \text{for any } r' \geq r, \quad (1.3.8)$$

so the approximation can only improve when we add further elements of the orthonormal system \mathcal{B} to the approximation.

Clearly, orthonormal systems and bases are extremely useful. We next discuss how to obtain an orthonormal system from any system of vectors.

Gram-Schmidt Orthonormalization

Assume that we have a system of vectors (perhaps a basis) (v_1, \dots, v_n) in an inner product vector space V . We wish to construct a new system (w_1, \dots, w_n) that is orthonormal. We start with v_1 and norm it, defining

$$w_1 := \frac{v_1}{\|v_1\|}$$

Next, we want to obtain from v_2 a vector w_2 such that $w_1 \perp w_2$. By Theorem 1.3.21, v_2 has a unique representation as a sum $v_2 = x + y$, where $x \in \text{span}\{w_1\}$ and $y \in (\text{span}\{w_1\})^\perp$. Now $x = \langle w_1, v_2 \rangle w_1$, so

$$y = v_2 - \langle w_1, v_2 \rangle w_1 \in (\text{span}\{w_1\})^\perp.$$

(Of course, y is independent and even orthogonal to w_1 .) It just remains to norm y , and we define

$$w_2 := \frac{v_2 - \langle w_1, v_2 \rangle w_1}{\|v_2 - \langle w_1, v_2 \rangle w_1\|}.$$

Gram-Schmidt Orthonormalization

Now we can write

$$v_3 = \langle w_1, v_3 \rangle w_1 + \langle w_2, v_3 \rangle w_2 + y,$$

where $y \in (\text{span}\{w_1, w_2\})^\perp$. Thus

$$w_3 := \frac{v_3 - \langle w_2, v_3 \rangle w_2 - \langle w_1, v_3 \rangle w_1}{\|v_3 - \langle w_2, v_3 \rangle w_2 - \langle w_1, v_3 \rangle w_1\|}$$

will be normed and orthogonal to w_1 and w_2 . Proceeding in this way, we set

$$w_1 := \frac{v_1}{\|v_1\|}$$

$$w_k := \frac{v_k - \sum_{j=1}^{k-1} \langle w_j, v_k \rangle w_j}{\|v_k - \sum_{j=1}^{k-1} \langle w_j, v_k \rangle w_j\|}, \quad k = 2, \dots, n,$$

and hence obtain an orthonormal system as desired.

Gram-Schmidt Orthonormalization

1.3.26. Example. Suppose we are given

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}.$$

Then $\|v_1\| = \sqrt{2}$ so

$$w_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Next, we calculate the projection of v_2 onto w_1 and subtract it from v_2 :

$$v_2 - \langle w_1, v_2 \rangle w_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \left(\frac{1}{\sqrt{2}} \cdot 1 + \frac{1}{\sqrt{2}} \cdot 1 \right) \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

Gram-Schmidt Orthonormalization

Since the norm of this vector is already one, we have

$$w_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Next, we calculate

$$v_3 - \langle w_2, v_3 \rangle w_2 - \langle w_1, v_3 \rangle w_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} - 2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

Norming,

$$w_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

Projections and Gram-Schmidt

We can use the **Normalize** command to create a normed vector:

```
v1 = {1, 0, 1};  
w1 = Normalize[v1]
```

$$\left\{ \frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right\}$$

The projection of v_2 onto w_1 can be calculated through the **Projection** command.

```
v2 = {1, 1, 1};  
w2 =  $\frac{v2 - \text{Projection}[v2, w1]}{\text{Norm}[v2 - \text{Projection}[v2, w1]]}$   
{0, 1, 0}
```

Note that addition of vectors and multiplication with numbers work naturally.



Projections and Gram-Schmidt

Mathematica has a built-in command for the Gram-Schmidt procedure,
Orthogonalize:

```
v3 = {1, 2, 0};  
Orthogonalize[{v1, v2, v3}]
```

$$\left\{ \left\{ \frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right\}, \{0, 1, 0\}, \left\{ \frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}} \right\} \right\}$$

Systems of Linear Equations

Finite-Dimensional Vector Spaces

Inner Product Spaces

Linear Maps

Matrices

Theory of Systems of Linear Equations

Determinants

Linear maps on Vector Spaces

In calculus, physics and engineering applications, a fundamental role is played by functions between vector spaces that are linear:

1.4.1. Definition. Let (U, \oplus, \odot) and (V, \boxplus, \boxdot) be vector spaces that are either both real or both complex. Then a map $L: U \rightarrow V$ is said to be **linear** if it is both **homogeneous**, i.e.,

$$L(\lambda \odot u) = \lambda \boxdot L(u) \quad (1.4.1a)$$

and **additive**, i.e.,

$$L(u \oplus u') = L(u) \boxplus L(u'), \quad (1.4.1b)$$

for all $u, u' \in U$ and $\lambda \in \mathbb{F}$. The set of all linear maps $L: U \rightarrow V$ is denoted by $\mathcal{L}(U, V)$.

1.4.2. Remark. A linear map $L: U \rightarrow V$ satisfies $L(0) = 0$, where we use the same symbol 0 for the zero in U or in V .

Linear maps on Vector Spaces

1.4.3. Examples.

- (i) All linear maps $\mathbb{R} \rightarrow \mathbb{R}$ are of the form $x \mapsto \alpha x$ for some $\alpha \in \mathbb{R}$.
- (ii) For $I \subset \mathbb{R}$, the map $\frac{d}{dx}: f \mapsto f'$ is a linear map $C^1(I) \rightarrow C(I)$.
- (iii) The map $(a_n) \mapsto a_0$ is a linear map from the space of all sequences to \mathbb{C} .
- (iv) The map $(a_n) \mapsto \lim_{n \rightarrow \infty} a_n$ is linear map from the space of all convergent sequences to \mathbb{C} .
- (v) If \mathbb{C} is regarded as a real vector space, the map $z \mapsto \bar{z}$ is linear $\mathbb{C} \rightarrow \mathbb{C}$. It is not linear if \mathbb{C} is regarded as a complex vector space.
- (vi) For any real or complex vector space V , the map $V \ni x \mapsto c \in \mathbb{F}$ ($\mathbb{F} = \mathbb{R}$ or \mathbb{C}) is linear if and only if $c = 0$.

For linear maps, we often write simply Lu instead of $L(u)$.

Linear Maps are Structure-Preserving

A linear map

$$L: U \rightarrow V$$

between vector spaces (U, \oplus, \odot) and (V, \boxplus, \boxdot) is a **structure-preserving map**. What does this mean?

Suppose (for the moment) that $L: U \rightarrow V$ is also bijective. Consider the scalar multiplication of a vector $x \in U$ with a number λ . There are now two ways of doing this: Either we calculate $\lambda \odot x$ directly, or we use the map L to form $Lx \in V$, then multiply by λ , then use the inverse map $L^{-1}: V \rightarrow U$ to regain an element of U :

$$\begin{array}{ccc} U & \xrightarrow{L} & V \\ \downarrow \lambda \odot & & \downarrow \lambda \boxdot \\ U & \xleftarrow{L^{-1}} & V \end{array} \tag{1.4.2}$$

Homomorphisms

The validity of (1.4.2) follows from

$$L^{-1}(\lambda \square Lx) = L^{-1}(L(\lambda \odot x)) = \lambda \odot x.$$

From now on, we will use \cdot instead of symbols like \odot or \square and $+$ instead of \oplus or \boxplus . It is up to the reader (you!) to determine which operation in which space is indicated.

Since linear maps have this important property of structure preservation, they deserve an appropriately fancy name: they are also known as (vector space) **homomorphisms** (the greek prefix **homo** means “same”, while **morphos** means “shape”). Thus, “homomorphism” and “linear map” both denote the same thing.

In fact linear maps are so intertwined with the linear structure of vector spaces that in the finite-dimensional case it suffices to know how a linear map acts on basis vectors to determine it completely.

Homomorphisms and Finite-Dimensional Spaces

1.4.4. Theorem. Let U, V be real or complex vector spaces and (b_1, \dots, b_n) a basis of U (in particular, it is assumed that $\dim U = n < \infty$). Then for every n -tuple $(v_1, \dots, v_n) \in V^n$ there exists a unique linear map $L: U \rightarrow V$ such that $Lb_k = v_k$, $k = 1, \dots, n$.

Proof.

We first show the uniqueness of L : Assume there exists a second homomorphism $M \in \mathcal{L}(U, V)$ with $Mb_k = v_k$. For any $u \in U$ we have numbers $\lambda_1, \dots, \lambda_n$ such that $u = \sum \lambda_k b_k$. Then

$$\begin{aligned} Lu &= L(\lambda_1 b_1 + \dots + \lambda_n b_n) = \sum_{k=1}^n \lambda_k L(b_k) \\ &= \sum_{k=1}^n \lambda_k v_k = \sum_{k=1}^n \lambda_k M(b_k) = M(\lambda_1 b_1 + \dots + \lambda_n b_n) = Mu. \end{aligned}$$

Since this is true for any $u \in U$, we have $L = M$.

Homomorphisms and Finite-Dimensional Spaces

Proof (continued).

We now prove the existence of such a linear map, i.e., given the tuple (v_1, \dots, v_n) we want to show how to define L . We define L by defining it for each $u \in U$. Every $u \in U$ has a unique basis decomposition $u = \sum \lambda_k b_k$ with numbers $\lambda_1, \dots, \lambda_n \in \mathbb{F}$. We hence define Lu in the obvious way,

$$Lu := \sum_{k=1}^n \lambda_k v_k.$$

It remains to check that L is linear: if $u, u' \in U$ have coordinates $(\lambda_k)_{k=1}^n$ and $(\lambda'_k)_{k=1}^n$, respectively, we have

$$L(u + u') = \sum_{k=1}^n (\lambda_k + \lambda'_k) v_k = \sum_{k=1}^n \lambda_k v_k + \sum_{k=1}^n \lambda'_k v_k = Lu + Lu'.$$

The homogeneity of L can be shown similarly. □

Coordinate Map

1.4.5. Remarks.

- (i) The identity map $\text{id}: V \rightarrow V$, $\text{id}(v) = v$, is linear.
- (ii) The set $\mathcal{L}(U, V)$ is again a vector space when endowed with pointwise addition and scalar multiplication.
- (iii) If $L_1 \in \mathcal{L}(U, V)$ and $L_2 \in \mathcal{L}(V, W)$, then $L_2 \circ L_1 \in \mathcal{L}(U, W)$. (The composition of linear maps is linear.)

1.4.6. Examples.

- (i) If V is a real or complex vector space and (b_1, \dots, b_n) a basis of V , then the **coordinate map**

$$\varphi: V \rightarrow \mathbb{F}^n, \quad v = \sum_{k=1}^n \lambda_k b_k \mapsto \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$$

is linear (and bijective).

Dual Space

1.4.7. Examples.

- (ii) Let V be a real or complex vector space. Then $\mathcal{L}(V, \mathbb{F})$ is known as the **dual space** of V and denoted by V^* . The dual space of V is of course itself a vector space.

Let $\dim V = n < \infty$ and $\mathcal{B} = (b_1, \dots, b_n)$ be a basis of V . Then for every $k = 1, \dots, n$ there exists a unique map

$$b_k^*: V \rightarrow \mathbb{F}, \quad b_k^*(b_j) = \delta_{jk} = \begin{cases} 1, & j = k, \\ 0, & j \neq k. \end{cases}$$

It turns out (see exercises) that the tuple of maps $\mathcal{B}^* = (b_1^*, \dots, b_n^*)$ is a basis of $V^* = \mathcal{L}(V, \mathbb{F})$ (called the **dual basis** of \mathcal{B}) and thus $\dim V^* = \dim V = n$.

Range and Kernel

1.4.8. Definition. Let U, V be real or complex vector spaces and $L \in \mathcal{L}(U, V)$. Then we define the range of L by

$$\text{ran } L := \left\{ v \in V : \exists_{u \in U} v = Lu \right\}$$

and the **kernel** of L by

$$\ker L := \{u \in U : Lu = 0\}.$$

It is easy to see that $\text{ran } L \subset V$ and $\ker L \subset U$ are subspaces.

1.4.9. Remark. It is not difficult to see that $L \in \mathcal{L}(U, V)$ is injective if and only if $\ker L = \{0\}$.

Nomenclature

According to their properties, there are several fancy names for linear maps. A homomorphism $L \in \mathcal{L}(U, V)$ is said to be

- ▶ an **isomorphism** if L is bijective;
- ▶ an **endomorphism** if $U = V$;
- ▶ an **automorphism** if $U = V$ and L is bijective;
- ▶ **epimorph** if L is surjective;
- ▶ **monomorph** if L is injective.

1.4.10. Remark. If L is an isomorphism, then its inverse, L^{-1} is also linear and hence also an isomorphism.

Isomorphisms

1.4.11. Theorem. Let U, V be finite-dimensional vector spaces and $L \in \mathcal{L}(U, V)$. Then L is an isomorphism if and only if for every basis (b_1, \dots, b_n) of U the tuple (Lb_1, \dots, Lb_n) is a basis of V .

Proof.

(\Rightarrow) Assume that L is bijective. Then for $y \in V$ the pre-image $x = L^{-1}y$ is uniquely determined. Let $x = \sum \lambda_k b_k$ be the representation of x in the basis $\mathcal{B} = (b_1, \dots, b_n)$. Now

$$y = L\left(\sum_{k=1}^n \lambda_k b_k\right) = \sum_{k=1}^n \lambda_k \cdot Lb_k$$

where the λ_k are uniquely determined by x , which is uniquely determined by y . Thus for any y we can find a representation in terms of (Lb_1, \dots, Lb_n) by considering the pre-image $x = L^{-1}y$.

Isomorphisms

Proof (continued).

We still need to show that this representation is unique, i.e., if $y = \sum \mu_k \cdot Lb_k$, then $\mu_k = \lambda_k$. Applying L^{-1} , we see that

$$L^{-1}y = x = \sum_{k=1}^n \lambda_k b_k, \quad L^{-1}y = L^{-1} \sum_{k=1}^n \mu_k \cdot Lb_k = \sum_{k=1}^n \mu_k b_k$$

and because (b_1, \dots, b_n) is a basis we see that $\mu_k = \lambda_k$.

(\Leftarrow) We need to show that L is injective and surjective. Since any $y \in V$ may be written as $y = \sum \lambda_k \cdot Lb_k$, y is obviously the image of $x = \sum \lambda_k b_k \in U$. Thus L is surjective.

To see that L is injective, we show that $\ker L = \{0\}$ (see Remark 1.4.9). Now $Lx = 0$ for $x = \sum \lambda_k b_k$ implies $\sum \lambda_k \cdot Lb_k = 0$. Since (Lb_1, \dots, Lb_n) is a basis, this means that $\lambda_1 = \dots = \lambda_n = 0$, so $x = 0$. □

Isomorphisms

1.4.12. Definition. Two vector spaces U and V are called **isomorphic**, written $U \cong V$, if there exists an isomorphism $\varphi: U \rightarrow V$.

1.4.13. Lemma. Two finite-dimensional vector spaces U and V are isomorphic if and only if they have the same dimension:

$$U \cong V \iff \dim U = \dim V$$

Proof.

(\Rightarrow) Let $\varphi: U \rightarrow V$ be an isomorphism and (b_1, \dots, b_n) a basis of U ($\dim U = n$). Then $(\varphi(b_1), \dots, \varphi(b_n))$ is a basis of V and thus $\dim V = n = \dim U$.

(\Leftarrow) If (a_1, \dots, a_n) and (b_1, \dots, b_n) are bases of U and V , respectively, define an isomorphism φ by $\varphi(a_k) = b_k$, $k = 1, \dots, n$. □

The Dimension Formula

We can now prove a deep and fundamental result on linear maps:

1.4.14. Dimension Formula. Let U, V be real or complex vector spaces, $\dim U < \infty$. Let $L \in \mathcal{L}(U, V)$. Then

$$\dim \text{ran } L + \dim \ker L = \dim U. \quad (1.4.3)$$

Proof.

Let $\dim U =: n < \infty$. Since $\ker L \subset U$, we have $\dim \ker L =: r \leq n$. We choose a basis (a_1, \dots, a_r) of the kernel, and use the Basis Completion Theorem 1.2.21 to construct a basis $(a_1, \dots, a_r, a_{r+1}, \dots, a_n)$ of U . Then for any $x = \sum \lambda_k a_k \in U$,

$$Lx = L(\lambda_1 a_1 + \dots + \lambda_n a_n) = \lambda_{r+1} \underbrace{La_{r+1}}_{=: b_1} + \dots + \lambda_n \underbrace{La_n}_{=: b_{n-r}}.$$

Thus $\text{ran } L = \text{span}\{b_1, \dots, b_{n-r}\}$.

The Dimension Formula

Proof.

We now claim that the vectors b_1, \dots, b_{n-r} are independent; in that case they form a basis of $\text{ran } L$ and $\dim \text{ran } L = n - r$, proving (1.4.3). Consider the equality

$$0 = \mu_1 b_1 + \cdots + \mu_{n-r} b_{n-r} = L(\mu_1 a_{r+1} + \cdots + \mu_{n-r} a_n). \quad (1.4.4)$$

If (1.4.4) holds, then $\mu_1 a_{r+1} + \cdots + \mu_{n-r} a_n \in \ker L = \text{span}\{a_1, \dots, a_r\}$. Thus, there exist $\lambda_1, \dots, \lambda_r$ such that

$$\mu_1 a_{r+1} + \cdots + \mu_{n-r} a_n - (\lambda_1 a_1 + \cdots + \lambda_r a_r) = 0.$$

Since (a_1, \dots, a_n) is a basis of U , we thence obtain

$$\mu_1 = \cdots = \mu_{n-r} = 0, \quad \lambda_1 = \cdots = \lambda_r = 0. \quad (1.4.5)$$

Thus (1.4.4) implies (1.4.5) and b_1, \dots, b_{n-r} are independent. □

The Dimension Formula

1.4.15. Corollary. Let U, V be real or complex finite-dimensional vector spaces with $\dim U = \dim V$. Then a linear map $L \in \mathcal{L}(U, V)$ is injective if and only if it is surjective.

Proof.

$$\begin{aligned} L \text{ injective} &\Leftrightarrow \ker L = \{0\} \\ &\Leftrightarrow \dim \ker L = 0 \\ &\Leftrightarrow \dim \text{ran } L = \dim U = \dim V \\ &\Leftrightarrow \text{ran } L = V \\ &\Leftrightarrow L \text{ surjective} \end{aligned}$$



Normed Vector Spaces and Bounded Linear Maps

1.4.16. Definition. Let $(U, \|\cdot\|_U)$ and $(V, \|\cdot\|_V)$ be normed vector spaces. Then a linear map $L: U \rightarrow V$ is said to be **bounded** if there exists some constant $c > 0$ (called a **bound** for L) such that

$$\|Lu\|_V \leq c \cdot \|u\|_U \quad \text{for all } u \in U. \quad (1.4.6)$$

1.4.17. Remark. It can be shown that if U is a finite-dimensional vector space, then any linear map is bounded.

1.4.18. Examples.

1. The map $L_\alpha: \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto \alpha x$ is bounded with $c = |\alpha|$.

Bounded Linear Maps

2. The map

$$L: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} 2x_2 \\ -x_1 \end{pmatrix}$$

is linear and bounded. If we take $\|x\|_2 = \sqrt{x_1^2 + x_2^2}$, we can see that $c = 2$ is a bound for L .

3. Take the space $C^1([0, 1])$ of the continuously differentiable functions on the interval $[0, 1]$ and imbue it with the norm given by $\|f\|_\infty = \sup_{x \in [0, 1]} |f(x)|$. Then the map

$$\frac{d}{dx}: C^1([0, 1]) \rightarrow C([0, 1]), \quad f \mapsto f'$$

is **not bounded**. To see this, consider the function $f(x) = e^{-nx}$ for $n \in \mathbb{N}$. Clearly, $\|f\|_\infty = 1$ but $\|f'\|_\infty = n$. Since we can choose n as large as we like, there can exist no $c > 0$ such that $\|f'\|_\infty \leq c \cdot \|f\|_\infty$.

The Operator Norm

By (1.4.6), for every bounded linear map there exists an upper bound $c > 0$ such that

$$\frac{\|Lu\|_V}{\|u\|_U} \leq c \quad \text{for } u \neq 0.$$

We are now interested in the **least upper bound** c .

1.4.19. Definition and Theorem. Let U, V be normed vector spaces. Then the set of bounded linear maps $\mathcal{L}(U, V)$ is also a vector space and

$$\|L\| := \sup_{\substack{u \in U \\ u \neq 0}} \frac{\|Lu\|_V}{\|u\|_U} = \sup_{\substack{u \in U \\ \|u\|_U=1}} \|Lu\|_V. \quad (1.4.7)$$

defines a norm, the so-called **operator norm** or **induced norm** on $\mathcal{L}(U, V)$.

The proof of the norm properties is left to the reader. The operator norm also has the additional, very useful, property that

$$\|L_2 L_1\| \leq \|L_2\| \cdot \|L_1\|, \quad L_1 \in \mathcal{L}(U, V), \quad L_2 \in \mathcal{L}(V, W).$$

Systems of Linear Equations

Finite-Dimensional Vector Spaces

Inner Product Spaces

Linear Maps

Matrices

Theory of Systems of Linear Equations

Determinants

A Calculus of Linear Maps

We have seen in Lemma 1.4.13 that two vector spaces are isomorphic if their dimensions are equal. In particular:

- ▶ Every real n -dimensional vector space is isomorphic to \mathbb{R}^n
- ▶ Every complex n -dimensional vector space is isomorphic to $\mathbb{C}^n \cong \mathbb{R}^{2n}$

This means that if we can find a **calculus** for linear maps $\mathbb{R}^n \rightarrow \mathbb{R}^m$, we can automatically treat maps from an n -dimensional space U to an m -dimensional space V :

$$\begin{array}{ccc} U & \xrightarrow{L} & V \\ \varphi_1 \downarrow & & \downarrow \varphi_2 \\ \mathbb{R}^n & \xrightarrow{A} & \mathbb{R}^m \end{array}$$

Here $L \in \mathcal{L}(U, V)$, φ_1, φ_2 are isomorphisms and $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. If

$$L = \varphi_2^{-1} \circ A \circ \varphi_1$$

we obtain all relevant properties of L (range, kernel) by analyzing A .

A Calculus of Linear Maps

The word **calculus** means a “scheme of calculating” that transforms a procedure that otherwise needs to be performed individually into an algorithm that can be (easily) applied in general.

For example, the revolutionary aspect of Newton/Leibniz’s calculus was the fact that areas under curves, which earlier had been calculated by hand for each individual type of curve, could suddenly easily be computed through inverse differentiation (by finding the primitive of a function).

This was exemplified by the fundamental theorem of calculus. As an example, compare Exercises 4 and 19 of Chapter 14 of Spivak’s book with the simplicity of applying the Fundamental Theorem of Calculus.

In the following, we will establish an analogous calculus for linear maps, where we are able, due to Lemma 1.4.13, to concentrate on those in $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$.

Matrices

1.5.1. Definition. An $m \times n$ matrix over the complex numbers is a map

$$a: \{1, \dots, m\} \times \{1, \dots, n\} \rightarrow \mathbb{C}, \quad (i, j) \mapsto a_{ij}.$$

We represent the graph of a through

$$A := \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = (a_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}.$$

We denote the set of all $m \times n$ matrices over \mathbb{C} by $\text{Mat}(m \times n; \mathbb{C})$.

1.5.2. Remarks.

1. With the usual pointwise addition and scalar multiplication of maps, $\text{Mat}(m \times n; \mathbb{C})$ becomes a complex vector space.
2. Matrices over \mathbb{R} instead of \mathbb{C} are defined in the same way.
Occasionally, we may also replace \mathbb{C} by a real or complex vector space.

Matrices as Linear Maps

Matrices turn out to be important tools in the analysis of linear maps: every linear map between finite-dimensional vector spaces may be expressed as a matrix, and every matrix corresponds (in a certain way) to some such linear map. We first restrict ourselves to the case of \mathbb{R}^n .

1.5.3. Theorem. Each matrix $A \in \text{Mat}(m \times n; \mathbb{R})$ uniquely determines a linear map $j(A) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ such that the columns $a_{\cdot k}$ are the images of the standard basis vectors $e_k \in \mathbb{R}^n$; in particular,

$$j: \text{Mat}(m \times n; \mathbb{R}) \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$$

is an isomorphism, $\text{Mat}(m \times n; \mathbb{R}) \cong \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, so every map $L \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ corresponds to a matrix $j^{-1}(L)$ whose columns $a_{\cdot k}$ are the images of the standard basis vectors $e_k \in \mathbb{R}^n$.

Matrices as Linear Maps

Proof.

Given a matrix A with columns $a_{\cdot k}$, $k = 1, \dots, n$, we simply define $j(A)$ by

$$j(A) : \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad e_k \mapsto a_{\cdot k}, \quad k = 1, \dots, n.$$

Given a map $L \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ we define $j^{-1}(L) \in \text{Mat}(m \times n; \mathbb{R})$ by

$$j^{-1}(L) = (a_{\cdot 1}, \dots, a_{\cdot n}), \quad a_{\cdot k} = L(e_k), \quad k = 1, \dots, n.$$

Obviously, j^{-1} is actually the inverse of j ; hence j is bijective. It remains to show that j is linear. Let $A = (a_{ik})$, $B = (b_{ik})$. Then

$$j(A + B)e_k = (a + b)_{\cdot k} = a_{\cdot k} + b_{\cdot k} = j(A)e_k + j(B)e_k,$$

so j is additive. The homogeneity can be shown analogously. □

Matrices as Linear Maps

We have thus established that every matrix $A = (a_{ik})$ represents a linear map $j(A)$. In particular,

$$j(A)e_k = \begin{pmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{pmatrix} = \sum_{i=1}^n a_{ik} e_i, \quad k = 1, \dots, n.$$

We also note that we can represent $x \in \mathbb{R}^n$ as

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + x_n \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \sum_{k=1}^n x_k e_k.$$

Matrices as Linear Maps

Then $j(A) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ acts on a general $x \in \mathbb{R}^n$ as follows,

$$\begin{aligned} j(A)x &= j(A)\left(\sum_{k=1}^n x_k e_k\right) = \sum_{k=1}^n x_k j(A)e_k \\ &= \sum_{k=1}^n x_k \begin{pmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{pmatrix} \\ &= \begin{pmatrix} x_1 a_{11} + \cdots + x_n a_{1n} \\ \vdots \\ x_1 a_{m1} + \cdots + x_n a_{mn} \end{pmatrix}. \end{aligned}$$

Matrices as Linear Maps

From a practical point of view, we start with $A \in \text{Mat}(m \times n; \mathbb{R})$ and some $x \in \mathbb{R}^n$ and obtain

$$\begin{pmatrix} x_1 a_{11} + \cdots + x_n a_{1n} \\ \vdots \\ x_1 a_{m1} + \cdots + x_n a_{mn} \end{pmatrix} \in \mathbb{R}^m.$$

It seems unnecessary to include the map $j: \text{Mat}(n \times m; \mathbb{R}) \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ in this. In fact, we might directly **interpret the matrix A as a linear map without mentioning j !**

The isomorphism j can be simply left out; mathematicians routinely consider sets of objects that are isomorphic as being **actually identical**. In this way, A has a **double meaning**: it is on the one hand a matrix, and on the other hand a linear map. This avoids always mentioning a superfluous isomorphism and greatly simplifies the formulation of statements.

Matrices as Linear Maps

We therefore write Ax instead of $j(A)x$; in particular,

$$Ax = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 a_{11} + \cdots + x_n a_{1n} \\ \vdots \\ x_1 a_{m1} + \cdots + x_n a_{mn} \end{pmatrix}. \quad (1.5.1)$$

We can interpret (1.5.1) as the action of a matrix $A \in \text{Mat}(m \times n; \mathbb{R})$ on a vector $x \in \mathbb{R}^n$, yielding a vector $Ax \in \mathbb{R}^m$.

This is the beginning of our ***calculus of linear maps***. We now need to develop this further to deal with (e.g.) compositions and inverses of linear maps.

Compositions

Let $\mathbb{R}^n \xrightarrow{j(A)} \mathbb{R}^m \xrightarrow{j(B)} \mathbb{R}^l$ be linear maps and consider their composition $j(B) \circ j(A)$. We want to find a matrix C such that $j(B) \circ j(A) = j(C)$. Now

$$\begin{aligned} j(B) \circ j(A)e_k &= j(B) \sum_{s=1}^m a_{sk} e_s = \sum_{s=1}^m a_{sk} j(B)e_s = \sum_{s=1}^m a_{sk} \sum_{t=1}^l b_{ts} e_t \\ &= \sum_{t=1}^l \underbrace{\left(\sum_{s=1}^m b_{ts} a_{sk} \right)}_{=: c_{tk}} e_t \end{aligned}$$

where $C = (c_{tk}) \in \text{Mat}(l \times n; \mathbb{R})$. We thus introduce C as the **matrix product** of B and A .

Matrix Product

1.5.4. Definition. Let $A \in \text{Mat}(l \times m; \mathbb{C})$ and $B \in \text{Mat}(m \times n; \mathbb{C})$. Then we define the **product of A** = (a_{ik}) **and B** = (b_{kj}) by

$$AB \in \text{Mat}(l \times n; \mathbb{C}), \quad AB := \left(\sum_{k=1}^m a_{ik} b_{kj} \right)_{i=1, \dots, l; j=1, \dots, n}$$

We have seen that the matrix product satisfies $j(A) \circ j(B) = j(AB)$. Furthermore, the product is **associative**, i.e.,

$$\begin{aligned} A(BC) &= j^{-1}(j(A) \circ j(BC)) = j^{-1}(j(A) \circ (j(B) \circ j(C))) \\ &= j^{-1}((j(A) \circ j(B)) \circ j(C)) = j^{-1}(j(AB) \circ j(C)) \\ &= (AB)C \end{aligned}$$

If $A, B \in \text{Mat}(n \times n; \mathbb{C})$ both products AB and BA exist; however

$$AB \neq BA,$$

so the matrix product is **not commutative**.

Matrix Product

The matrix product is easily memorized through “row-by-column multiplication,” as seen in the following examples:

1.5.5. Examples.

$$1. \quad A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 5 & 6 & 7 \\ 1 & 0 & 2 \end{pmatrix},$$

$$AB = \begin{pmatrix} 1 \cdot 5 + 2 \cdot 1 & 1 \cdot 6 + 2 \cdot 0 & 1 \cdot 7 + 2 \cdot 2 \\ 3 \cdot 5 + 4 \cdot 1 & 3 \cdot 6 + 4 \cdot 0 & 3 \cdot 7 + 4 \cdot 2 \end{pmatrix} = \begin{pmatrix} 7 & 6 & 11 \\ 19 & 18 & 29 \end{pmatrix}$$

$$2. \quad A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}, \quad AB = \begin{pmatrix} 3 & 1 \\ 6 & 2 \end{pmatrix}, \quad BA = \begin{pmatrix} 4 & 4 \\ 1 & 1 \end{pmatrix}$$

$$3. \quad A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}, \quad AB = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} = BA$$

$$4. \quad A = \begin{pmatrix} 0 & \alpha \\ 0 & 0 \end{pmatrix}, \quad A^2 = AA = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \forall \alpha \in \mathbb{C}$$

Matrix Transpose

For $A = (a_{ij}) \in \text{Mat}(m \times n; \mathbb{F})$ we define the **transpose** of A by

$$A^T \in \text{Mat}(n \times m; \mathbb{F}), \quad A^T = (a_{ji}).$$

For example,

$$\begin{pmatrix} 5 & 6 & 7 \\ 1 & 0 & 2 \end{pmatrix}^T = \begin{pmatrix} 5 & 1 \\ 6 & 0 \\ 7 & 2 \end{pmatrix}.$$

We also define the **adjoint**

$$A^* \in \text{Mat}(n \times m; \mathbb{F}), \quad A^* = \overline{A}^T = (\overline{a_{ji}}).$$

where in addition to the transpose the complex conjugate of each entry is taken.

It is easy to see (in the assignments) that for $A \in \text{Mat}(m \times n; \mathbb{F})$, $x \in \mathbb{F}^m$, $y \in \mathbb{F}^n$,

$$\langle x, Ay \rangle = \langle A^*x, y \rangle.$$

Matrices

In Mathematica, a matrix is defined as follows:

```
In[6]:= A = Table[ai,j, {i, 4}, {j, 3}]
```

```
Out[6]= {{a1,1, a1,2, a1,3}, {a2,1, a2,2, a2,3}, {a3,1, a3,2, a3,3}, {a4,1, a4,2, a4,3}}
```

The `MatrixForm` command can be used for nicer formatting.

```
In[7]:= MatrixForm[A]
```

```
Out[7]//MatrixForm=
```

$$\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \\ a_{4,1} & a_{4,2} & a_{4,3} \end{pmatrix}$$



Matrix Multiplication

Matrix multiplication works using the same dot as for the inner product:

```
In[3]:= A = {{1, 1}, {2, 2}};  
B = {{2, 1}, {1, 0}};  
MatrixForm[A.B]
```

Out[5]/MatrixForm=

$$\begin{pmatrix} 3 & 1 \\ 6 & 2 \end{pmatrix}$$

The `Transpose` command gives the transpose:

```
MatrixForm[Transpose[A.B]]
```

$$\begin{pmatrix} 3 & 6 \\ 1 & 2 \end{pmatrix}$$

Matrix Multiplication

There are two very useful facts to keep in mind:

- (i) When a vector $x \in \mathbb{R}^n$ is multiplied by a matrix $A \in \text{Mat}(m \times n)$, the result is a linear combination of the column vectors of A . For illustration, in the case of $n = 3$ and $m = 2$, we have

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \end{pmatrix}$$
$$= x_1 \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} + x_3 \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix}$$

- (ii) When a matrix B is multiplied by a matrix A , the result is a matrix whose columns are the products of the columns of B multiplied with A . Again, for illustration, we give a simple example.

Matrix Multiplication

Write

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix} = (b_1, b_2, b_3)$$

where

$$b_1 = \begin{pmatrix} b_{11} \\ b_{21} \end{pmatrix}, \quad b_2 = \begin{pmatrix} b_{12} \\ b_{22} \end{pmatrix}, \quad b_3 = \begin{pmatrix} b_{13} \\ b_{23} \end{pmatrix}.$$

Then

$$\begin{aligned} AB &= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix} \\ &= \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} & a_{11}b_{13} + a_{12}b_{23} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} & a_{21}b_{13} + a_{22}b_{23} \end{pmatrix} \\ &= (Ab_1, Ab_2, Ab_3). \end{aligned}$$

Matrix of a Linear Map

We are now able to properly define the matrix of a linear map between two finite-dimensional vector spaces.

Let U, V be finite-dimensional real or complex vector spaces with bases

$$\mathcal{A} = (a_1, \dots, a_n) \subset U \quad \text{and} \quad \mathcal{B} = (b_1, \dots, b_m) \subset V.$$

Define the isomorphisms

$$\varphi_{\mathcal{A}}: U \xrightarrow{\cong} \mathbb{R}^n, \quad \varphi_{\mathcal{A}}(a_j) = e_j, \quad j = 1, \dots, n,$$

$$\varphi_{\mathcal{B}}: V \xrightarrow{\cong} \mathbb{R}^m, \quad \varphi_{\mathcal{B}}(b_j) = e_j, \quad j = 1, \dots, m.$$

Then any linear map $L \in \mathcal{L}(U, V)$ induces a matrix $A = \Phi_{\mathcal{A}}^{\mathcal{B}}(L) \in \text{Mat}(m \times n; \mathbb{R})$ through

$$\begin{array}{ccc}
 U & \xrightarrow{L} & V \\
 \varphi_{\mathcal{A}} \downarrow & & \downarrow \varphi_{\mathcal{B}} \\
 \mathbb{R}^n & \xrightarrow{A} & \mathbb{R}^m
 \end{array}
 \qquad
 \Phi_{\mathcal{A}}^{\mathcal{B}}(L) = A = \varphi_{\mathcal{B}} \circ L \circ \varphi_{\mathcal{A}}^{-1}$$

Matrix of Complex Conjugation

1.5.6. Example. Consider \mathbb{C} as a real two-dimensional vector space with basis $\mathcal{B} = (1, i)$. The complex conjugation $L: \mathbb{C} \rightarrow \mathbb{C}$, $z \mapsto \bar{z}$ is then a linear map. We want to determine the matrix of this map with respect to the basis \mathcal{B} . The isomorphism is

$$\varphi_{\mathcal{B}}: \mathbb{C} \rightarrow \mathbb{R}^2, \quad 1 \mapsto \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad i \mapsto \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Thus $\varphi_{\mathcal{B}}(a + bi) = \begin{pmatrix} a \\ b \end{pmatrix}$. The most convenient way to determine $A = \Phi_{\mathcal{B}}^{\mathcal{B}}(L)$ is to calculate

$$\varphi_{\mathcal{B}}(a + bi) = \begin{pmatrix} a \\ b \end{pmatrix}, \quad \varphi_{\mathcal{B}}(L(a + bi)) = \varphi_{\mathcal{B}}(a - bi) = \begin{pmatrix} a \\ -b \end{pmatrix}$$

and then find $A \in \text{Mat}(2 \times 2; \mathbb{R})$ such that $A \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a \\ -b \end{pmatrix}$. It is easily seen that

$$A = \Phi_{\mathcal{B}}^{\mathcal{B}}(L) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Matrix of Complex Conjugation

1.5.7. Example. If we change the basis we used in the previous example, we get a different matrix. Let us take the basis $\mathcal{A} = (1+i, 1-i)$ for \mathbb{C} . Then the isomorphism is

$$\varphi_{\mathcal{A}}: \mathbb{C} \rightarrow \mathbb{R}^2, \quad 1+i \mapsto \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad 1-i \mapsto \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Thus

$$\begin{aligned}\varphi_{\mathcal{A}}(a+bi) &= \frac{1}{2}\varphi_{\mathcal{A}}((a+b)(1+i) + (a-b)(1-i)) \\ &= \frac{a+b}{2}\varphi_{\mathcal{A}}(1+i) + \frac{a-b}{2}\varphi_{\mathcal{A}}(1-i) = \frac{1}{2} \begin{pmatrix} a+b \\ a-b \end{pmatrix}.\end{aligned}$$

Hence we need to find $A \in \text{Mat}(2 \times 2; \mathbb{R})$ such that $A \begin{pmatrix} a+b \\ a-b \end{pmatrix} = \begin{pmatrix} a-b \\ a+b \end{pmatrix}$, i.e.,

$$A = \Phi_{\mathcal{A}}^{\mathcal{A}}(L) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Systems of Equations

Before we proceed, we take a step back to the beginning of the course.
Recall that a system of linear equations was given by

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{1.5.2}$$

We can express (1.5.2) using vectors and matrices by writing

$$Ax = b,$$

where

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Elementary Matrix Manipulations

The Gauß-Jordan algorithm introduced ***elementary row manipulation***, which we now reformulate in the context of matrices:

1.5.8. Elementary Matrix Manipulations. An elementary row manipulation of a matrix is one of the following:

- (i) Swapping (interchanging) of two rows,
- (ii) Multiplication of a row with a non-zero number,
- (iii) Addition of a multiple of one row to another row.

The additions and multiplications are performed componentwise in each row.

If the word “row” is replaced by “column” these operations are termed elementary column operations

The Gauß-Jordan algorithm uses only row operations. We seek to find matrices that implement these row manipulations through multiplication of $Ax = b$ from the left.

Elementary Matrix Manipulations

For illustration, we consider the case $n = 4, m = 3$.

Consider the most trivial operation possible: we do nothing. This would be represented by multiplying with the unit matrix,

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

and

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix}$$

Note that we do not need to mention x at all in these equations! This is the true underlying philosophy of the notational scheme used in the Gauß-Jordan algorithm.

Elementary Matrix Manipulations

Now how would we swap the first and second row? We can see that

$$\underbrace{\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{=:S_{12}} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} b_2 \\ b_1 \\ b_3 \end{pmatrix}$$

and

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix} = \begin{pmatrix} a_{21} & a_{22} & a_{23} & a_{24} \\ a_{11} & a_{12} & a_{13} & a_{14} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix}$$

Note that we have swapped the first and second row of the unit matrix to obtain S_{12} !

Elementary Matrix Manipulations

Furthermore, in order to add 3 times the second row to the third row, we would use

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 + 3b_2 \end{pmatrix}$$

and

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} + 3a_{21} & a_{32} + 3a_{22} & a_{33} + 3a_{23} & a_{34} + 3a_{24} \end{pmatrix}$$

Again we have performed the elementary row operation on the unit matrix to obtain the matrix that implements this operation.

Elementary Matrix Manipulations

In conclusion, we remark that

- (i) An elementary row operation on a system of equations may be simply considered as a multiplication of $Ax = b$ from the left with a suitable matrix, a so-called **elementary matrix**.
- (ii) An elementary matrix is obtained by applying the desired elementary row operation to the unit matrix. [Why must this be the case?]
- (iii) If we apply two elementary operations, the product of their respective matrices gives the matrix corresponding to these two operations, in order.

This means that in solving a system $Ax = b$, the sum of all row operations in forward elimination and backward substitution may be represented by a single matrix $S \in \text{Mat}(m \times m; \mathbb{R})$. We thus have

$$SAx = Sb.$$

If $m = n$, the system $Ax = b$ may have a unique solution; in that case, SA is a diagonal matrix of the form (1.1.3), i.e., $SA = \text{id}$.

Inverse of a Matrix

Let us now return to the question of finding the inverse of a linear map $A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ (of course, we must assume that A is an isomorphism, so $m = n$). Of course, we say that a matrix is invertible if the corresponding linear map is invertible and the inverse of a matrix is just the matrix of the inverse map. However, it may be useful to clarify this.

1.5.9. Definition. A matrix $A \in \text{Mat}(n \times n; \mathbb{R})$ is called **invertible** if there exists some $B \in \text{Mat}(n \times n; \mathbb{R})$ such that

$$AB = BA = \text{id} = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}. \quad (1.5.3)$$

We then write $B = A^{-1}$ and say that A^{-1} is the **inverse** of A .

1.5.10. Remark. The inverse is of course unique; if B and \tilde{B} both satisfy (1.5.3) for some A , then

$$B = (\tilde{B}A)B = \tilde{B}(AB) = \tilde{B}.$$

Inverse of a Matrix

1.5.11. Remark. It is obvious that the matrix S corresponding to a series of elementary row manipulations will be invertible, because the operations themselves are invertible. Thus the matrix $S: \mathbb{R}^m \rightarrow \mathbb{R}^m$ represents an isomorphism.

1.5.12. Remark. Given a matrix $A \in \text{Mat}(n \times n; \mathbb{R})$ (identified with a linear map $L \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$) and a putative inverse matrix $B = A^{-1} \in \text{Mat}(n \times n; \mathbb{R})$ it is sufficient to verify that

$$BA = \text{id}.$$

In this case, B corresponds to a linear map M such that $M \circ L$ is the identity map. Thus $\dim \text{ran } M = n$, so M is bijective. Hence M is invertible and $L = M^{-1}$ is bijective. Then $L \circ M = M \circ L = \text{id}$, so we have $AB = BA = \text{id}$.

Inverse of a Matrix

1.5.13. Lemma. Let $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$. Then A is **invertible** if and only if there exists an elementary matrix S corresponding to elementary row operations that transform A into the unit matrix $SA = \text{id}$.

Proof.

(\Rightarrow) If A is bijective, for every $y \in \mathbb{R}^n$ there exists a unique solution x to $Ax = y$. Thus there exists a matrix S corresponding to row operations such that

$$SAx = x = Sy.$$

For every x there exists a unique y such that $y = Ax$. Thus $SAx = x$ for every $x \in \mathbb{R}^n$, and so $SA = \text{id}$.

(\Leftarrow) By Remark 1.5.12, $SA = AS = \text{id}$, and by Remark 1.5.10, $S = A^{-1}$ so A is invertible. □

Finding the Inverse

Lemma 1.5.13 tells us how to actually find the inverse of a matrix A : it is simply the elementary matrix that transforms A into the unit matrix. If this transformation is not possible, A is not invertible.

1.5.14. Example. Consider the matrix

$$A = \begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix}.$$

In order to find the inverse, we transform A into the unit matrix through a sequence of elementary row operations S , keeping track of the elementary matrix that implements these operations.

Finding the Inverse

$$\left| \begin{array}{cc|cc}
 SA & S \\
 \left(\begin{array}{cc} 2 & 3 \\ 2 & 1 \end{array} \right) & \left(\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right) & \xrightarrow{\substack{| : 2 \\ \leftarrow : (-2)}} & \\
 \left(\begin{array}{cc} 1 & 3/2 \\ 0 & -2 \end{array} \right) & \left(\begin{array}{cc} 1/2 & 0 \\ -1 & 1 \end{array} \right) & \xrightarrow{\substack{| : (-2) \\ \leftarrow : (-3/2)}} & \\
 \left(\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right) & \left(\begin{array}{cc} -1/4 & 3/4 \\ 1/2 & -1/2 \end{array} \right) & & \\
 & & = A^{-1} &
 \end{array} \right|$$

We may immediately check that

$$A^{-1}A = \begin{pmatrix} -1/4 & 3/4 \\ 1/2 & -1/2 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$AA^{-1} = \begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} -1/4 & 3/4 \\ 1/2 & -1/2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

 Matrix Inverse

Mathematica has a command for finding the inverse:

```
In[7]:= MatrixForm[Inverse[{{2, 3}, {2, 1}}]]
```

```
Out[7]//MatrixForm=
```

$$\begin{pmatrix} -\frac{1}{4} & \frac{3}{4} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$$

Inverse Maps

1.5.15. Remark. We note that if $A, B \in \text{Mat}(n \times n; \mathbb{R})$ are invertible, then so is their product $AB \in \text{Mat}(n \times n; \mathbb{R})$ and $(AB)^{-1} = B^{-1}A^{-1}$.

We can use this procedure to find the inverse of any vector space isomorphism L :

$$\begin{array}{ccc}
 U & \xrightarrow{L} & V \\
 \varphi_{\mathcal{A}} \downarrow & & \downarrow \varphi_{\mathcal{B}} \\
 \mathbb{R}^n & \xrightarrow{A} & \mathbb{R}^m
 \end{array}
 \qquad L^{-1} = \varphi_{\mathcal{A}}^{-1} \circ A^{-1} \circ \varphi_{\mathcal{B}}$$

1.5.16. Example. Let \mathcal{P}_2 be the space of polynomials of degree not more than 2. Consider the linear map

$$L: \mathcal{P}_2 \rightarrow \mathcal{P}_2, \quad ax^2 + bx + c \mapsto \frac{a+b+c}{3}x^2 + \frac{a+b}{2}x + \frac{a-c}{2}$$

Inverse Maps

We choose a basis (any will do) of \mathcal{P}_2 : $\mathcal{B} = (x^2, x, 1)$. Then

$$\varphi_{\mathcal{B}}(ax^2 + bx + c) = \begin{pmatrix} a \\ b \\ c \end{pmatrix}, \quad \varphi_{\mathcal{B}}(L(ax^2 + bx + c)) = \begin{pmatrix} \frac{a+b+c}{3} \\ \frac{a+b}{2} \\ \frac{a-c}{2} \end{pmatrix}$$

We can read off that

$$A = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/2 & 1/2 & 0 \\ 1/2 & 0 & -1/2 \end{pmatrix}$$

and (with a little bit of work) calculate

$$A^{-1} = \begin{pmatrix} 3 & -2 & 2 \\ -3 & 4 & -2 \\ 3 & -2 & 0 \end{pmatrix}$$

Inverse Maps

Now we are able to calculate the inverse of L :

$$L^{-1}(ax^2 + bx + c) = \varphi_{\mathcal{B}}^{-1} \circ A^{-1} \circ \varphi_{\mathcal{B}}(ax^2 + bx + c)$$

$$= \varphi_{\mathcal{B}}^{-1} \circ A^{-1} \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

$$= \varphi_{\mathcal{B}}^{-1} \begin{pmatrix} 3 & -2 & 2 \\ -3 & 4 & -2 \\ 3 & -2 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

$$= \varphi_{\mathcal{B}}^{-1} \begin{pmatrix} 3a - 2b + 2c \\ -3a + 4b - 2c \\ 3a - 2b \end{pmatrix}$$

$$= (3a - 2b + 2c)x^2 + (-3a + 4b - 2c)x + 3a - 2b$$

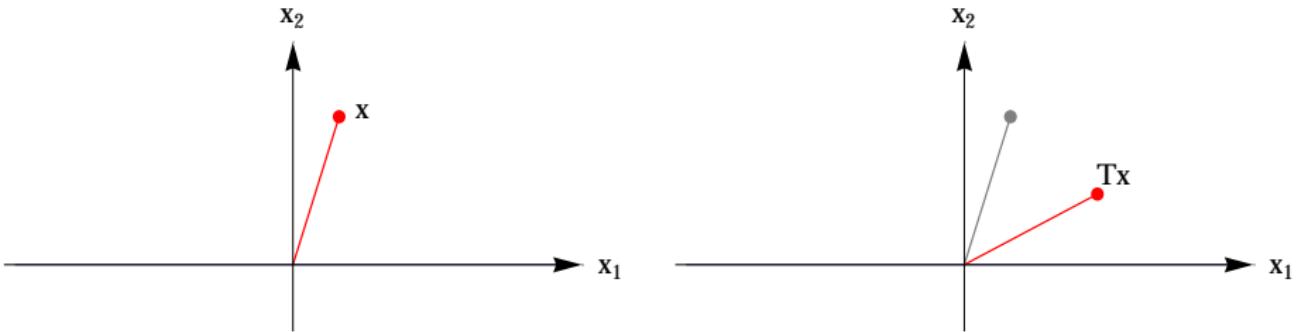
Linear Maps - Active and Passive Points of View

Consider the rotation of the plane \mathbb{R}^2 by 45° in the clockwise direction. In the **active** point of view we apply a linear map

$$T: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

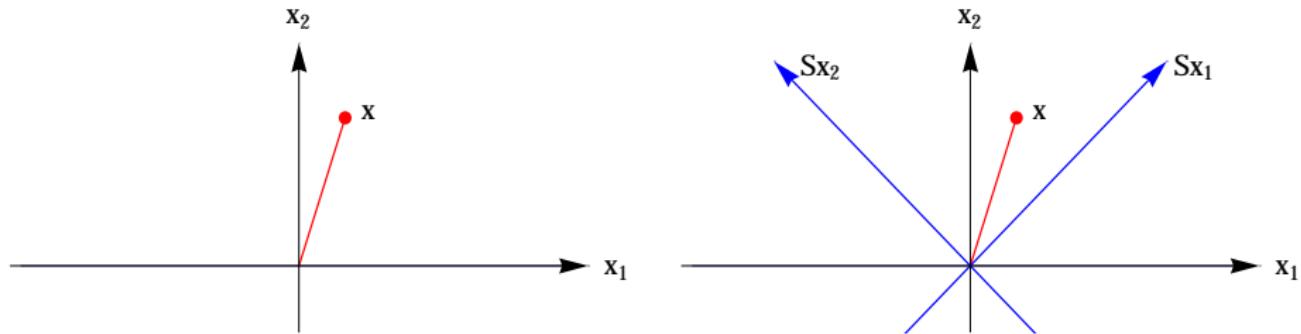
so

$$T = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$



Linear Maps - Active and Passive Points of View

In the **passive** point of view we do not apply a map T to x , but rather transform the coordinate system, i.e., the basis vectors of \mathbb{R}^2 :



The basis vectors must be transformed with the map $S = T^{-1}$. If $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ are the original standard basis vectors and $e'_1 = T^{-1}e_1$, $e'_2 = T^{-1}e_2$ are the transformed basis vectors, we have

$$x = \sum x_i e_i = \sum x_i T T^{-1} e_i = T \left(\sum x_i e'_i \right).$$

Thus, the point x remains unchanged, but with respect to the new basis (e'_1, e'_2) it appears that T has been applied.

Linear Maps - Active and Passive Points of View

Both points of view are equally valid; beginners usually prefer to work with the active point of view (we are **doing something**, e.g., rotating the vectors in the plane), while more experienced mathematicians consider the passive point of view (we are **not doing anything**, just expressing the same vectors in a different basis) more elegant and frequently more useful.

Change of Basis - Active Point of View

A special case that is of particular interest is the changing of a basis. (Of course, any isomorphism can be regarded as a basis change.) Suppose we have a basis $\mathcal{A} = (e_1, \dots, e_n)$ of \mathbb{R}^n and we wish to represent a vector x not as

$$x = \sum_{i=1}^n x_i e_i, \quad x_1, \dots, x_n \in \mathbb{R}$$

but rather as

$$x = \sum_{i=1}^n x'_i e'_i, \quad x'_1, \dots, x'_n \in \mathbb{R}$$

where $\mathcal{B} = (e'_1, \dots, e'_n)$ is another basis of \mathbb{R}^n . Our goal is to calculate the x'_i based on the x_i .

Change of Basis - Active Point of View

Assume that there exists some matrix T such that $e'_i = Te_i$, $i = 1, \dots, n$.
Then

$$T^{-1}x = \sum_{i=1}^n x'_i T^{-1}e'_i = \sum_{i=1}^n x'_i e_i$$

Thus, if we apply T^{-1} to x we obtain the coordinates x'_i of x with respect to the basis vectors $e'_i = Te_i$. Alternatively, we could say that the change of basis $e_i \mapsto Te_i$ is accomplished by applying T^{-1} to x .

Change of Basis - Passive Point of View

We can also view the passive point of view as follows: we consider the identity map $\text{id}: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, but we imbue the original space \mathbb{R}^2 with the standard basis $\mathcal{A} = (e_1, e_2)$, while the image space gets the basis $\mathcal{B} = (e'_1, e'_2) = (Te_1, Te_2)$. Then we have $\varphi_{\mathcal{A}} = \text{id}$ and $\varphi_{\mathcal{B}} = T^{-1}$ consider the matrix $A = \Phi_{\mathcal{A}}^{\mathcal{B}}(\text{id})$:

$$\begin{array}{ccc} \mathbb{R}^2 & \xrightarrow{\text{id}} & \mathbb{R}^2 \\ \varphi_{\mathcal{A}} \downarrow & & \downarrow \varphi_{\mathcal{B}} \\ \mathbb{R}^2 & \xrightarrow{A} & \mathbb{R}^2 \end{array} \quad A \circ \varphi_{\mathcal{A}} = A = \varphi_{\mathcal{B}} \circ \text{id} = \varphi_{\mathcal{B}} = T^{-1}$$

Thus, T^{-1} is just the representing matrix of the basis change.

Regardless of the point of view, we arrive at a change to a basis (Te_1, Te_2) by applying T^{-1} to x . While we have focused on \mathbb{R}^2 here, this generalizes to \mathbb{R}^n in a natural way.

Reflection in \mathbb{R}^2

1.5.17. Example. Consider the reflection of vectors in \mathbb{R}^2 by the x_1 axis, i.e., the map

$$A: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix}.$$

The matrix representation of A is simply

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{1.5.4}$$

Now we want to consider the reflection in \mathbb{R}^2 about the line through the vector

$$y = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Of course, the correct matrix for this reflection can be found geometrically. However, here we want to illustrate how a change of basis can help us determine this matrix algebraically.

Reflection in \mathbb{R}^2

Denote by L the reflection about the line through $y = b_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$. A vector perpendicular to y is $b_2 = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$ and so we can choose $\mathcal{A} = (b_1, b_2)$ as a basis. These vectors have the property that

$$Lb_1 = b_1 \quad \text{and} \quad Lb_2 = -b_2,$$

so in this basis we know how L acts. Now $b_1 = Te_1$ and $b_2 = Te_2$ where

$$T = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} \quad \text{and we note that} \quad T^{-1} = \frac{1}{5} \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}.$$

The strategy for calculating the action of the reflection L is now as follows:

1. Change to the basis (b_1, b_2) ;
2. Execute the reflection in this basis. It is given by the matrix A of (1.5.4);
3. Change back to the basis (e_1, e_2) .

Reflection in \mathbb{R}^2

Implementing these steps requires applying, in order, T^{-1} to change the basis, A for the reflection in this basis, T to change back:

$$L = TAT^{-1} = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \frac{1}{5} \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} -3 & 4 \\ 4 & 3 \end{pmatrix}$$

It is easily verified that $Lb_1 = b_1$ and $Lb_2 = -b_2$, as expected.

Systems of Linear Equations

Finite-Dimensional Vector Spaces

Inner Product Spaces

Linear Maps

Matrices

Theory of Systems of Linear Equations

Determinants

The Solution Set of Systems of Equations

We briefly return to the theory of solvability of linear systems of equations $Ax = b$. We define the **solution set**

$$\text{Sol}(A, b) = \{x \in \mathbb{R}^n : Ax = b\}.$$

If $x_0 \in \mathbb{R}^n$ satisfies

$$Ax_0 = b$$

we say that x_0 is a **particular solution** of $Ax = b$. The **associated homogeneous solution set** is

$$\text{Sol}(A, 0) = \{x \in \mathbb{R}^n : Ax = 0\} = \ker A.$$

A very important, fundamental result states:

The solution set of $Ax = b$ is the sum of the homogeneous solution set and a particular solution.

Structure of the Solution Set

1.6.1. Lemma. Let $x_0 \in \mathbb{R}^n$ be a particular solution of $Ax = b$. Then

$$\text{Sol}(A, b) = \{x_0\} + \ker A = \{y \in \mathbb{R}^n : y = x_0 + x, x \in \ker A\}.$$

where the sum of sets is understood as in Definition 1.2.24.

Proof.

(i) $\text{Sol}(A, b) \supset \{x_0\} + \ker A$: Let $x \in \ker A$. Then

$$A(x_0 + x) = Ax_0 + Ax = Ax_0 = b,$$

so $x_0 + x \in \text{Sol}(a, b)$.

(ii) $\text{Sol}(A, b) \subset \{x_0\} + \ker A$: Let $v \in \text{Sol}(A, b)$. Then

$$A(v - x_0) = Av - Ax_0 = b - b = 0,$$

so $v - x_0 \in \ker A$, implying $v \in \{x_0\} + \ker A$. □

Solvability of Systems of Equations

The following results follow immediately:

1.6.2. Corollary. If x_0 is a solution of $Ax = b$ and $\{v_1, \dots, v_r\}$ a basis of $\ker A$, then

$$\text{Sol}(A, b) = \{x \in \mathbb{R}^n : x = x_0 + \lambda_1 v_1 + \dots + \lambda_r v_r : \lambda_1, \dots, \lambda_r \in \mathbb{R}\}.$$

Here $r = \dim \ker A$.

1.6.3. Corollary. Suppose that the linear system of equations $Ax = b$ has a solution. Then the solution is unique if and only if $\ker A = \{0\}$.

Solvability of Systems of Equations

This gives rise to a further, fundamentally important result:

1.6.4. Fredholm Alternative. Let A be an $n \times n$ matrix. Then

- ▶ either $Ax = b$ has a unique solution for any $b \in \mathbb{R}^n$
- ▶ or $Ax = 0$ has a non-trivial solution.

Proof.

Either $\ker A = \{0\}$ (in which case $Ax = b$ has the solution $x = A^{-1}b$ for any $b \in \mathbb{R}^n$) or $x_0 \in \ker A$ is a non-trivial solution of $Ax = 0$. □

The Fredholm alternative occurs in many more complicated contexts. This is the most basic case.

Matrix Rank

1.6.5. Definition. Let $A \in \text{Mat}(m \times n; \mathbb{F})$ be a matrix with columns $a_j \in \mathbb{F}^m$, $1 \leq j \leq n$, and rows $a_i \in \mathbb{F}^n$, $1 \leq i \leq m$. Then we define

- ▶ the **column rank** of A to be

$$\text{column rank } A := \dim \text{span}\{a_{\cdot 1}, \dots, a_{\cdot n}\}$$

- ▶ and the **row rank** of A to be

$$\text{row rank } A := \dim \text{span}\{a_{1 \cdot}, \dots, a_{m \cdot}\}.$$

1.6.6. Remarks.

- ▶ The column rank is the greatest number of independent column vectors a_j that can be selected from all columns. This is analogously true for the row rank.
- ▶ $\text{column rank } A = \text{row rank } A^T$.
- ▶ $\text{column rank } A = \dim \text{ran } A$.

Matrix Rank

1.6.7. Definition and Theorem. Let $A \in \text{Mat}(m \times n; \mathbb{F})$. Then the column rank is equal to the row rank and we define the **rank** of A by

$$\text{rank } A := \text{column rank } A = \text{row rank } A.$$

Proof.

In the assignments it will be shown that

$$\text{ran } \bar{A}^T = (\ker A)^\perp.$$

Then, using Corollary 1.3.24 and the dimension formula (1.4.3),

$$\begin{aligned}\text{row rank } A &= \text{column rank } A^T = \text{column rank } \bar{A}^T = \dim \text{ran } \bar{A}^T \\ &= \dim(\ker A)^\perp = n - \dim \ker A = \dim \text{ran } A = \text{column rank } A.\end{aligned}$$

Here we have used that complex conjugation is a linear, bijective map $\mathbb{C} \rightarrow \mathbb{C}$ if \mathbb{C} is regarded as a real vector space. □

Existence of Solutions

The fundamental theorem on the existence of solutions to a linear system of equations is the following:

1.6.8. Theorem. There exists a solution x for $Ax = b$ if and only if $\text{rank } A = \text{rank}(A | b)$, where

$$(A | b) = \begin{pmatrix} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{pmatrix} \in \text{Mat}((n+1) \times m).$$

Solvability of Systems of Equations

Proof.

We write $A = (a_{\cdot 1}, \dots, a_{\cdot n})$, where the $a_{\cdot k} \in \mathbb{R}^m$ are column vectors of A . Then we use that the range of a matrix is the span of its column vectors and the rank is the dimension of the range, so

$$\begin{aligned} & Ax = b \text{ has solution } x \in \mathbb{R}^n \\ \Leftrightarrow & b \in \text{ran } A \\ \Leftrightarrow & b \in \text{span}\{a_{\cdot 1}, \dots, a_{\cdot n}\} \\ \Leftrightarrow & b \text{ is not independent of } a_{\cdot 1}, \dots, a_{\cdot n} \\ \Leftrightarrow & \dim \text{span}\{a_{\cdot 1}, \dots, a_{\cdot n}\} = \dim \text{span}\{a_{\cdot 1}, \dots, a_{\cdot n}, b\} \\ \Leftrightarrow & \dim \text{ran } A = \dim \text{ran}(A \mid b) \\ \Leftrightarrow & \text{rank } A = \text{rank}(A \mid b) \end{aligned}$$



Manipulating Matrices

A matrix is just a list of lists. The command **Append** is used to add elements to a list:

```
Append[{a, b, c, d}, x]
```

```
{a, b, c, d, x}
```

We want to use this to check $\text{rank } A = \text{rank}(A | b)$. Define a matrix A and a vector b as follows::

```
A = Table[ai,j, {i, 2}, {j, 3}];  
B = Table[bi, {i, 2}];  
Print["A = " MatrixForm[A], ", b = " MatrixForm[B]]
```

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

 Manipulating Matrices

Since a matrix is a list of row vectors, it is easy to add a row:

```
Append[A, {x, y, z}] // MatrixForm
```

$$\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ x & y & z \end{pmatrix}$$

To add a column, we could transpose, add a row, and transpose again:

```
Transpose[Append[Transpose[A], B]] // MatrixForm
```

$$\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & b_1 \\ a_{2,1} & a_{2,2} & a_{2,3} & b_2 \end{pmatrix}$$

However, the repeated transposition is inefficient and may cost significant computing resources for large matrices.

Manipulating Matrices

There exists a specialized command to achieve the same result without transposition:

```
MapThread[Append, {A, B}] // MatrixForm
```

$$\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & b_1 \\ a_{2,1} & a_{2,2} & a_{2,3} & b_2 \end{pmatrix}$$

The rank of a matrix is found through the **MatrixRank** command.

 Manipulating Matrices

1.6.9. Example.

```
A = {{1, 2, 3}, {4, 5, 6}, {7, 8, 9}};  
MatrixRank[A]
```

2

```
b = {3, 4, 5};  
MatrixRank[MapThread[Append, {A, b}]]
```

2

```
b = {3, 4, 6};  
MatrixRank[MapThread[Append, {A, b}]]
```

3

 Manipulating Matrices

The kernel of a matrix is obtained from the **NullSpace** command:

```
NullSpace[{{1, 2, 3}, {4, 5, 6}, {7, 8, 9}}]  
{ {1, -2, 1} }
```

The output is a list of basis vectors of the kernel of the matrix.

Systems of Linear Equations

Finite-Dimensional Vector Spaces

Inner Product Spaces

Linear Maps

Matrices

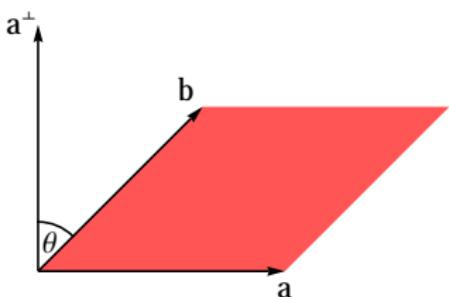
Theory of Systems of Linear Equations

Determinants

Parallelograms

We will motivate determinants geometrically (as areas of parallelograms/volumes of parallelepipeds) rather than algebraically (via solutions of systems of linear equations).

Consider a parallelogram $P(a, b)$ spanned by two non-colinear vectors $a, b \in \mathbb{R}^2$.



We are interested in the area $A(a, b)$ of the parallelogram, which is equal to the area of the rectangle with width $|a|$ and height given by $|b||\cos \theta|$. Let $a = (a_1, a_2)$, $a^\perp = (-a_2, a_1)$. Then $a \perp a^\perp$, i.e., $\langle a, a^\perp \rangle = 0$ and $|a^\perp| = |a|$.

From (1.3.3) it follows that

$$|b| \cos \theta = \left\langle \frac{a^\perp}{|a^\perp|}, b \right\rangle$$

The Determinant in \mathbb{R}^2

We obtain

$$A(a, b) = |a| \left| \left\langle \frac{a^\perp}{|a^\perp|}, b \right\rangle \right| = \frac{|a|}{|a^\perp|} |\langle a^\perp, b \rangle| = |\langle a^\perp, b \rangle| = |a_1 b_2 - a_2 b_1|.$$

We remark that

$$A(a, b) = |\langle a^\perp, b \rangle| = |a||b| \sin \sphericalangle(a, b)$$

We define the determinant as a map

$$\det: \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}, \quad \det \left(\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \right) = a_1 b_2 - a_2 b_1, \quad (1.7.1)$$

so that $A(a, b) = |\det(a, b)|$. The determinant is an **oriented area**.

Equivalently, the determinant may be regarded as a map

$$\det: \text{Mat}(2 \times 2; \mathbb{R}) \rightarrow \mathbb{R}, \quad \det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} = a_1 b_2 - a_2 b_1. \quad (1.7.2)$$

Properties of the Determinant

Both interpretations of the determinant will be used frequently.

1.7.1. Remark. We note the following properties of the determinant:

1. det is **normed**, i.e.,

$$\det(e_1, e_2) = \det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1.$$

2. det is **bilinear**:

$$\det(\lambda a, b) = \lambda \det(a, b) = \det(a, \lambda b),$$

$$\det(a + b, c) = \det(a, c) + \det(b, c),$$

$$\det(a, b + c) = \det(a, b) + \det(a, c).$$

This can be easily seen geometrically by considering the volumes of the parallelograms.

Properties of the Determinant

3. det is **alternating**, i.e., $\det(a, a) = a_1a_2 - a_2a_1 = 0$. Note that this implies that $\det(a, b) = -\det(b, a)$, since (using the bilinearity)

$$\begin{aligned} 0 &= \det(a + b, b + a) \\ &= \det(a, a) + \det(b, b) + \det(a, b) + \det(b, a) \\ &= \det(a, b) + \det(b, a). \end{aligned}$$

(In the case of two variables, an alternating map is often called **antisymmetric**.)

Vector Product in \mathbb{R}^3

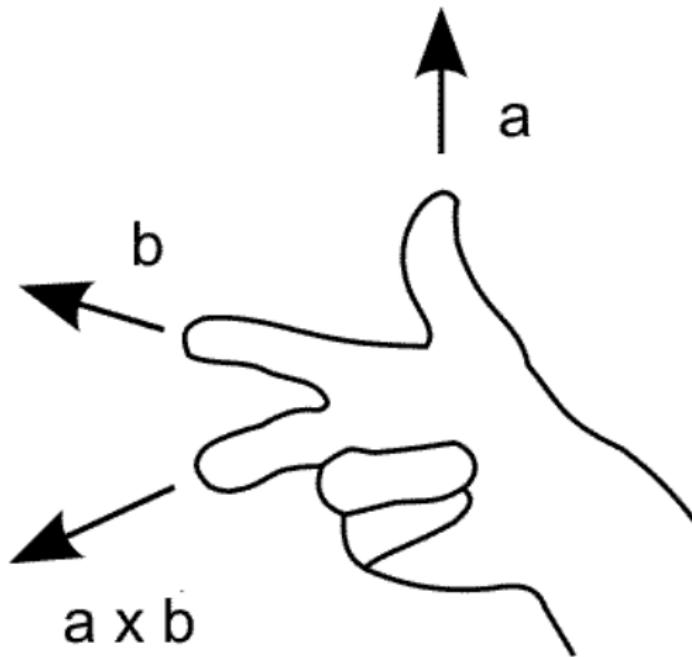
We now introduce the “vector product” $a \times b$ of two vectors $a, b \in \mathbb{R}^3$.
The vector $a \times b \in \mathbb{R}^3$ is determined by

1. its length: we set $|a \times b| = A(a, b)$, the area of the parallelogram spanned by a and b (if a and b are linearly dependent, we set $a \times b = 0$);
2. its direction: we want $a \times b$ to be orthogonal to a and b - in other words, $a \times b \perp \text{span}\{a, b\}$;
3. its orientation: $(a, b, a \times b)$ should form a “right-hand system” (defined using the thumb, index finger and middle finger of the right hand).

This is sufficient to define a unique vector $a \times b$ for $a, b \in \mathbb{R}^3$, i.e., we have a map $\times: \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$.

1.7.2. Remark. In contradistinction to the scalar product, which can be defined on \mathbb{R}^n for $n = 1, 2, \dots$, the vector product is only defined on \mathbb{R}^3 .

The Right Hand Rule



Rechte Hand Regel [modified]. Wikimedia Commons. Wikimedia Foundation. Web. 9 May 2012.

Properties of the Vector Product in \mathbb{R}^3

Note that the vector product is

1. **bilinear**: the homogeneity $(\lambda a) \times b = \lambda(a \times b) = a \times (\lambda b)$ follows from the definition of the cross product, the additivity $a \times (b + c) = a \times b + a \times c$, $(a + b) \times c = a \times c + b \times c$ is easy to see geometrically when a, b, c are coplanar and slightly more difficult to show when they are not.
2. **antisymmetric**: $a \times a = 0$, or $a \times b = -b \times a$.

We can compute the vector product of the standard basis vectors with each other:

$$e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$e_1 \times e_2 = e_3 = -e_2 \times e_1$$

$$e_2 \times e_3 = e_1 = -e_3 \times e_2$$

$$e_3 \times e_1 = e_2 = -e_1 \times e_3$$

$$e_1 \times e_1 = 0 = e_2 \times e_2 = e_3 \times e_3$$
(1.7.3)

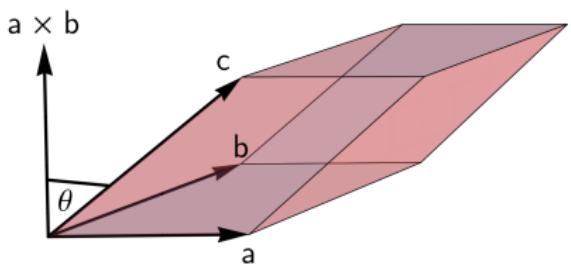
Calculating the Vector Product in \mathbb{R}^3

Using the bilinearity and (1.7.3), we can now calculate $a \times b$ for arbitrary $a, b \in \mathbb{R}^3$:

$$\begin{aligned} a \times b &= (a_1 e_1 + a_2 e_2 + a_3 e_3) \times (b_1 e_1 + b_2 e_2 + b_3 e_3) \\ &= \sum_{i,j=1}^3 a_i b_j (e_i \times e_j) \\ &= (a_2 b_3 - a_3 b_2) e_1 + (a_3 b_1 - a_1 b_3) e_2 + (a_1 b_2 - a_2 b_1) e_3 \\ &= \begin{pmatrix} + \det \begin{pmatrix} a_2 & b_2 \\ a_3 & b_3 \end{pmatrix} \\ - \det \begin{pmatrix} a_1 & b_1 \\ a_3 & b_3 \end{pmatrix} \\ + \det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \end{pmatrix} \end{aligned} \tag{1.7.4}$$

Parallel Epipeds

We now consider the problem of finding the volume of a parallel epiped spanned by three vectors $a, b, c \in \mathbb{R}^3$.



The volume is given by the base area (the area of the parallelogram spanned by a, b) multiplied with the height, $|c||\cos \theta|$. Using the fact that $a \times b$ is orthogonal to a and b we have

$$|c| \cos \theta = \left\langle \frac{a \times b}{|a \times b|}, c \right\rangle,$$

so the volume is given by

$$V(a, b, c) = |a \times b| \left| \left\langle \frac{a \times b}{|a \times b|}, c \right\rangle \right| = |\langle a \times b, c \rangle|.$$

The Determinant in \mathbb{R}^3

We therefore define the determinant as an ***oriented volume***,

$$\det: \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}, \quad \det(a, b, c) = \langle a \times b, c \rangle. \quad (1.7.5)$$

(Again, we note that it can be equivalently defined $\text{Mat}(3 \times 3; \mathbb{R}) \rightarrow \mathbb{R}$.)

Note that

$$\det(a, b, c) > 0 \quad \text{if } (a, b, c) \text{ form a right-hand system,} \quad (1.7.6)$$

$$\det(a, b, c) < 0 \quad \text{if } (a, b, c) \text{ form a left-hand system,} \quad (1.7.7)$$

$$\det(a, b, c) = 0 \quad \text{if } a = \lambda b \text{ or } a = \lambda c \text{ or } b = \lambda c \text{ for any } \lambda \in \mathbb{R}. \quad (1.7.8)$$

The last property follows from the properties of the vector and scalar products: if $a = \lambda b$ then $a \times b = 0$, and since $a \times b$ is orthogonal both a and b , the scalar product $\langle a \times b, c \rangle$ will vanish if $a = \lambda c$ or $b = \lambda c$.

Cyclic Permutations

Let (x_1, \dots, x_n) be an ordered list of elements. Define a relation \prec by $x_1 \prec x_2 \prec x_3 \prec \cdots \prec x_{n-1} \prec x_n \prec x_1$ (" x_1 precedes x_2 precedes x_3 etc."). Let $\pi: \{x_1, \dots, x_n\} \rightarrow \{x_1, \dots, x_n\}$ be a bijective map (such a map is called a **permutation**). Then any list $(\pi(x_1), \dots, \pi(x_n))$ is called a cyclic permutation of (x_1, \dots, x_n) if $\pi(x_1) \prec \pi(x_2) \prec \cdots \prec \pi(x_n) \prec \pi(x_1)$.

Furthermore, if (a, b, c) form a right-hand system, then $V(a, b, c) = \det(a, b, c)$. Since the volume is independent of the designation of the vectors, we observe that a **cyclic permutation** of (a, b, c) preserves the right-handedness and

$$\det(a, b, c) = \det(c, a, b) = \det(b, c, a)$$

or

$$\langle a \times b, c \rangle = \langle c \times a, b \rangle = \langle b \times c, a \rangle.$$

Calculating Determinants in \mathbb{R}^3

Note that by (1.7.4),

$$\begin{aligned}
 \det(a, b, c) &= \langle b \times c, a \rangle = \sum_{i=1}^3 a_i(b \times c)_i \\
 &= a_1 \det \begin{pmatrix} b_2 & c_2 \\ b_3 & c_3 \end{pmatrix} - a_2 \det \begin{pmatrix} b_1 & c_1 \\ b_3 & c_3 \end{pmatrix} + a_3 \det \begin{pmatrix} b_1 & c_1 \\ b_2 & c_2 \end{pmatrix} \\
 &= \det \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}
 \end{aligned} \tag{1.7.9}$$

We may therefore calculate a 3×3 determinant $\det A$ by calculating 2×2 **subdeterminants**. Denote by A_{kj} the 2×2 matrix obtained from A by deleting the k th row and the j th column, (1.7.9) can be written as

$$\det A = \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \sum_{k=1}^3 (-1)^{k+1} a_{k1} \det A_{k1} \tag{1.7.10}$$

Calculating Determinants in \mathbb{R}^3

We will prove later that in fact

$$\det \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} = \det \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}.$$

This (together with (1.7.9)) motivates the mnemonic

$$a \times b = \det \begin{pmatrix} e_1 & e_2 & e_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix},$$

where e_1, e_2, e_3 are the standard unit basis vectors and $a = (a_1, a_2, a_3)$, $b = (b_1, b_2, b_3)$.

Properties of the Determinant in \mathbb{R}^3

We once more note the following properties of the determinant:

1. det is **normed**, i.e.,

$$\det(e_1, e_2, e_3) = \det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 1.$$

2. det is **trilinear**:

$$\det(\lambda a, b, c) = \lambda \det(a, b, c) = \det(a, \lambda b, c) = \det(a, b, \lambda c),$$

$$\det(a + b, c, d) = \det(a, c, d) + \det(b, c, d),$$

$$\det(a, b + c, d) = \det(a, b, d) + \det(a, c, d),$$

$$\det(a, b, c + d) = \det(a, b, c) + \det(a, b, d)$$

These properties follow from the corresponding properties of the scalar and vector products.

3. det is **alternating** (see (1.7.8)).

Preview of Determinants in \mathbb{R}^n

Our goal is now to find a generalization of the determinant that has the three main properties of being

- ▶ multilinear,
- ▶ alternating and
- ▶ normed.

It will turn out that these three properties are actually sufficient to define the determinant uniquely; there is only one map $\text{Mat}(n \times n; \mathbb{R}) \rightarrow \mathbb{R}$ with these properties, and for $n = 2, 3$ it is given by (1.7.2) and (1.7.5), respectively.

In the case of $n = 1$, $\text{Mat}(1 \times 1; \mathbb{R})$ is equivalent to \mathbb{R} and we define $\det(a) = a$. This definition trivially has the properties of being normed and linear; while it makes no sense to define what alternating means.

Preview of Determinants in \mathbb{R}^n

We will further see that one possible formula for determinants

$\det: \text{Mat}(n \times n) \rightarrow \mathbb{R}$ can be constructed recursively, similar to (1.7.10).

In fact, if $A \in \text{Mat}(n \times n; \mathbb{R})$, we define $A_{kj} \in \text{Mat}((n - 1) \times (n - 1); \mathbb{R})$ as the matrix obtained from A by deleting the k th row and the j th column.

Then for any $j = 1, \dots, n$ we will obtain the recursion formula

$$\det A = \sum_{k=1}^n (-1)^{k+j} a_{kj} \det A_{kj} \quad (1.7.11)$$

In order to understand the extension of determinants to \mathbb{R}^n better, we need to formalize the concept of permutations.

Groups

1.7.3. Definition. A **group** is a pair (G, \circ) consisting of a set G and a **group operation** $\circ: G \times G \rightarrow G$ such that

1. $a \circ (b \circ c) = (a \circ b) \circ c$ for all $a, b, c \in G$ (associativity),
2. there exists an element $e \in G$ such that $a \circ e = e \circ a = a$ for all $a \in G$ (existence of a unit element),
3. for every $a \in G$ there exists an element $a^{-1} \in G$ such that $a \circ a^{-1} = a^{-1} \circ a = e$ (existence of an inverse).

A group is called commutative if in addition to the above properties

4. $a \circ b = b \circ a$ for all $a, b \in G$ (commutativity).

Groups and Permutations

1.7.4. Examples.

1. Any vector space $(V, +, \cdot)$ may be regarded as a commutative group $(V, +)$ with the additional operation of scalar multiplication.
2. The set of invertible matrices,

$$\mathrm{GL}(n; \mathbb{R}) := \{A \in \mathrm{Mat}(n \times n; \mathbb{R}): A \text{ is invertible}\}$$

is a group with the group operation given by matrix multiplication (composition of maps).

1.7.5. Definition. The set of all *permutations of n elements*

$$S_n = \{\pi: \{x_1, \dots, x_n\} \rightarrow \{x_1, \dots, x_n\}: \pi \text{ bijective}\}$$

together with the group operation “composition of maps”,
 $\pi_1 \circ \pi_2(x) = \pi_1(\pi_2(x))$ is called the *symmetric group*.

Permutations

It is easy to check that (S_n, \circ) in fact has properties i) – iii), but not property iv). We will often denote a group by G instead of (G, \circ) if no confusion arises therefrom.

A permutation of n elements is a **finite** map; recall that a function f is defined by pairs of the form $(x, f(x))$, where x is the independent variable. A permutation is defined on a set of n elements; instead of $\{x_1, \dots, x_n\}$ we can also simply write $\{1, \dots, n\}$, replacing the permutation of elements with a permutation of indices. Then we might define a permutation π through a set of pairs $\{(1, \pi(1)), \dots, (n, \pi(n))\}$. In fact, we do represent permutations in this way, but us a different notation, writing

$$\pi = \begin{pmatrix} 1 & 2 & \dots & n \\ \pi(1) & \pi(2) & \dots & \pi(n) \end{pmatrix}$$

Transpositions

For example, if $n = 2$, there are only two permutations $\pi_1, \pi_2 \in S_2$,

$$\begin{aligned}\pi_1: 1 &\mapsto 1, & 2 &\mapsto 2, & \pi_1 &= \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \\ \pi_2: 1 &\mapsto 2, & 2 &\mapsto 1, & \pi_2 &= \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}. \end{aligned}\tag{1.7.12}$$

1.7.6. Definition. A permutation in S_n that leaves exactly $n - 2$ elements invariant is called a **transposition**.

A transposition $\tau \in S_n$ has the form

$$\tau(k) = \begin{cases} i & \text{if } k = j \\ j & \text{if } k = i \\ k & \text{otherwise} \end{cases}\tag{1.7.13}$$

for some $i, j \in \{1, \dots, n\}$.

Permutations as Transpositions

1.7.7. Lemma. Every permutation $\pi \in S_n$, $n \geq 2$, is a composition of transpositions, $\pi = \tau_1 \circ \cdots \circ \tau_k$.

Note that the transpositions τ_j and the number k are **not** uniquely defined.

Proof.

We proceed by induction. For $n = 2$ there are only two permutations π_1 and π_2 (see (1.7.12)); π_2 is a transposition, and $\pi_1 = \pi_2 \circ \pi_2$. We now assume that any permutation in S_n can be written as a composition of transpositions and prove that this is also true for any permutation in S_{n+1} .

Let $\pi \in S_n$. Then we can consider

$$\tilde{\pi} = \begin{pmatrix} 1 & \dots & n & n+1 \\ \pi(1) & \dots & \pi(n) & n+1 \end{pmatrix} \quad (1.7.14)$$

as an element of S_{n+1} . Also, every element $\tilde{\pi} \in S_{n+1}$ of the form (1.7.14) can be regarded as an element $\pi \in S_n$.

Permutations as Transpositions

Proof (continued).

Now let $\sigma \in S_{n+1}$ and let τ be the transposition that exchanges $n+1$ and $\sigma^{-1}(n+1)$. Then

$$\sigma \circ \tau : n+1 \xrightarrow{\tau} \sigma^{-1}(n+1) \xrightarrow{\sigma} n+1$$

so that

$$\sigma \circ \tau = \begin{pmatrix} 1 & \dots & n & n+1 \\ \pi(1) & \dots & \pi(n) & n+1 \end{pmatrix}$$

for some values $\pi(1), \dots, \pi(n)$, $\pi \in S_n$. It follows that $\sigma \circ \tau$ can be written as a composition of transpositions $\tau_1 \circ \dots \circ \tau_k$,

$$\sigma \circ \tau = \tau_1 \circ \dots \circ \tau_k,$$

so $\sigma = \tau_1 \circ \dots \circ \tau_k \circ \tau^{-1}$, which proves the assertion. □

Sign of a Permutation

While the number of transpositions that make up a permutation is not unique, we do have the following:

1.7.8. Definition and Theorem. Let $\pi \in S_n$ be represented as a composition of k transpositions, $\pi = \tau_1 \circ \cdots \circ \tau_k$. Then the **sign** of π ,

$$\operatorname{sgn} \pi := (-1)^k$$

does not depend on the representation chosen.

In order to prove this, we need an additional concept from group theory, which we introduce on the following slide.

In advance, we note that the sign is “well-behaved”:

$$\operatorname{sgn}(\pi_1 \circ \pi_2) = \operatorname{sgn} \pi_1 \operatorname{sgn} \pi_2$$

for any $\pi_1, \pi_2 \in S_n$.

Group Actions

1.7.9. Definition. Let (G, \circ) be a group and X a set. Then an ***action (or operation) of G on X from the left*** is a map

$$\Phi: G \times X \rightarrow X \quad (g, x) \mapsto \Phi(g, x) = \Phi_g x = gx$$

with the properties

1. $ex = x$ ($e \in G$ is the unit element),
2. $(a \circ b)x = a(bx)$ for $a, b \in G, x \in X$.

We say that G acts (operates) on X .

1.7.10. Proposition. Let X be the set of all maps $f: \mathbb{R}^n \rightarrow \mathbb{R}$. Then S_n acts on X via

$$(\pi f)(x_1, \dots, x_n) = f(x_{\pi(1)}, \dots, x_{\pi(n)}), \quad \pi \in S_n.$$

Group Actions

Proof.

We need to show the properties i) and ii) of Definition 1.7.9. The unit element of S_n is

$$\pi_e = \begin{pmatrix} 1 & \dots & n \\ 1 & \dots & n \end{pmatrix},$$

so trivially $\pi_e f = f$, since

$$(\pi_e f)(x_1, \dots, x_n) = f(x_{\pi_e(1)}, \dots, x_{\pi_e(n)}) = f(x_1, \dots, x_n).$$

Furthermore, let $\sigma, \pi \in S_n$. Then

$$\begin{aligned} [\sigma(\pi f)](x) &= (\sigma f)(x_{\pi(1)}, \dots, x_{\pi(n)}) = f(x_{\sigma(\pi(1))}, \dots, x_{\sigma(\pi(n))}) \\ &= f(x_{(\sigma \circ \pi)(1)}, \dots, x_{(\sigma \circ \pi)(n)}) = [(\sigma \circ \pi)f](x_1, \dots, x_n), \end{aligned}$$

so $\sigma(\pi f) = (\sigma \circ \pi)f$.

□

Group Actions

1.7.11. Lemma. Denote by $\Delta: \mathbb{R}^n \rightarrow \mathbb{R}$ the function

$$\Delta(x_1, \dots, x_n) = \prod_{i < j} (x_j - x_i). \quad (1.7.15)$$

Then

$$\tau\Delta = -\Delta \quad \text{for any transposition } \tau \in S_n.$$

Proof.

Let $r, s \in \{1, \dots, n\}$, $r < s$, and τ the transposition exchanging r and s ,

$$\tau = \begin{pmatrix} 1 & \dots & r-1 & r & r+1 & \dots & s-1 & s & s+1 & \dots & n \\ 1 & \dots & r-1 & s & r+1 & \dots & s-1 & r & s+1 & \dots & n \end{pmatrix}.$$

Group Actions

Proof (continued).

Note that

$$\tau\Delta(x_1, \dots, x_n) = \prod_{i < j} \tau(x_j - x_i).$$

Then

$$\tau(x_r - x_s) = -(x_r - x_s). \quad (1.7.16)$$

All other factors in (1.7.15) either do not contain x_r or x_s (and are left unchanged by τ) or occur in one of the following pairings:

- ▶ $j < r$: $(x_r - x_j)(x_s - x_j)$
- ▶ $r < j < s$: $(x_s - x_j)(x_j - x_r)$
- ▶ $s < j$: $(x_j - x_s)(x_j - x_r)$

Each of these pairs is left invariant by τ , so the sign change in (1.7.16) is the only effect of τ on Δ . □

Sign of a Permutation

1.7.12. Corollary. For every permutation $\pi = \tau_1 \circ \cdots \circ \tau_k \in S_n$,

$$\pi\Delta = (\tau_1 \circ \cdots \circ \tau_k)\Delta = (-1)^k \Delta.$$

In particular,

$$\operatorname{sgn} \pi = (-1)^k,$$

does not depend on the decomposition of π into transpositions and is therefore well-defined.

Proof.

Let $\pi \in S_n$ and assume that there are transpositions $\tau_1, \dots, \tau_k, \tilde{\tau}_1, \dots, \tilde{\tau}_l$ such that

$$\pi = \tau_1 \circ \cdots \circ \tau_k = \tilde{\tau}_1 \circ \cdots \circ \tilde{\tau}_l.$$

Then $\pi\Delta(x_1, \dots, x_n) = (-1)^k \Delta(x_1, \dots, x_n) = (-1)^l \Delta(x_1, \dots, x_n)$. Choosing some x_1, \dots, x_n such that $\Delta(x_1, \dots, x_n) \neq 0$ we obtain $(-1)^k = (-1)^l$. \square

p -Multilinear Maps

1.7.13. Definition. A function $f: \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{p \text{ times}} \rightarrow \mathbb{R}$ is said to be a **p -multilinear map** (or **p -multilinear form**) if f is linear in each entry, i.e.,

$$f(\lambda a_1, a_2, \dots, a_p) = \lambda f(a_1, a_2, \dots, a_p)$$

and

$$f(a_1 + b, a_2, \dots, a_p) = f(a_1, a_2, \dots, a_p) + f(b, a_2, \dots, a_p)$$

for $b, a_1, \dots, a_p \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$ and analogous equations hold for the other entries.

The form is said to be **alternating** if $f(a_1, \dots, a_p) = 0$ whenever $a_j = a_k$ for any $j \neq k$.

An n -multilinear form is said to be **normed** if $f(e_1, \dots, e_n) = 1$, where e_1, \dots, e_n are the standard basis vectors in \mathbb{R}^n .

Characterization of Alternating Forms

We will prove that the properties of being multilinear, alternating and normed are sufficient to uniquely define the determinant in \mathbb{R}^n . First, however, we give a useful result:

1.7.14. Lemma. Let $f: \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{p \text{ times}} \rightarrow \mathbb{R}$ be a p -multilinear map. Then

the following are equivalent:

- (i) f is alternating
- (ii) $f(a_1, \dots, a_{j-1}, a_j, a_{j+1}, \dots, a_{k-1}, a_k, a_{k+1}, \dots, a_p) = -f(a_1, \dots, a_{j-1}, a_k, a_{j+1}, \dots, a_{k-1}, a_j, a_{k+1}, \dots, a_p)$
- (iii) $f(a_1, \dots, a_p) = 0$ if a_1, \dots, a_p are linearly dependent.

The proof is not difficult and left as an exercise!

Determinants in \mathbb{R}^n

We will now define the determinant as an alternating, normed, n -multilinear function for column vectors in \mathbb{R}^n and corresponding square matrices whose columns consist of these vectors, using the notation

$$a_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}, \quad (j = 1, \dots, n), \quad A = (a_1, \dots, a_n) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

1.7.15. Theorem. For every $n \in \mathbb{N}$, $n > 1$, there exists a unique, normed, alternating n -multilinear form $\det: \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{n \text{ times}} \cong \text{Mat}(n \times n; \mathbb{R}) \rightarrow \mathbb{R}$.

Furthermore,

$$\det(a_1, \dots, a_n) = \det A = \sum_{\pi \in S_n} \text{sgn } \pi \ a_{\pi(1)1} \cdots a_{\pi(n)n}. \quad (1.7.17)$$

Determinants in \mathbb{R}^n

Proof.

We will first show that the determinant defined in (1.7.17) in fact has the required properties.

1. (det is multilinear) Let $a_1, \dots, a_n, b \in \mathbb{R}^n$. Then we show the additivity in the first entry (the proof for all other entries is completely analogous)

$$\begin{aligned}& \det(a_1 + b, a_2, \dots, a_n) \\&= \sum_{\pi \in S_n} \operatorname{sgn} \pi (a_{\pi(1)1} + b_{\pi(1)}) a_{\pi(2)2} \cdots a_{\pi(n)n} \\&= \sum_{\pi \in S_n} \operatorname{sgn} \pi a_{\pi(1)1} \cdots a_{\pi(n)n} + \sum_{\pi \in S_n} \operatorname{sgn} \pi b_{\pi(1)} \cdots a_{\pi(n)n} \\&= \det(a_1, a_2, \dots, a_n) + \det(b, a_2, \dots, a_n)\end{aligned}$$

The homogeneity is shown analogously

Determinants in \mathbb{R}^n

Proof (continued).

2. (\det is normed) Let

$$e_j = \begin{pmatrix} \delta_{1j} \\ \vdots \\ \delta_{nj} \end{pmatrix} \quad (j = 1, \dots, n), \quad \delta_{ij} = \begin{cases} 1 & i = j, \\ 0 & i \neq j. \end{cases}$$

Then for any permutation $\pi \in S_n$,

$$\delta_{1\pi(1)} \cdots \delta_{n\pi(n)} = \begin{cases} 1 & \pi(k) = k, \quad k = 1, \dots, n, \\ 0 & \text{otherwise.} \end{cases}$$

Determinants in \mathbb{R}^n

Proof (continued).

2. It follows that in the summation of the permutations only the summand with

$$\pi = \begin{pmatrix} 1 & 2 & \dots & n-1 & n \\ 1 & 2 & \dots & n-1 & n \end{pmatrix}, \quad \operatorname{sgn} \pi = 1,$$

survives. Thus

$$\det(e_1, \dots, e_n) = \sum_{\pi \in S_n} \operatorname{sgn} \pi \delta_{\pi(1)1} \cdots \delta_{\pi(n)n} = 1.$$

Determinants in \mathbb{R}^n

Proof (continued).

3. (\det is alternating) We will show that

$\det(a_1, a_2, \dots, a_{n-1}, a_n) = -\det(a_n, a_2, \dots, a_{n-1}, a_1)$ (again, the proof is similar when any other entries are exchanged). Let

$$\tau = \begin{pmatrix} 1 & 2 & \dots & n-1 & n \\ n & 2 & \dots & n-1 & 1 \end{pmatrix} \in S_n. \quad (1.7.18)$$

be the transposition exchanging 1 and n . We will use that $\text{sgn } \tau = -1$ and that summing over all permutations $\pi \in S_n$ is the same as summing over all $\tau \circ \pi \in S_n$, when τ is fixed by (1.7.18).

Determinants in \mathbb{R}^n

Proof (continued).

3. Then

$$\begin{aligned}& \det(a_n, a_2, \dots, a_{n-1}, a_1) \\&= \sum_{\pi \in S_n} \operatorname{sgn} \pi \ a_{\pi(1)n} a_{\pi(2)2} \cdots a_{\pi(n-1)(n-1)} a_{\pi(n)1} \\&= \sum_{\pi \in S_n} \operatorname{sgn} \pi \ a_{\pi(n)1} a_{\pi(2)2} \cdots a_{\pi(n-1)(n-1)} a_{\pi(1)n} \\&= - \sum_{\tau \circ \pi \in S_n} \operatorname{sgn}(\tau \circ \pi) \ a_{\tau \circ \pi(1)1} a_{\tau \circ \pi(2)2} \cdots a_{\tau \circ \pi(n-1)(n-1)} a_{\tau \circ \pi(n)n} \\&= - \det(a_1, a_2, \dots, a_{n-1}, a_n).\end{aligned}$$

Determinants in \mathbb{R}^n

Proof (continued).

We next show that the properties of the determinant imply the formula (1.7.17). By multilinearity we have

$$\begin{aligned}\det(a_1, \dots, a_n) &= \det\left(\sum_{j_1=1}^n a_{j_1 1} e_{j_1}, \dots, \sum_{j_n=1}^n a_{j_n n} e_{j_n}\right) \\ &= \sum_{j_1, \dots, j_n=1}^n a_{j_1 1} \cdots a_{j_n n} \det(e_{j_1}, \dots, e_{j_n})\end{aligned}$$

Since \det is supposed to be alternating, all summands vanish where any j_k occurs more than once. We therefore sum only over permutations of $\{1, \dots, n\}$,

$$\det(a_1, \dots, a_n) = \sum_{\pi \in S_n} a_{\pi(1)1} \cdots a_{\pi(n)n} \det(e_{\pi(1)}, \dots, e_{\pi(n)})$$

Determinants in \mathbb{R}^n

Proof (continued).

Again, because \det is alternating and assuming each π is composed of k transpositions,

$$\det(e_{\pi(1)}, \dots, e_{\pi(n)}) = (-1)^k \det(e_1, \dots, e_n) = \operatorname{sgn} \pi \det(e_1, \dots, e_n).$$

Since \det is normed, $\det(e_1, \dots, e_n) = 1$, so we finally have

$$\det(a_1, \dots, a_n) = \sum_{\pi \in S_n} a_{\pi(1)1} \cdots a_{\pi(n)n} \operatorname{sgn} \pi.$$

□

Determinants and Elementary Column Operations

Since the determinant is alternating and multilinear, we see that the Elementary Column Operations 1.5.8 affect the determinant as follows:

- ▶ The determinant of a matrix A changes sign if two columns of A are interchanged, e.g.,

$$\det(a_2, a_1, \dots, a_n) = -\det(a_1, a_2, \dots, a_n)$$

- ▶ Multiplying all the entries in a column with a number λ leads to the determinant being multiplied by this constant:

$$\det(a_1, \dots, \lambda a_j, \dots, a_n) = \lambda \det(a_1, \dots, a_j, \dots, \dots, a_n)$$

- ▶ Adding a multiple of a column to another column does not change the value of the determinant:

$$\det(a_1, \dots, a_j, \dots, a_k + \lambda a_j, \dots, a_n) = \det(a_1, \dots, a_j, \dots, a_k, \dots, a_n)$$

Determinants of Transposed Matrices

1.7.16. Lemma. Let $A \in \text{Mat}(n \times n; \mathbb{R})$. Then

$$\det A = \det A^T.$$

Proof.

We first note that for every $\pi \in S_n$, $\text{sgn } \pi = \text{sgn } \pi^{-1}$ and the sum over all π is equal to the sum over all π^{-1} . Then we can reorder the terms in each summand, so that

$$\begin{aligned}\det A &= \sum_{\pi \in S_n} \text{sgn } \pi \ a_{\pi(1)1} \cdots a_{\pi(n)n} = \sum_{\pi \in S_n} \text{sgn } \pi \ a_{1\pi^{-1}(1)} \cdots a_{n\pi^{-1}(n)} \\ &= \sum_{\pi^{-1} \in S_n} \text{sgn } \pi^{-1} \ a_{1\pi^{-1}(1)} \cdots a_{n\pi^{-1}(n)} = \sum_{\pi \in S_n} \text{sgn } \pi \ a_{1\pi(1)} \cdots a_{n\pi(n)} \\ &= \det A^T.\end{aligned}$$



Determinants and Elementary Row Operations

As a corollary, we can rewrite (1.7.17) in a more commonly seen form:

1.7.17. Leibnitz Formula.

$$\det A = \sum_{\pi \in S_n} \operatorname{sgn} \pi \ a_{1\pi(1)} \cdots a_{n\pi(n)} \quad (1.7.19)$$

1.7.18. Corollary. Elementary row manipulations of a matrix A affect the determinant of A in the same way as the corresponding elementary column manipulations.

Proof.

$$\begin{array}{ccc} \det A & \xrightarrow{\text{row manipulation}} & \det B \\ \parallel & & \parallel \\ \det A^T & \xrightarrow{\text{column manipulation}} & \det B^T \end{array}$$



Triangular Determinants

1.7.19. Proposition. Let $A \in \text{Mat}(n \times n)$ have upper triangular form, i.e.,

$$A = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

for diagonal elements $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ and arbitrary values (denoted by *) above the diagonal. Then

$$\det A = \lambda_1 \cdots \lambda_n.$$

Triangular Determinants

Proof.

By multilinearity,

$$\det A = \left(\prod_{i=k}^n \lambda_k \right) \det \begin{pmatrix} 1 & & * \\ & \ddots & \\ 0 & & 1 \end{pmatrix}.$$

The matrix in the determinant on the right can be transformed into the unit matrix through elementary row manipulations that do not change the value of the determinant. Therefore its determinant is 1, proving the result. □

Proposition 1.7.19 can be applied to calculate determinants of matrices $A \in \text{Mat}(n \times n)$ when A is first transformed to upper triangular form using elementary matrix manipulations. This is of practical use for $n \geq 4$.

Determinants and Invertibility of Matrices

The following result is of fundamental importance for many applications:

1.7.20. Proposition. A matrix $A \in \text{Mat}(n \times n)$ is invertible if and only if $\det A \neq 0$.

Proof.

We first show that if A is not invertible, then $\det A = 0$. The linear map $A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is invertible if and only if $\text{ran } A = \mathbb{R}^n$. Since $\text{ran } A$ is the span of the column vectors, A is invertible if and only if the column vectors are independent. But if the column vectors are not independent, then $\det A$ vanishes.

Now let $A = (a_1, \dots, a_n)$ be invertible. By Lemma 1.5.13 A can be transformed into the unit matrix by elementary row operations. These only change the value of the determinant by a non-zero factor. Since the determinant of the unit matrix is 1, it follows that $\det A \neq 0$. □

Determinants and Systems of Equations

The determinant can be used to give another formulation of Fredholm's Alternative 1.6.4:

1.7.21. **Fredholm Alternative.** Let $A \in \text{Mat}(n \times n)$. Then either

- ▶ $\det A = 0$, in which case $Ax = 0$ has a non-zero solution $x \in \ker A$, or
- ▶ $\det A \neq 0$, then $Ax = b$ has a unique solution $x = A^{-1}b$ for any $b \in \mathbb{R}^n$.

The proof is a straightforward application of the definitions and left to the reader.

1.7.22. **Cramer's Rule.** Let $A = (a_1, \dots, a_n) \in \text{Mat}(n \times n)$, $a_1, \dots, a_n \in \mathbb{R}^n$, be invertible. Then the system $Ax = b$, $b \in \mathbb{R}^n$, has the solution

$$x_i = \frac{1}{\det A} \det(a_1, \dots, a_{i-1}, b, a_{i+1}, \dots, a_n), \quad i = 1, \dots, n. \quad (1.7.20)$$

Determinants and Systems of Equations

Proof.

We note that $Ax = \sum_{k=1}^n x_k a_k$ for $A = (a_1, \dots, a_n) \in \text{Mat}(n \times n)$. Therefore,

$$\begin{aligned}& \det(a_1, \dots, a_{i-1}, b, a_{i+1}, \dots, a_n) \\&= \det(a_1, \dots, a_{i-1}, Ax, a_{i+1}, \dots, a_n) \\&= \det\left(a_1, \dots, a_{i-1}, \sum_{k=1}^n x_k a_k, a_{i+1}, \dots, a_n\right) \\&= \sum_{k=1}^n x_k \det\left(a_1, \dots, a_{i-1}, a_k, a_{i+1}, \dots, a_n\right) \\&= x_i \det(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_n) + 0 \\&= x_i \det A.\end{aligned}$$



Minors and Cofactors

1.7.23. Definition. Let $A = (a_{ij}) \in \text{Mat}(n \times n)$. Denote the $(n - 1) \times (n - 1)$ matrix obtained from A by deleting the i th row and j th column by

$$A_{ij} = (a_{kl})_{\substack{1 \leq k, l \leq n \\ k \neq i \\ l \neq j}}$$

Then

$$m_{ij} := \det A_{ij}$$

is called the (i, j) th minor of A . The number

$$c_{ij} := (-1)^{i+j} m_{ij} = (-1)^{i+j} \det A_{ij}$$

is called the (i, j) th cofactor of A and the matrix

$$\text{Cof } A := (c_{ij})_{1 \leq i, j \leq n}$$

is called the cofactor matrix of A .

Determinants and Inversion of Matrices

1.7.24. Definition. Let $A = (a_{ij}) \in \text{Mat}(n \times n)$. The transpose of the cofactor matrix of A is called the **adjugate** of A , denoted by

$$A^\sharp := (\text{Cof } A)^T$$

1.7.25. Theorem. Let $A = (a_{ij}) \in \text{Mat}(n \times n)$ be invertible. Then

$$A^{-1} = \frac{1}{\det A} A^\sharp$$

The proof is based on a useful lemma, which we first establish.

Determinants and Inversion of Matrices

1.7.26. Lemma. Let $A = (a_1, \dots, a_n) \in \text{Mat}(n \times n)$ and e_i be the i th standard basis vector in \mathbb{R}^n . Then

$$\det(a_1, \dots, a_{j-1}, e_i, a_{j+1}, \dots, a_n) = (-1)^{i+j} \det A_{ij} = c_{ij}$$

where c_{ij} is the (i, j) th cofactor of A .

Proof.

Since the determinant is multilinear, we have

$$\begin{aligned} & \det(a_1, \dots, a_{j-1}, e_i, a_{j+1}, \dots, a_n) \\ &= -\det(a_1, \dots, a_{j-1}, a_{j+1}, e_i, a_{j+2}, \dots, a_n) \\ &= (-1)^{n-j-1} \det(a_1, \dots, a_{j-1}, a_{j+1}, \dots, a_n, e_i). \end{aligned}$$

Determinants and Inversion of Matrices

Proof (continued).

Swapping i th and the $(i + 1)$ st row, etc., we obtain

$$\begin{aligned}\det(a_1, \dots, a_{j-1}, e_i, a_{j+1}, \dots, a_n) &= (-1)^{n-j-1+n-i-1} \det \begin{pmatrix} A_{ij} & 0 \\ * & 1 \end{pmatrix} \\ &= (-1)^{i+j} \det B.\end{aligned}$$

where the entries in $*$ represent the elements of the i th row of A (with the j th entry deleted). Now from the definition (1.7.17),

$$\det \begin{pmatrix} A_{ij} & 0 \\ * & 1 \end{pmatrix} = \det B = \sum_{\pi \in S_n} \operatorname{sgn} \pi \ b_{\pi(1)1} \cdots b_{\pi(n)n}$$

Determinants and Inversion of Matrices

Proof (continued).

Since $b_{\pi(n)n} = \delta_{n\pi(n)}$, we can write

$$\det \begin{pmatrix} A_{ij} & 0 \\ * & 1 \end{pmatrix} = \det B = \sum_{\pi \in S_{n-1}} \operatorname{sgn} \pi \ b_{\pi(1)1} \cdots b_{\pi(n-1)(n-1)} \underbrace{b_{nn}}_{=1} = \det A_{ij},$$

completing the proof. □

Proof of Theorem 1.7.25.

Let $A^{-1} = (x_1, \dots, x_n) = (x_{ij})$ be a matrix of column vectors x_1, \dots, x_n . The inverse of A satisfies $AA^{-1} = \text{id}$, so we need to find columns x_j of A^{-1} satisfying $Ax_j = e_j$, $j = 1, \dots, n$.

By Cramer's rule and Lemma 1.7.26,

$$x_{ij} = \frac{1}{\det A} \det(a_1, \dots, a_{i-1}, e_j, a_{i+1}, \dots, a_n) = \frac{1}{\det A} (-1)^{i+j} A_{ji}. \quad \square$$

Laplace Expansion

Another application of Lemma 1.7.26 is the expansion of $\det A$ in terms of the minors of A :

1.7.27. Laplace Expansion. For $A \in \text{Mat}(n \times n)$ and any $j = 1, \dots, n$ the recursion formula

$$\det A = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det A_{ij} \quad (1.7.21)$$

holds

Note that when using this expansion to calculate the determinant of an $n \times n$ matrix, n determinants of $(n - 1) \times (n - 1)$ matrices need to be evaluated. This means the number of computational steps required is much larger than when Proposition 1.7.19 is used.

Laplace Expansion

Proof.

Let $A = (a_1, \dots, a_n)$, $a_k \in \mathbb{R}^n$, $k = 1, \dots, n$. Then the j th column has the representation $a_j = \sum_{i=1}^n a_{ij} e_i$ and

$$\begin{aligned}\det A &= \det\left(a_1, \dots, a_{j-1}, \sum_{i=1}^n a_{ij} e_i, a_{j+1}, \dots, a_n\right) \\ &= \sum_{i=1}^n a_{ij} \det\left(a_1, \dots, a_{j-1}, e_i, a_{j+1}, \dots, a_n\right) \\ &= \sum_{i=1}^n a_{ij} (-1)^{i+j} \det A_{ij},\end{aligned}$$

where the last equality follows from Lemma 1.7.26. □

 Determinants and Minors

We can obtain the determinant of a matrix as follows:

```
A = {{1, 2, 3}, {4, 5, 6}, {7, 8, 8}};  
MatrixForm[A]
```

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 8 \end{pmatrix}$$

```
Det[A]
```

3

The Mathematica command **Minors** gives the matrix of minors of A . However, the command returns the determinants of the submatrix found by deleting the $(n - i - 1)$ th row and $(m - j - 1)$ th column. To conform to our definition, the command needs to be modified slightly.



Determinants and Minors

```
MatrixForm[Map[Reverse, Minors[A], {0, 1}]]
```

$$\begin{pmatrix} -8 & -10 & -3 \\ -8 & -13 & -6 \\ -3 & -6 & -3 \end{pmatrix}$$

The adjugate matrix can be defined as follows:

```
adj[m_] := Map[Reverse,
  Minors[Transpose[m], Length[m] - 1], {0, 1}]
Table[(-1)^i+j, {i, Length[m]}, {j, Length[m]}]
```

```
MatrixForm[adj[A]]
```

$$\begin{pmatrix} -8 & 8 & -3 \\ 10 & -13 & 6 \\ -3 & 6 & -3 \end{pmatrix}$$

Product Rule for Determinants

1.7.28. Proposition. Let $A, B \in \text{Mat}(n \times n)$. Then $\det(AB) = \det A \det B$.

Proof.

If $A = (a_{ik})$, $B = (b_{kj})$ then $AB = C = (c_{ij})$ with column vectors c_j ,

$$c_j = \begin{pmatrix} c_{1j} \\ \vdots \\ c_{nj} \end{pmatrix}, \quad c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}.$$

Let b_j denote the columns of B , then

$$c_j = Ab_j.$$

We can assume that A is bijective (otherwise $\det(AB) = 0 = \det A \det B$).

Product Rule for Determinants

Proof (continued).

Hence we can write

$$\begin{aligned}\frac{1}{\det A} \det AB &= \frac{1}{\det A} \det(c_1, \dots, c_n) = \frac{1}{\det A} \det(Ab_1, \dots, Ab_n) \\ &= \frac{1}{\det A} \det(A(\cdot), \dots, A(\cdot))[b_1, \dots, b_n] =: f(b_1, \dots, b_n)\end{aligned}$$

The so defined function f is clearly multilinear, because A is linear and \det is multilinear. It is also alternating, because $Ab_k = Ab_j$ if $b_k = b_j$ and \det is alternating. Finally,

$$f(e_1, \dots, e_n) = \frac{1}{\det A} \det(Ae_1, \dots, Ae_n) = 1$$

The function f is multilinear, normed and alternating. Therefore, by the uniqueness of the determinant, it must be the determinant.

Product Rule for Determinants

Proof (continued).

That means that

$$f(B) = \det B \quad \Leftrightarrow \quad \frac{1}{\det A} \det AB = \det B.$$

□

1.7.29. Corollary. Let $A \in \text{Mat}(n \times n)$ be invertible. Then

$$\det A^{-1} = \frac{1}{\det A}.$$

First Midterm Exam

The preceding material completes the first third of the course material. It encompasses everything that will be the subject of the **First Midterm Exam**.

The exam date will be posted on Canvas.

No calculators or other aids will be permitted during the exam. A sample exam with solutions has been uploaded to Canvas. Please study it carefully, including the instructions on the cover page.

Part II

Differential Calculus

Convergence and Continuity

Functions and Derivatives

Curves in Vector Spaces

Potential Functions

The Second Derivative

Extrema of Potential Functions

Constrained Extrema

Convergence and Continuity

Functions and Derivatives

Curves in Vector Spaces

Potential Functions

The Second Derivative

Extrema of Potential Functions

Constrained Extrema

Finite-Dimensional Vector Spaces

For the rest of the term, we will focus on functions of several variables, e.g., functions

$$f: \mathbb{R}^n \rightarrow \mathbb{R}^m.$$

Our previously developed knowledge of linear algebra is essential for this, since, for example, the derivative of such a function at a point x turns out to be a matrix. More precisely, the derivative is a map

$$Df: \mathbb{R}^n \rightarrow \text{Mat}(m \times n; \mathbb{R}).$$

It is not sufficient to restrict ourselves to functions defined on \mathbb{R}^n with values in \mathbb{R}^m . On the one hand, the second derivative of f is then the derivative of Df , a matrix-valued function. Another aspect occurs in the study of ordinary differential equations, when we need to differentiate functions such as the determinant of a matrix. Therefore, we need to define concepts such as continuity for arbitrary vector spaces.

Open Balls

The basic ingredient in our discussion are open balls:

2.1.1. Definition. Let $(V, \|\cdot\|)$ be a normed vector space. Then

$$B_\varepsilon(a) := \{x \in V : \|x - a\| < \varepsilon\}, \quad a \in V, \quad \varepsilon > 0, \quad (2.1.1)$$

is called an **open ball** of radius ε about a .

Of course, the “shape” of an open ball depends on the vector space V and the norm $\|\cdot\|$. For instance, the open balls in \mathbb{R}^2 with norms

$$\begin{aligned}\|x\|_1 &= |x_1| + |x_2|, \\ \|x\|_2 &= \sqrt{|x_1|^2 + |x_2|^2}, \\ \|x\|_\infty &= \max\{|x_1|, |x_2|\}\end{aligned} \quad (2.1.2)$$

all have quite different shapes.

Furthermore, if $V = \mathcal{P}_n$, for example, open balls do not have an obvious “shape” at all.

Open Sets

2.1.2. Definition. Let $(V, \|\cdot\|)$ be a normed vector space. A set $U \subset V$ is called **open** if for every $a \in U$ there exists an $\varepsilon > 0$ such that $B_\varepsilon(a) \subset U$.

2.1.3. Examples.

- (i) Any open ball $B_\varepsilon(a)$, $\varepsilon > 0$, $a \in V$, is an open set.
(For any $b \in B_\varepsilon(a)$ take $\delta < \varepsilon - \|a - b\|$. Then $B_\delta(b) \subset B_\varepsilon(a)$.)
- (ii) The empty set $\emptyset \subset V$ is open.
(Since there is no $a \in \emptyset$ for which we need to check that $B_\varepsilon(a) \subset \emptyset$, this is an example of a **vacuously true statement**.)
- (iii) The entire space V is an open set in V .

Open Sets

We will see that open sets are fundamental for understanding properties of continuous functions, convergence in vector spaces and much more.

Therefore, it becomes important to answer a basic question:

If a set is open in a vector space $(V, \|\cdot\|)$, is it also open if $\|\cdot\|$ is replaced by some other norm?

2.1.4. Example. If a set $\Omega \subset \mathbb{R}^2$ is open with respect to any one of the norms (2.1.2), it is also open with respect to any of the other norms given in (2.1.2) here. Why?

Equivalent Norms

2.1.5. Definition. Let V be a vector space on which we may define two norms $\|\cdot\|_1$ and $\|\cdot\|_2$. Then the two norms are called **equivalent** if there exists two constants $C_1, C_2 > 0$ such that

$$C_1\|x\|_1 \leq \|x\|_2 \leq C_2\|x\|_1 \quad \text{for all } x \in V. \quad (2.1.3)$$

2.1.6. Example. In \mathbb{R}^n we have (amongst others) the following two possible choices of norms:

$$\|x\|_2 := \left(\sum_{i=1}^n |x_i|^2 \right)^{1/2}, \quad \|x\|_\infty := \max_{1 \leq i \leq n} |x_i|. \quad (2.1.4)$$

It is easily verified that for all $x \in \mathbb{R}^n$,

$$\frac{1}{\sqrt{n}}\|x\|_2 \leq \|x\|_\infty \leq \|x\|_2,$$

so the two norms are equivalent.

Convergence of Sequences

2.1.7. Remark. It is obvious from the definition that if two norms on a vector space are equivalent, then any set that is open with respect to the first norm is also open with respect to the second norm.

We recall the following from Vv186:

2.1.8. Definition. Let $(V, \|\cdot\|)$ be a normed vector space and (v_n) a sequence in V . Then (v_n) converges to a (unique) limit $v \in V$,

$$v_n \xrightarrow{n \rightarrow \infty} v \quad \text{if and only if} \quad \|v_n - v\| \xrightarrow{n \rightarrow \infty} 0.$$

For later use, we note:

2.1.9. Remark. If a sequence (v_n) in $(V, \|\cdot\|)$ converges to $v \in V$, then $\|v_n\| \rightarrow \|v\|$. This follows from

$$\|v_n - v\| \leq \|v_n\| + \|v\| \rightarrow 0.$$

Equivalence of All Norms

2.1.10. Remark. It is again easy to see from the definition that if two norms on a vector space are equivalent, then a sequence that converges to a limit with respect to the first norm also converges to the same limit with respect to the second norm.

Therefore, the following theorem is of fundamental importance:

2.1.11. Theorem. In a finite-dimensional vector space, all norms are equivalent.

A major consequence of Theorem 2.1.11 is that if we have several norms at our disposal in a finite-dimensional space, then we can freely choose a convenient one in order to show openness of sets, convergence of sequences, etc.

The proof of Theorem 2.1.11 requires some preliminary work.

The Theorem of Bolzano-Weierstraß

We recall two basic facts from the theory of sequences of real numbers:

- (i) Every bounded and monotonic sequence of real numbers converges.
- (ii) Every sequence of real numbers has a monotonic subsequence.

Together, these yield the following fundamental result (cf.
186 Theorem 2.2.35):

2.1.12. Theorem of Bolzano-Weierstraß. Every bounded sequence of real numbers has a convergent subsequence.

We remark that the Theorem of Bolzano-Weierstraß easily implies that every Cauchy sequence of real numbers converges, because every Cauchy sequence that has a convergent subsequence must itself converge. Thus the basic ingredient in proving that the real numbers (with the usual metric) are complete is the fact that a bounded, monotonic sequence converges.

The Theorem of Bolzano-Weierstraß in \mathbb{R}^n

2.1.13. Theorem of Bolzano-Weierstraß in \mathbb{R}^n . Let $(x^{(m)})_{m \in \mathbb{N}}$ be a sequence of vectors in \mathbb{R}^n , i.e., $x^{(m)} = (x_1^{(m)}, \dots, x_n^{(m)})$. Suppose that there exists a constant $C > 0$ such that $|x_k^{(m)}| < C$ for all $m \in \mathbb{N}$ and each $k = 1, \dots, n$. Then there exists a subsequence $(x^{(m_j)})_{j \in \mathbb{N}}$ that converges to a vector $y \in \mathbb{R}^n$ in the sense that

$$x_k^{(m_j)} \xrightarrow{j \rightarrow \infty} y_k \quad \text{for } k = 1, \dots, n.$$

Proof.

Consider the real coordinate sequence $(x_1^{(m)})_{m \in \mathbb{N}}$. By assumption, this sequence is bounded, so by the Theorem of Bolzano-Weierstraß 2.1.12 there exists a convergent subsequence $(x_1^{(m_{j_1})})$ with some limit, say $y_1 \in \mathbb{R}$.

The second coordinate sequence $(x_2^{(m)})$ is also bounded and has a convergent subsequence, but this subsequence does not need to have the same indices as that for $(x_1^{(m)})$.

The Theorem of Bolzano-Weierstraß in \mathbb{R}^n

Proof (continued).

We therefore employ a trick: The subsequence $(x_2^{(m_{j_1})})$ that uses the indices from our above subsequence for the first coordinate is of course also bounded and hence has a sub-subsequence $(x_2^{(m_{j_2})})$ that converges, say to $y_2 \in \mathbb{R}$. Taking the corresponding sub-subsequence for the first coordinate, $(x_1^{(m_{j_2})})$ still converges to y_1 .

Similarly, we a sub-sub-subsequence of the third coordinate will converge to some $y_3 \in \mathbb{R}$ while the corresponding sub-sub subsequences of the first two coordinates will still converge to y_1 and y_2 , respectively. Repeating the procedure n times, the n -fold subsequence $(x_k^{(m_{j_n})})$ converges to some $y_k \in \mathbb{R}$, $k = 1, \dots, n$. Hence, the subsequence $(x^{(m_{j_n})})$ converges to some $y \in \mathbb{R}^n$. □

A Basic Norm inequality

2.1.14. Lemma. Let $(V, \|\cdot\|)$ be a finite- or infinite-dimensional normed vector space and $\{v_1, \dots, v_n\}$ an independent set in V . Then there exists a $C > 0$ such that for any $\lambda_1, \dots, \lambda_n \in \mathbb{F}$

$$\|\lambda_1 v_1 + \dots + \lambda_n v_n\| \geq C(|\lambda_1| + \dots + |\lambda_n|). \quad (2.1.5)$$

Proof.

Let $s := |\lambda_1| + \dots + |\lambda_n|$. If $s = 0$, then all $\lambda_k = 0$ and the inequality (2.1.5) holds trivially for any C , so we can assume $s > 0$. Dividing by s , (2.1.5) becomes

$$\|\mu_1 v_1 + \dots + \mu_n v_n\| \geq C, \quad \sum_{k=1}^n |\mu_k| = 1, \quad (2.1.6)$$

with $\mu_k = \lambda_k/s$.

A Basic Norm inequality

Proof (continued).

Hence, we need to show

$$\exists C > 0 \quad \forall \begin{array}{c} \mu_1, \dots, \mu_n \in \mathbb{F} \\ |\mu_1| + \dots + |\mu_n| = 1 \end{array} \quad \|\mu_1 v_1 + \dots + \mu_n v_n\| \geq C.$$

Suppose that this is false, i.e.,

$$\forall C > 0 \quad \exists \begin{array}{c} \mu_1, \dots, \mu_n \in \mathbb{F} \\ |\mu_1| + \dots + |\mu_n| = 1 \end{array} \quad \|\mu_1 v_1 + \dots + \mu_n v_n\| < C.$$

In particular, choosing $C = 1/m$, $m = 1, 2, 3, \dots$, we can find a sequence of vectors

$$u^{(m)} := \mu_1^{(m)} v_1 + \dots + \mu_n^{(m)} v_n$$

such that $\|u^{(m)}\| \rightarrow 0$ as $m \rightarrow \infty$ and $|\mu_1^{(m)}| + \dots + |\mu_n^{(m)}| = 1$ for all m .

A Basic Norm inequality

Proof (continued).

Hence, for each $k = 1, \dots, n$, $|\mu_k^{(m)}| \leq 1$ and so each coefficient sequence $(\mu_k^{(m)})$ is bounded. Write

$$\mu^{(m)} := (\mu_1^{(m)}, \dots, \mu_n^{(m)})$$

By the Theorem of Bolzano Weierstraß in \mathbb{R}^n , there exists a subsequence of vectors $(\mu^{(m_j)})_{j \in \mathbb{N}}$ that converges to some $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$. This corresponds to a subsequence $u^{(m_j)}$ of $u^{(m)}$ such that

$$u^{(m_j)} \xrightarrow{j \rightarrow \infty} \alpha_1 v_1 + \cdots + \alpha_n v_n =: u \quad \text{with } |\alpha_1| + \cdots + |\alpha_n| = 1.$$

Since the vectors v_1, \dots, v_n are independent and not all α_k vanish, it follows that $u \neq 0$.

A Basic Norm inequality

Proof (continued).

Remark 2.1.9 then implies,

$$\|u^{(m_j)}\| \xrightarrow{j \rightarrow \infty} \|u\| \neq 0.$$

But by our construction, $\|u^{(m)}\| \rightarrow 0$ as $m \rightarrow \infty$, so the subsequence $(\|u^{(m_j)}\|)$ must also converge to zero. This gives a contradiction. □

We can now proceed to prove Theorem 2.1.11.

Equivalence of Norms

Proof of Theorem 2.1.11.

Let V be a finite-dimensional vector space, $\|\cdot\|$ be any norm on V and $\{v_1, \dots, v_n\}$ a basis of V . Let $v \in V$ have the representation $v = \lambda_1 v_1 + \dots + \lambda_n v_n$ with $\lambda_1, \dots, \lambda_n \in \mathbb{F}$. By the triangle inequality,

$$\|v\| = \|\lambda_1 v_1 + \dots + \lambda_n v_n\| \leq \sum_{i=1}^n |\lambda_i| \|v_i\| \leq C \sum_{i=1}^n |\lambda_i|$$

where $C := \max_{1 \leq i \leq n} \|v_i\|$ depends only on the basis and not on v . We hence see that for any norm there are constants $C_1, C_2 > 0$ such that

$$C_1 \sum_{i=1}^n |\lambda_i| \leq \|v\| \leq C_2 \sum_{i=1}^n |\lambda_i|, \quad (2.1.7)$$

where the first inequality is just (2.1.5). Given two norms $\|\cdot\|_1$ and $\|\cdot\|_2$, it follows from their respective inequalities (2.1.7) that (2.1.3) holds. \square

Equivalence of Norms

It is essential that Theorem 2.1.11 assumes that V is a finite-dimensional vector space. In an infinite-dimensional vector space, it is possible to define non-equivalent norms.

2.1.15. Example. Consider the space of continuous functions on $[0, 1]$, $C([0, 1])$. We can define the two norms

$$\|f\|_{\infty} = \sup_{x \in [0,1]} |f(x)|, \quad \|f\|_1 = \int_0^1 |f(x)| dx.$$

You will show in the assignments that these two norms are not equivalent.

Interior, Exterior and Boundary Points

2.1.16. Definition. Let $(V, \|\cdot\|)$ be a normed vector space and $M \subset V$.

- (i) A point $x \in M$ is called an **interior point of M** if there exists an $\varepsilon > 0$ such that $B_\varepsilon(x) \subset M$.
- (ii) The set of interior points of M is denoted by $\text{int } M$.
- (iii) A point $x \in V$ is called a **boundary point of M** if for every $\varepsilon > 0$ $B_\varepsilon(x) \cap M \neq \emptyset$ and $B_\varepsilon(x) \cap (V \setminus M) \neq \emptyset$.
- (iv) The set of boundary points of M is denoted by ∂M .
- (v) A point that is neither a boundary nor an interior point of M is called an **exterior point of M** .

2.1.17. Remarks.

- (i) An exterior point of M is an interior point of $V \setminus M$. (Check this!)
- (ii) For given M , any point of V is either an interior, boundary or exterior point of M .

Closed Sets

2.1.18. Definition. Let $(V, \|\cdot\|)$ be a normed vector space and $M \subset V$. Then M is said to be **closed** if its complement $V \setminus M$ is open.

2.1.19. Remark. Of course, a set M does not need to be either open or closed. Some sets are open and closed at the same time.

2.1.20. Examples.

- (i) A set consisting of a single point, $M = \{a\} \subset V$, is a closed set.
- (ii) The empty set $\emptyset \subset V$ is closed.
- (iii) The entire space V is a closed set in V .

Closed Sets

2.1.21. Lemma. Let $(V, \|\cdot\|)$ be a normed vector space and $M \subset V$.

- (i) The set M is open if and only if $M = \text{int } M$.
- (ii) The set M is closed if and only if $\partial M \subset M$.

Proof.

- (i) This is just a restatement of the definition of an open set.
- (ii) Suppose that M is closed. Then $V \setminus M$ is open. An open set can not contain a boundary point, since all its points are interior points. Hence, $\partial M \cap (V \setminus M) = \emptyset$ and so $\partial M \subset M$.

Suppose that $\partial M \subset M$. Then $V \setminus M$ contains only exterior points of M . But an exterior point of M is an interior point of $V \setminus M$, so $V \setminus M$ is open. Hence, M is closed. □

The Closure

2.1.22. Definition. Let $(V, \|\cdot\|)$ be a normed vector space and $M \subset V$. Then

$$\overline{M} := M \cup \partial M$$

is called the ***closure*** of M .

2.1.23. Remark. It is not hard to show that the closure of a set M is a closed set. In fact, it is the smallest set that both contains M and is closed.

The closure of a set may also be characterized in terms of sequences:

2.1.24. Lemma. Let $(V, \|\cdot\|)$ be a normed vector space and $M \subset V$. Then

$$\overline{M} = \left\{ x \in V : \exists_{(x_n)_{n \in \mathbb{N}}} x_n \in M \text{ and } x_n \rightarrow x \right\} \quad (2.1.8)$$

The Closure

Proof.

- (i) Suppose that $x \in V$ is such that there exists a sequence (x_n) with $x_n \in M$ and $x_n \rightarrow x$. Then for every $\varepsilon > 0$, $B_\varepsilon(x)$ contains at least one x_n . Hence, $B_\varepsilon(x) \cap M \neq \emptyset$ and so x can not be an exterior point. This implies $x \in M \cup \partial M$.
- (ii) Suppose $x \in M \cup \partial M$. Then for every $\varepsilon > 0$, $B_\varepsilon(x) \cap M \neq \emptyset$. Choose $\varepsilon = 1/n$ for $n \in \mathbb{N} \setminus \{0\}$ to find a sequence of points $x_n \in B_{1/n}(x) \cap M$. This sequence converges to x , so x is in the set on the right-hand side of (2.1.8). □

Continuous Functions

Recall the following definition of continuity in normed vector spaces:

2.1.25. Definition. Let $(U, \|\cdot\|_1)$ and $(V, \|\cdot\|_2)$ be normed vector spaces and $f: U \rightarrow V$ a function. Then f is **continuous at $a \in U$** if

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x \in U \quad \|x - a\|_1 < \delta \quad \Rightarrow \quad \|f(x) - f(a)\|_2 < \varepsilon. \quad (2.1.9)$$

Of course, we can prove as usual the following:

2.1.26. Theorem. Let $(U, \|\cdot\|_1)$ and $(V, \|\cdot\|_2)$ be normed vector spaces and $f: U \rightarrow V$ a function. Then f is **continuous at $a \in U$** if and only if

$$\forall_{(x_n)_{n \in \mathbb{N}} \atop x_n \in U} \quad x_n \rightarrow a \quad \Rightarrow \quad f(x_n) \rightarrow f(a). \quad (2.1.10)$$

Image and Pre-Image of Sets

Suppose that $f: M \rightarrow N$, where M, N are any sets. Let $A \subset M$. Then we define the **image of A** by

$$f(A) := \{y \in N : y = f(x) \text{ for some } x \in A\}.$$

In particular, we can write

$$\text{ran } f = f(M).$$

Similarly, for $B \subset N$ we define the **pre-image of B** by

$$f^{-1}(B) := \{x \in M : f(x) = y \text{ for some } y \in B\}. \quad (2.1.11)$$

2.1.27. Examples.

- (i) Let $f: \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = \sin x$. Then $f([0, \pi]) = [0, 1]$.
- (ii) Let $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = x^2 + y^2$. Then

$$f^{-1}(\{1\}) = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$$

(This is the unit circle in \mathbb{R}^2).

Continuous Functions

It is often useful to characterize continuous maps by using open sets:

2.1.28. Theorem. Let $(U, \|\cdot\|_1)$ and $(V, \|\cdot\|_2)$ be normed vector spaces and $f: U \rightarrow V$ a function. Then f is continuous if and only if the pre-image $f^{-1}(\Omega)$ of every open set $\Omega \subset V$ is open.

Proof.

(\Rightarrow) Let f be continuous and $\Omega \subset V$ open. We will show that $f^{-1}(\Omega)$ is open. Let $a \in f^{-1}(\Omega)$. Then $f(a) \in \Omega$, and since Ω is open we can find $\varepsilon > 0$ such that $B_\varepsilon(f(a)) \subset \Omega$.

Now let $\delta > 0$. By the continuity of f we can choose δ small enough to ensure that $f(B_\delta(a)) \subset B_\varepsilon(f(a))$. But then $B_\delta(a) \subset f^{-1}(\Omega)$. Since this is true for any $a \in f^{-1}(\Omega)$, it follows that $f^{-1}(\Omega)$ is open.

Continuous Functions

Proof (continued).

(\Leftarrow) Let $f: U \rightarrow V$ be such that the pre-image $f^{-1}(\Omega)$ of every open set $\Omega \subset V$ is open. We will show that f is continuous. Let $a \in U$ be arbitrary and fix $\varepsilon > 0$. We want to show that there exists a $\delta > 0$ such that

$$x \in B_\delta(a) \quad \Rightarrow \quad f(x) \in B_\varepsilon(f(a)). \quad (2.1.12)$$

The set $B_\varepsilon(f(a))$ is open, and by assumption $f^{-1}(B_\varepsilon(f(a))) \ni a$ is also open. Thus, we can find $\delta > 0$ such that $B_\delta(a) \subset f^{-1}(B_\varepsilon(f(a)))$. But then (2.1.12) holds and we are finished. □

Continuous Functions

2.1.29. Example. We show that the function

$$\det: \text{Mat}(n \times n; \mathbb{C}) \rightarrow \mathbb{C}, \quad \det A = \sum_{\pi \in S_n} \operatorname{sgn} \pi \ a_{\pi(1)1} \cdots a_{\pi(n)n}$$

is continuous.

In particular, we can choose to use the norm $\|A\| = \max_{i,j} |a_{ij}|$. Then fix $A = (a_{ij}) \in \text{Mat}(n \times n; \mathbb{C})$ and suppose that (A_m) is a sequence converging to A . Our choice of norm implies that all coefficients converge, $a_{ij}^{(m)} \rightarrow a_{ij}$. Since $\det A$ is a polynomial in the coefficients a_{ij} , $\det A_m \rightarrow \det A$ and therefore \det is continuous at $A \in \text{Mat}(n \times n; \mathbb{C})$.

Note that the pre-image of the set of non-zero complex numbers is

$$\det^{-1}(\mathbb{C} \setminus \{0\}) = \text{GL}(n; \mathbb{C}),$$

the **general linear** group of invertible matrices. Since $\mathbb{C} \setminus \{0\}$ is an open set, Theorem 2.1.28 implies that $\text{GL}(n; \mathbb{C})$ is an open set in $\text{Mat}(n \times n; \mathbb{C})$.

Compact Sets

We are now interested in generalizing the results of Vv186 that apply to continuous functions on closed intervals to vector spaces. Note that a closed interval in \mathbb{R} is always bounded in the following sense

2.1.30. Definition. Let $(V, \|\cdot\|)$ be a normed vector space and $M \subset V$. Then M is said to be **bounded** if there exists some $R > 0$ such that $M \subset B_R(0)$.

It turns out that the natural generalization of a closed interval is a little more complicated than just requiring a set to be closed and bounded.

2.1.31. Definition. Let $(V, \|\cdot\|)$ be a normed vector space and $K \subset V$. Then K is said to be **compact** if every sequence in K has a convergent subsequence with limit contained in K .

Compact Sets are Closed and Bounded

2.1.32. Theorem. Let $(V, \|\cdot\|)$ be a (possibly infinite-dimensional) normed vector space and $K \subset V$ be compact. Then K is closed and bounded.

Proof.

We first show that K is closed by establishing $K = \overline{K}$. Let $x \in \overline{K}$. Then there exists a sequence (x_n) in K converging to x . Since K is compact, (x_n) has a subsequence (x_{n_k}) that converges to $x' \in K$. Since (x_n) converges to x , $x = x' \in K$, so $K = \overline{K}$ and K is closed.

Now suppose that K is unbounded. Then for any $n \in \mathbb{N}$ there exists an $x_n \in K$ such that $\|x_n\| > n$. This gives rise to an unbounded sequence (x_n) . Furthermore, any subsequence of (x_n) is unbounded. Since a convergent sequence is bounded, we conclude that (x_n) can not have a convergent subsequence. This implies that K is not compact. By contraposition, if K is compact, then K must be bounded. □

Closed and Bounded Sets are Sometimes Compact

2.1.33. Theorem. Let $(V, \|\cdot\|)$ be a **finite-dimensional** vector space and let $K \subset V$ be closed and bounded. Then K is compact.

Proof.

Suppose that (b_1, \dots, b_n) be a basis of V and K closed and compact. Let (v_m) be a sequence in K . Then each sequence term has the representation

$$v_m = \lambda_1^{(m)} b_1 + \cdots + \lambda_n^{(m)} b_n, \quad \lambda_1^{(m)}, \dots, \lambda_n^{(m)} \in \mathbb{F}, \quad m \in \mathbb{N}.$$

By Lemma 2.1.14 and the boundedness of K , there exist constants $C_1, C_2 > 0$ such that

$$C_1 \geq \|v_m\| \geq C_2 \sum_{k=1}^n |\lambda_k^{(m)}|.$$

Closed and Bounded Sets are Sometimes Compact

Proof (continued).

It follows that for each k , the sequence $(\lambda_k^{(m)})$ is bounded. Write

$$\lambda^{(m)} = (\lambda_1^{(m)}, \dots, \lambda_n^{(m)}).$$

By the Theorem of Bolzano-Weierstraß in \mathbb{R}^n , $(\lambda^{(m)})$ has a convergent subsequence $(\lambda^{(m_j)})$ so that (v_{m_j}) converges to some element $v \in \overline{K}$. Since K is closed, $v \in K$. This implies that K is compact. □

Closed and Bounded Sets are Sometimes Compact

Theorem 2.1.33 is in general false in infinite-dimensional spaces:

2.1.34. Example. Consider the vector space of summable complex sequences,

$$l^1 := \left\{ (a_n) : \mathbb{N} \rightarrow \mathbb{C} : \sum_{n=0}^{\infty} |a_n| < \infty \right\}.$$

The natural norm is given by

$$\|(a_n)\|_1 := \sum_{n=0}^{\infty} |a_n|.$$

Then the set

$$\overline{B_1(0)} = \left\{ (a_n) \in l^1 : \sum_{n=0}^{\infty} |a_n| \leq 1 \right\}$$

is closed and bounded, but not compact. (You will prove this in the assignments.)

Compact Sets and Continuity

Why are we so interested in compact sets? Well, it turns out that compact sets are natural extensions of closed intervals in \mathbb{R} for the purpose of generalizing some major theorems on continuous functions.

2.1.35. Proposition. Let $(U, \|\cdot\|_1)$, $(V, \|\cdot\|_2)$ be normed vector spaces and $K \subset U$ compact. Let $f: K \rightarrow V$ be continuous. Then $\text{ran } f = f(K)$ is compact in V .

Proof.

Let (y_n) be a sequence in $f(K)$. Then there exists a sequence (x_n) in K with $y_n = f(x_n)$. Since K is compact, a subsequence (x_{n_k}) of (x_n) converges to some $a \in K$. But because f is continuous the subsequence $(f(x_{n_k}))$ of (y_n) converges to $f(a) \in f(K)$. Hence, (y_n) has a convergent subsequence and $f(K)$ is compact. □

Extrema of Continuous Functions on Compact Sets

2.1.36. Theorem. Let $(V, \|\cdot\|)$ be a normed vector space and $K \subset V$ compact. Let $f: K \rightarrow \mathbb{R}$ be continuous. Then f has a maximum in K , i.e., there exists an $x \in K$ such that $f(y) \leq f(x)$ for all $y \in K$.

Proof.

The range $\text{ran } f = f(K)$ is compact by Proposition 2.1.35, so it is closed and bounded by Theorem 2.1.32. The least upper bound $b = \sup f(K)$ exists because $f(K)$ is bounded.

Since b is the **least** upper bound, b can not be an exterior point of $f(K)$, so $b \in \overline{f(K)}$. Since $f(K)$ is closed, $\overline{f(K)} = f(K)$ and $b \in f(K)$.

Hence, there exists an $x \in K$ with $f(x) = b$ and $f(y) \leq b$ for all $y \in K$. □

Uniform Continuity on Compact Sets

Recall the definition of uniform continuity for functions in vector spaces:

2.1.37. Definition. Let $(U, \|\cdot\|_1)$ and $(V, \|\cdot\|_2)$ be normed vector spaces, $\Omega \subset U$ and $f: \Omega \rightarrow V$ a function. Then f is **uniformly continuous in Ω** if

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x, y \in \Omega \quad \|x - y\|_1 < \delta \quad \Rightarrow \quad \|f(x) - f(y)\|_2 < \varepsilon. \quad (2.1.13)$$

(Compare with Definition 2.1.25.)

2.1.38. Theorem. Let $(U, \|\cdot\|_1)$ and $(V, \|\cdot\|_2)$ be normed vector spaces, $K \subset U$ a compact set and $f: K \rightarrow V$ continuous on K . Then f is uniformly continuous on K .

Uniform Continuity on Compact Sets

Proof.

Suppose that f is continuous but not uniformly continuous on K . Then

$$\exists \varepsilon > 0 \quad \forall \delta > 0 \quad \exists x, y \in K \quad \|x - y\|_1 < \delta \quad \wedge \quad \|f(x) - f(y)\|_2 \geq \varepsilon.$$

Denote this ε by ε_0 . Then for each $\delta = 1/n$ there exist vectors $x_n, y_n \in K$ such that

$$\|x_n - y_n\|_1 < \frac{1}{n} \quad \wedge \quad \|f(x_n) - f(y_n)\|_2 \geq \varepsilon_0.$$

Since K is compact, there exist subsequences (x_{n_k}) and (y_{n_k}) that converge, say to ξ and η , respectively. Since $\|x_{n_k} - y_{n_k}\|_1 < \frac{1}{n_k}$, we see that $\xi = \eta$. However, then

$$x_{n_k} \rightarrow \xi \quad \wedge \quad y_{n_k} \rightarrow \xi \quad \wedge \quad \|f(x_{n_k}) - f(y_{n_k})\|_2 \geq \varepsilon_0 \not\rightarrow 0.$$

which contradicts the continuity of f at ξ . □

Convergence and Continuity

Functions and Derivatives

Curves in Vector Spaces

Potential Functions

The Second Derivative

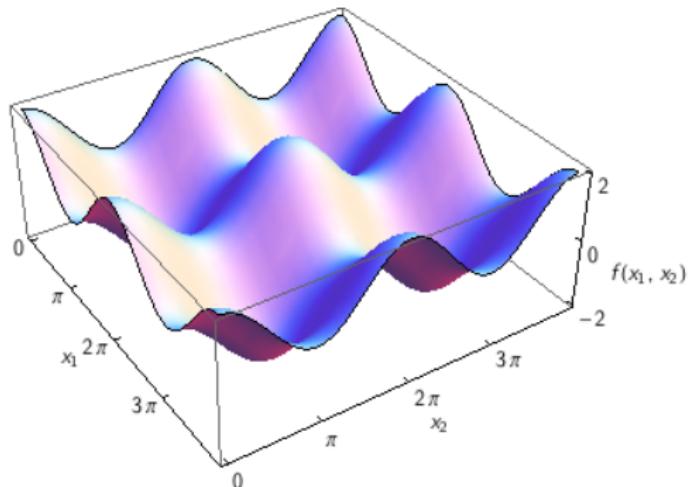
Extrema of Potential Functions

Constrained Extrema

Representing Functions $f: \mathbb{R}^2 \rightarrow \mathbb{R}$

Suppose we have a function $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, i.e., a real function of two variables. One method of graphing such a function is using a three-dimensional graph showing the (x_1, x_2, z) -axes and plotting $z = f(x_1, x_2)$. For example, the graph below shows the function

$$f: [0, 4\pi] \times [0, 4\pi] \rightarrow \mathbb{R}, \quad f(x_1, x_2) = \cos x_1 + \cos x_2$$



3D Plots with Mathematica

2.2.1. Example. The following Mathematica command creates the three-dimensional plot on the previous slide:

```
Plot3D[Cos[x] + Cos[y], {x, 0, 4 Pi}, {y, 0, 4 Pi},  
Mesh → False,  
Ticks → {{0, π, 2 π, 3 π}, {0, π, 2 π, 3 π}, {-2, 0, 2}},  
AxesLabel → {x1, x2, f[x1, x2]},  
BaseStyle → {FontSize → 12, FontFamily → "CMU Sans Serif"}]
```

Contour Plots

Another representation for functions $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ is the so-called **contour plot**. In this two-dimensional graph we plot curves

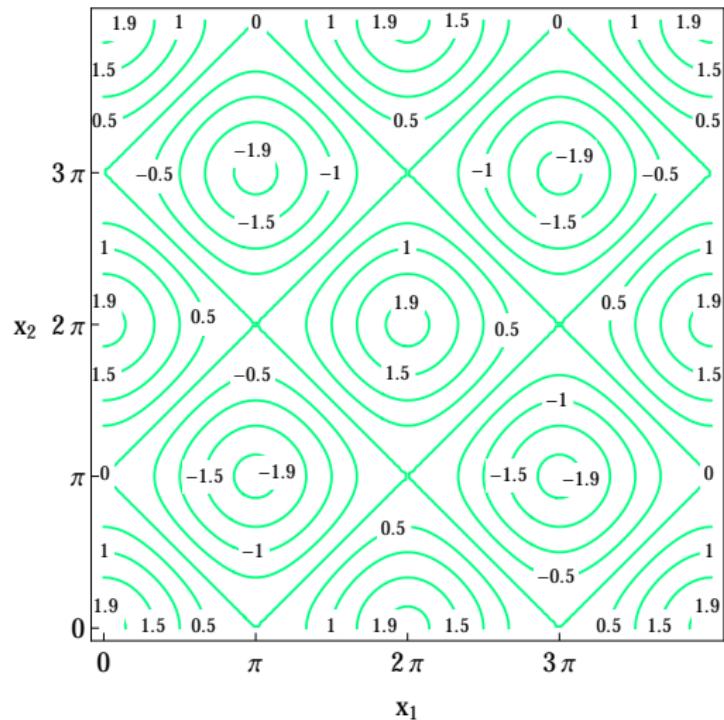
$$C_\alpha = f^{-1}(\{\alpha\})$$

for several values of α . These are the pre-image sets (see (2.1.11)) of $\{\alpha\}$.

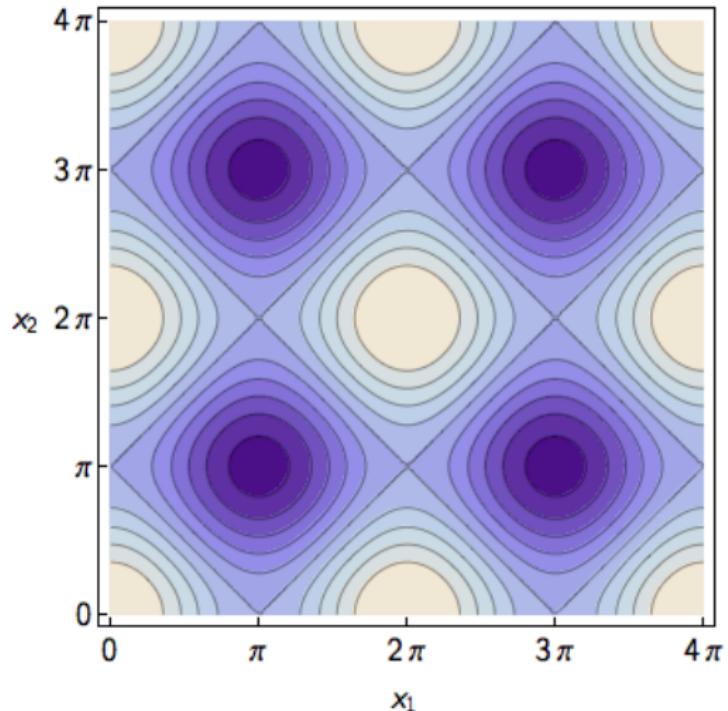
Contour Plots

To illustrate, we successively show contours of

$$f: [0, 4\pi] \times [0, 4\pi] \rightarrow \mathbb{R}, \quad f(x_1, x_2) = \cos x_1 + \cos x_2$$



Contour Plots Instead of labeling, Mathematica can also color-code the contours according to their values. Here, dark colors represent smaller values, light colors larger values.



 Contour Plots with Mathematica

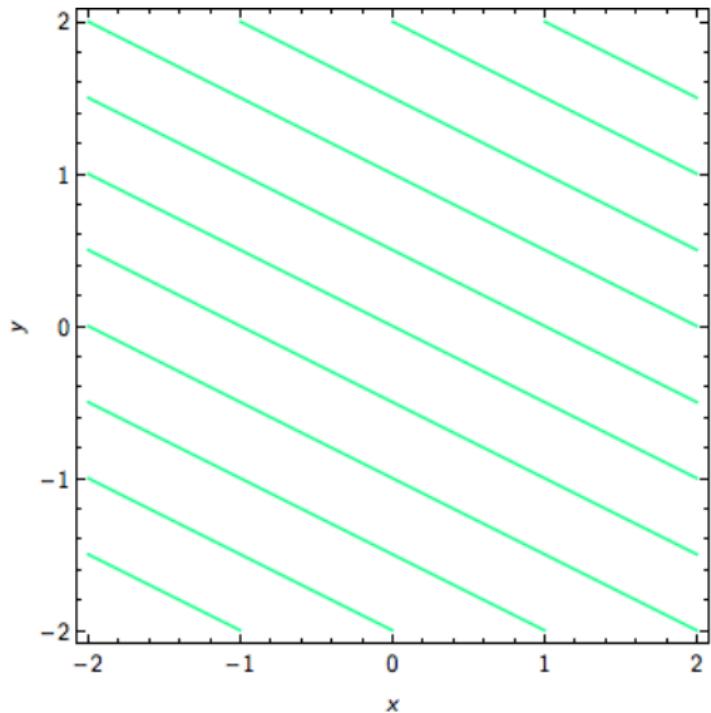
2.2.2. Example. The following Mathematica commands creates the contour plots on the previous slides:

```
ContourPlot[Cos[x] + Cos[y], {x, 0, 4 Pi}, {y, 0, 4 Pi},
FrameLabel → {x1, x2}, RotateLabel → False,
FrameTicks → {{0, π, 2 π, 3 π, 4 π}, {0, π, 2 π, 3 π, 4 π}, {}, {}},
ContourStyle → {{RGBColor[0, 1, 0.5], Thickness[0.004]}},
BaseStyle → {FontSize → 14, FontFamily → "CMU Sans Serif"},
PlotPoints → 50, ContourLabels →
  (Text[Framed[#, FrameStyle → White, FrameMargins → 0.2],
    {#1, #2}, Background → White, BaseStyle → {FontSize → 10}] &),
Contours → {0, 0.5, -0.5, -1, -1.5, 1, 1.5, 1.9, -1.9},
ContourShading → None]
```

```
ContourPlot[Cos[x] + Cos[y], {x, 0, 4 Pi}, {y, 0, 4 Pi},
FrameLabel → {x1, x2}, RotateLabel → False,
FrameTicks → {{0, π, 2 π, 3 π, 4 π}, {0, π, 2 π, 3 π, 4 π}, {}, {}},
BaseStyle → {FontSize → 16, FontFamily → "CMU Sans Serif"},
PlotPoints → 50, Contours → 10]
```

Contour Plots

2.2.3. Example. Sketch the contour plot of the function $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x) = x + y$.



Phase Curves

In the hamiltonian formulation of analytical mechanics, one defines a so-called **Hamilton function** H for a mechanical system. This function is the sum of the kinetic energy (T) and potential energy (V). It represents the total energy of the system, and remains constant if the system satisfies the law of energy conservation (there are not, for example, any frictional forces). We will assume this for our present discussion.

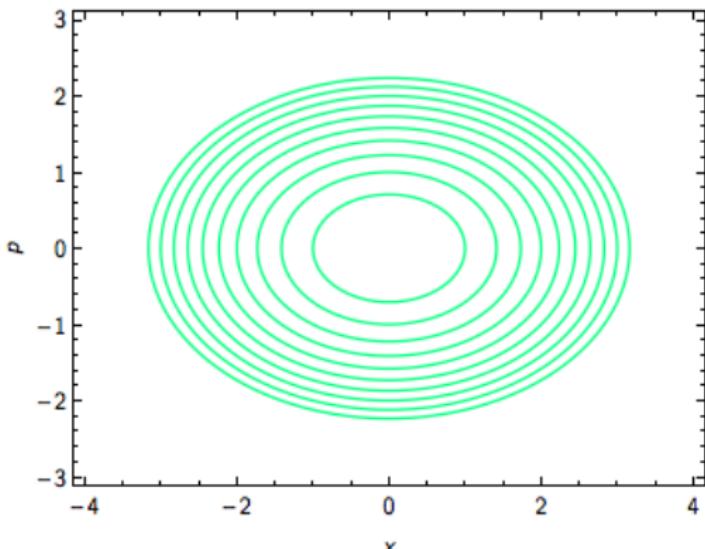
In the hamiltonian formulation of mechanics, the essential variables of a system are the position x and the momentum p . The variables are tracked in so-called **phase space** $\mathbb{R}_x^n \times \mathbb{R}_p^n = \mathbb{R}_{(x,p)}^{2n}$, where, typically, $n = 1, 2$ or 3 . The time-evolution of the system is represented through **phase curves** in \mathbb{R}^{2n} , which are given by the contour lines of H , which is regarded as a function $\mathbb{R}^{2n} \rightarrow \mathbb{R}$. In other words, a phase curve is the set $H^{-1}(E)$, where E is the conserved energy of the system.

Phase Curves

2.2.4. Example. For the simple harmonic oscillator, the kinetic energy is given by $T = \frac{1}{2}mv^2 = p^2/(2m)$ and the potential energy is given by $V = \frac{k}{2}x^2$, so

$$H(x, p) = \frac{1}{2m}p^2 + \frac{k}{2}x^2.$$

The phase curves of the system are ellipses in $\mathbb{R}_{x,p}^2$, with each ellipse describing the behavior of a harmonic oscillator at a fixed energy E .

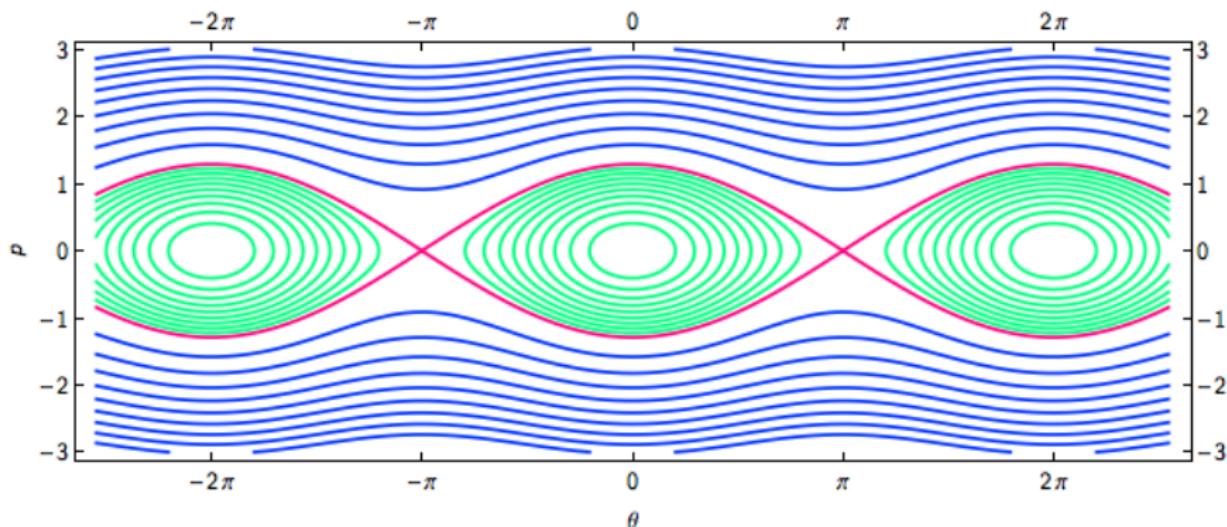


Phase Curves

2.2.5. Example. For a mathematical pendulum of length l with mass m ,
 $V = -mgl \cos \theta$, so

$$H(\theta, p) = \frac{1}{2m}p^2 - mgl \cos \theta.$$

Sketch the phase curves of the pendulum for different energies and interpret them physically!



Calculus on Vector Spaces

In the rest of this term we will develop calculus for “functions of multiple variables”. This generally means functions defined on (a subset of) \mathbb{R}^n , but it is not any more difficult to treat functions defined on finite-dimensional vector spaces.

Throughout the following discussion, we assume that V and X denote finite-dimensional, normed vector spaces. The concrete norm will be irrelevant, as all norms are equivalent (see Theorem 2.1.11). We will consider first the derivative of a function

$$f: X \rightarrow V.$$

2.2.6. Definition. Let $f: X \rightarrow V_1$, $g: X \rightarrow V_2$ and $x_0 \in X$. We say that

$$f(x) = o(g(x)) \quad \text{as } x \rightarrow x_0 \quad \Leftrightarrow \quad \lim_{x \rightarrow x_0} \frac{\|f(x)\|_{V_1}}{\|g(x)\|_{V_2}} = 0$$

The Derivative of a Function

2.2.7. Definition. Let X, V be finite-dimensional vector spaces and $\Omega \subset X$ an open set. Then a map $f: \Omega \rightarrow V$ is called **differentiable at $x \in \Omega$** if there exists a linear map $L_x \in \mathcal{L}(X, V)$ such that

$$f(x + h) = f(x) + L_x h + o(h) \quad \text{as } h \rightarrow 0. \quad (2.2.1)$$

In this case we call L_x the **derivative of f at x** and write

$$L_x = Df|_x = df|_x.$$

We say that f is differentiable on Ω if it is differentiable for every $x \in \Omega$.

2.2.8. Remarks.

- ▶ Just as in the proof of 186 Lemma 3.1.2 we can show that the derivative is uniquely defined by (2.2.1).
- ▶ We may also copy the proof of Lemma 186 3.1.8 to see that every differentiable function is continuous.

The Derivative of a Function

If f is differentiable on Ω , we may regard Df as a map

$$Df: \Omega \rightarrow \mathcal{L}(X, V), \quad x \mapsto Df|_x.$$

2.2.9. Definition. We define

$$C(\Omega, V) := \{f: \Omega \rightarrow V : f \text{ is continuous}\},$$

$$C^1(\Omega, V) := \{f: \Omega \rightarrow V : f \text{ is differentiable and } Df \text{ is continuous}\}.$$

We may thus regard the **derivative** D as a (linear) map

$$D: C^1(\Omega, V) \rightarrow C(\Omega, \mathcal{L}(X, V)), \quad f \mapsto Df.$$

The Derivative of a Function

2.2.10. Example. Let X, V be finite-dimensional vector spaces and $L \in \mathcal{L}(X, V)$ a linear map. Then

$$L(x + h) = Lx + Lh \stackrel{!}{=} Lx + DL|_x h + o(h) \quad (h \rightarrow 0),$$

so the derivative of L at any $x \in X$ is $DL|_x = L$.

2.2.11. Examples. Explicit instances of Example 2.2.10 are, e.g.,

- ▶ Let $X = V = \mathbb{C}$ be regarded as real vector spaces and $f: z \rightarrow \bar{z}$ be the (then linear) complex conjugation. Then for $z, h \in \mathbb{C}$

$$\overline{z + h} = \bar{z} + \bar{h},$$

so $Df|_z(h) = \bar{h}$.

The Derivative of a Function

- ▶ Regard $A \in \text{Mat}(2 \times 2; \mathbb{R})$ as a linear map $\mathbb{R}^2 \rightarrow \mathbb{R}^2$. Then for $x, h \in \mathbb{R}^2$

$$A(x + h) = Ax + Ah,$$

so $DA|_x(h) = Ah$.

- ▶ Let $\text{tr}: \text{Mat}(n \times n; \mathbb{C}) \rightarrow \mathbb{C}$ be the **trace** of a square matrix, i.e.,

$$\text{tr } A = \text{tr}(a_{ij})_{1 \leq i,j \leq n} = \sum_{i=1}^n a_{ii}.$$

Then the trace is linear and for $A, H \in \text{Mat}(n \times n; \mathbb{C})$

$$D \text{tr}|_A H = \text{tr } H.$$

The Derivative of a Function

2.2.12. Example. Some examples of derivatives of non-linear maps are as follows:

- ▶ Let $X = V = \mathbb{C}$ be regarded as real vector spaces and $f: z \rightarrow z^2$.
Then for $z, h \in \mathbb{C}$

$$(z + h)^2 = z^2 + 2zh + h^2,$$

so $Df|_z(h) = 2zh$.

- ▶ Let $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ be given by

$$f(x) = f(x_1, x_2) = x_1 + 2x_2x_1 + x_2^2.$$

Then, for $h = (h_1, h_2) \in \mathbb{R}^2$ and $x \in \mathbb{R}^2$,

$$\begin{aligned}f(x + h) &= f(x_1 + h_1, x_2 + h_2) \\&= x_1 + h_1 + 2(x_2 + h_2)(x_1 + h_1) + (x_2 + h_2)^2 \\&= f(x) + h_1 + 2(h_2x_1 + h_1x_2 + h_2x_2) + 2h_1h_2 + h_2^2.\end{aligned}$$

The Derivative of a Function as a Matrix

In

$$f(x+h) = f(x) + \underbrace{h_1 + 2(h_2x_1 + h_1x_2 + h_2x_2)}_{=:L_{(x_1,x_2)}h} + 2h_1h_2 + h_2^2$$

the term $L_{(x_1,x_2)}h$ is clearly linear in h , while

$$\lim_{h \rightarrow 0} \frac{\|2h_1h_2\|_{\mathbb{R}}}{\|h\|_{\mathbb{R}^2}} = \lim_{h_1, h_2 \rightarrow 0} \frac{|2h_1h_2|}{\sqrt{h_1^2 + h_2^2}} = 2 \lim_{h_1, h_2 \rightarrow 0} \frac{|h_2|}{\sqrt{1 + (h_2/h_1)^2}}.$$

Since $|h_2| \rightarrow 0$ as $h_2 \rightarrow 0$ and $1/\sqrt{1 + (h_2/h_1)^2}$ is bounded, we see that

$$\lim_{h \rightarrow 0} \frac{\|2h_1h_2\|_{\mathbb{R}}}{\|h\|_{\mathbb{R}^2}} = 0,$$

and so $2h_1h_2 = o(h)$ as $h \rightarrow 0$. Similarly, we show that $h_2^2 = o(h)$, so we conclude

$$Df|_x h = (1 + 2x_2)h_1 + 2(x_1 + x_2)h_2.$$

The Derivative of a Function as a Matrix

Notice that we may express the derivative as a 1×2 matrix,

$$Df|_x h = (1 + 2x_2, 2(x_1 + x_2)) \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}.$$

This is of course not surprising; if $X = \mathbb{R}^n$ and $V = \mathbb{R}^m$, i.e., we are considering a function

$$f: \mathbb{R}^n \supset \Omega \rightarrow \mathbb{R}^m, \quad f(x_1, \dots, x_n) = \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_m(x_1, \dots, x_n) \end{pmatrix},$$

then its derivative at $x \in \Omega$ (if it exists) is

$$Df|_x \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \simeq \text{Mat}(m \times n; \mathbb{R}).$$

How to obtain this matrix? Denote by e_j the j th standard basis vector in \mathbb{R}^n or \mathbb{R}^m . We now consider the columns of $Df|_x$, which are given by $Df|_x e_j, j = 1, \dots, n$.

The Derivative of a Function as a Matrix

Assuming that f is differentiable, for any $h \in \mathbb{R}$, $x \in \mathbb{R}^n$ and $j = 1, \dots, n$ we have

$$f(x + he_j) = f(x) + Df|_x(he_j) + o(h),$$

which we may rewrite as

$$Df|_x e_j = \frac{1}{h}(f(x + he_j) - f(x)) + o(1) = \frac{1}{h} \sum_{k=1}^m (f_k(x + he_j) - f_k(x))e_k + o(1).$$

The (i, j) th element of $Df|_x$ is given by $\langle e_i, Df|_x e_j \rangle$, so

$$(Df|_x)_{ij} = \langle e_i, Df|_x e_j \rangle = \frac{1}{h}(f_i(x + he_j) - f_i(x)) + o(1).$$

We now take the limit $h \rightarrow 0$ to obtain

$$(Df|_x)_{ij} = \langle e_i, Df|_x e_j \rangle = \lim_{h \rightarrow 0} \frac{f_i(x + he_j) - f_i(x)}{h}.$$

Partial Derivatives

2.2.13. Definition. Let $\Omega \subset \mathbb{R}^n$ and $f: \Omega \rightarrow \mathbb{R}$ be differentiable on Ω . We then define the ***partial derivative with respect to x_j at $x \in \Omega$*** by

$$\begin{aligned}\frac{\partial f}{\partial x_j} \Big|_x &:= \lim_{h \rightarrow 0} \frac{f(x + he_j) - f(x)}{h} \\ &= \lim_{h \rightarrow 0} \frac{f(x_1, \dots, x_{j-1}, x_j + h, x_{j+1}, \dots, x_n) - f(x)}{h}\end{aligned}$$

In this notation,

$$(Df|_x)_{ij} = \frac{\partial f_i}{\partial x_j}$$

or rather

$$Df|_x = \left(\begin{array}{ccc} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{array} \right) \Bigg|_x$$

Partial Derivatives

There are several notations for the partial derivatives of a function. If $f: \mathbb{R}^n \rightarrow \mathbb{R}$, we may use any of the following:

$$\frac{\partial f}{\partial x_j} = \partial_{x_j} f = \partial_j f = f_{x_j} = f_j$$

to denote differentiation w.r.t. the variable x_j . In practice, we calculate the partial derivative w.r.t. to x_j by holding all other variables constant and simply differentiating f as a function of x_j .

2.2.14. Example. Let $f(x_1, x_2, x_3) = x_1 \sin(x_1 x_2 x_3) + 3x_2^2 x_1$. Then

$$\frac{\partial f}{\partial x_1} = \sin(x_1 x_2 x_3) + x_1 x_2 x_3 \cos(x_1 x_2 x_3) + 3x_2^2,$$

$$\frac{\partial f}{\partial x_2} = x_1^2 x_3 \cos(x_1 x_2 x_3) + 6x_2 x_1,$$

$$\frac{\partial f}{\partial x_3} = x_1^2 x_2 \cos(x_1 x_2 x_3).$$

The Jacobian

Of course, if $Df|_x$ exists, we may write it as a matrix of partial derivatives. However, it is not clear whether the existence of all partial derivatives implies the existence of the derivative $Df|_x$. Thus it is useful to consider the matrix of partial derivatives on its own; in fact, it deserves a special designation.

2.2.15. Definition. Let $\Omega \subset \mathbb{R}^n$ and $f: \Omega \rightarrow \mathbb{R}^m$. Assume that all partial derivatives $\frac{\partial f_i}{\partial x_j}$ of f exist at $x \in \Omega$. The matrix

$$J_f(x) := \left(\begin{array}{ccc} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{array} \right) \Bigg|_x$$

called the **Jacobian** of f .

If the derivative $Df|_x \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ exists, $J_f(x) \in \text{Mat}(m \times n; \mathbb{R})$ is the representing matrix of $Df|_x$ w.r.t. the standard bases in \mathbb{R}^n and \mathbb{R}^m .

The Jacobian

2.2.16. Example. Let $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be given by

$f(x_1, x_2) = (x_1^2 + x_2^2, x_2 - x_1)$. Then the partial derivatives are

$$\frac{\partial f_1}{\partial x_1} = \frac{\partial}{\partial x_1}(x_1^2 + x_2^2) = 2x_1,$$

$$\frac{\partial f_1}{\partial x_2} = \frac{\partial}{\partial x_2}(x_1^2 + x_2^2) = 2x_2,$$

$$\frac{\partial f_2}{\partial x_1} = \frac{\partial}{\partial x_1}(x_2 - x_1) = -1,$$

$$\frac{\partial f_2}{\partial x_2} = \frac{\partial}{\partial x_2}(x_2 - x_1) = 1.$$

The Jacobian is given by

$$J_f(x_1, x_2) = \begin{pmatrix} 2x_1 & 2x_2 \\ -1 & 1 \end{pmatrix}$$

The natural question that arises is, “Does the existence of $J_f(x)$ imply the differentiability of f at x ?”

The Jacobian

Regrettably, the answer to that question is negative, as the following example shows:

2.2.17. Example. Let $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ be given by

$$g(x_1, x_2) = \begin{cases} \frac{x_1 x_2}{x_1^2 + x_2^2} & (x_1, x_2) \neq (0, 0) \\ 0 & (x_1, x_2) = (0, 0) \end{cases}$$

Then all partial derivatives of g exist at $x = 0$, since

$$\left. \frac{\partial g}{\partial x_1} \right|_{x=0} = \lim_{h \rightarrow 0} \frac{g(0+h, 0) - g(0)}{h} = 0,$$

$$\left. \frac{\partial g}{\partial x_2} \right|_{x=0} = \lim_{h \rightarrow 0} \frac{g(0, 0+h) - g(0)}{h} = 0.$$

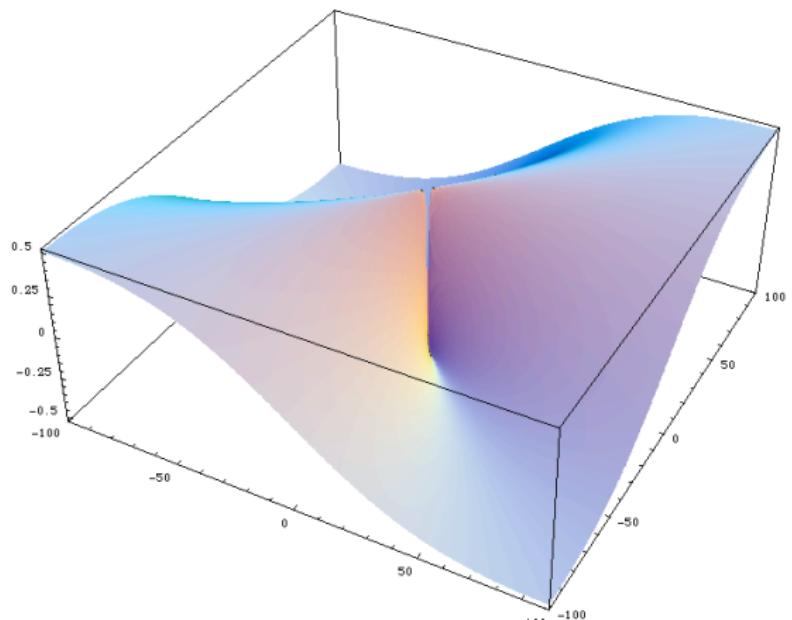
Thus both partial derivatives exist at $x = 0$ and in fact vanish.

The Jacobian

However, g is not even continuous at 0 since

$$\lim_{h \rightarrow 0} g(h, h) = \frac{h^2}{h^2 + h^2} = \frac{1}{2}, \quad \lim_{h \rightarrow 0} g(-h, h) = \frac{-h^2}{(-h)^2 + h^2} = -\frac{1}{2}.$$

Thus g can not be differentiable at $x = 0$.



The Jacobian

Thus the existence of the partial derivatives of a function $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ is not even enough to guarantee the continuity of f . However, we have the following result:

2.2.18. Theorem. Let $\Omega \subset \mathbb{R}^n$ be an open set and $f: \Omega \rightarrow \mathbb{R}^m$ such that all partial derivatives $\partial_{x_j} f_i$ exist on Ω .

- (i) If all partial derivatives are bounded (there exists a constant $M > 0$ such that $|\partial_{x_j} f_i| \leq M$ on Ω), then f is continuous i.e., $f \in C(\Omega, \mathbb{R}^m)$.
- (ii) If all partial derivatives are continuous on Ω , then f is continuously differentiable on Ω , i.e., $f \in C^1(\Omega, \mathbb{R}^m)$. In particular,

$$Df|_x = J_f(x) = \left(\begin{array}{ccc} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{array} \right) \Bigg|_x$$

for all $x \in \Omega$.

The Jacobian

Proof.

Let

$$f: \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad f(x) = \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_m(x_1, \dots, x_n) \end{pmatrix}.$$

For both statements of the theorem, we need to consider $f(x + h) - f(x)$.
To illustrate, let us first look at the case $n = 2$. Then

$$\begin{aligned} f_i(x + h) - f_i(x) &= f_i(x_1 + h_1, x_2 + h_2) - f_i(x_1, x_2) \\ &= [f_i(x_1 + h_1, x_2 + h_2) - f_i(x_1 + h_1, x_2)] \\ &\quad + [f_i(x_1 + h_1, x_2) - f_i(x_1, x_2)]. \end{aligned}$$

The Jacobian

Proof (continued).

For fixed h_1 , the first difference can be treated by the Mean Value Theorem 3.2.7 of Vv186: define

$$g: \mathbb{R} \rightarrow \mathbb{R}, \quad g(y) = f_i(x_1 + h_1, y).$$

Then there exists a $\theta_2 \in (x_2, x_2 + h_2)$ such that

$$\begin{aligned} & f_i(x_1 + h_1, x_2 + h_2) - f_i(x_1 + h_1, x_2) \\ &= g(x_2 + h_2) - g(x_2) \\ &= h_2 \cdot g'(\theta_2) = h_2 \partial_2 f_i(x_1 + h_1, x_2 + \tau_2 h_2) \end{aligned}$$

where we have chosen $\tau_2 \in (0, 1)$ such that $\theta_2 = x_2 + \tau_2 h_2$.

The Jacobian

Proof (continued).

Similarly, we find that

$$f_i(x_1 + h_1, x_2) - f_i(x_1, x_2) = h_1 \frac{\partial f_i}{\partial x_1}(x_1 + \tau_1 h_1, x_2)$$

for some $\tau_1 \in (0, 1)$. Generalizing to $n \geq 2$, we have constants $\tau_1, \dots, \tau_n \in (0, 1)$ such that

$$\begin{aligned} & f_i(x + h) - f_i(x) \\ &= f_i(x_1 + h_1, x_2 + h_2, \dots, x_n + h_n) - f_i(x_1, x_2 + h_2, \dots, x_n + h_n) \\ &\quad + f_i(x_1, x_2 + h_2, \dots, x_n + h_n) - f_i(x_1, x_2, x_3 + h_3, \dots, x_n + h_n) \\ &\quad + \cdots + f_i(x_1, x_2, \dots, x_{n-1}, x_n + h_n) - f_i(x_1, x_2, \dots, x_n) \\ &= h_1 \partial_1 f_i(x_1 + \tau_1 h_1, x_2 + h_2, \dots, x_n + h_n) \\ &\quad + h_2 \partial_2 f_i(x_1, x_2 + \tau_2 h_2, x_3 + h_3, \dots, x_n + h_n) \\ &\quad + \cdots + h_n \partial_n f_i(x_1, x_2, \dots, x_n + \tau_n h_n). \end{aligned}$$

The Jacobian

Proof (continued).

We proceed with the proof of the theorem.

- (i) Suppose that the partial derivatives are bounded. We want to prove that f is continuous at $x \in \Omega$, i.e.,

$$\lim_{h \rightarrow 0} f(x + h) = f(x)$$

where we are free to choose arbitrary norms in \mathbb{R}^n and \mathbb{R}^m for the convergence. In both spaces we choose the maximum norm $\|\cdot\|_\infty$ (see (2.1.4)):

$$\begin{aligned}\|f(x + h) - f(x)\|_\infty &= \max_{i=1,\dots,n} |f_i(x + h) - f_i(x)| \\ &\leq n \cdot \max_{j=1,\dots,n} |h_j| \max_{i,j=1,\dots,n} \sup_{x \in \Omega} \left| \frac{\partial f_i}{\partial x_j}(x) \right| \\ &\leq n \cdot M \cdot \|h\|_\infty \xrightarrow{h \rightarrow 0} 0.\end{aligned}$$

The Jacobian

Proof (continued).

(ii) Write

$$L = \left(\frac{\partial f_i}{\partial x_j} \right)_{\substack{i=1, \dots, m \\ j=1, \dots, n}} = (L_{ij})_{\substack{i=1, \dots, m \\ j=1, \dots, n}}$$

for the Jacobian. We want to show that

$$f(x + h) - f(x) - Lh = o(h) \quad \text{as } h \rightarrow 0.$$

We again choose the maximum norm $\|\cdot\|_\infty$ to establish the convergence and write

$$u^j = (x_1, \dots, x_{j-1}, x_j + \tau_j h_j, x_{j+1} + h_{j+1}, \dots, x_n + h_n).$$

for $j = 1, \dots, n$. We have the following estimate:

The Jacobian

Proof (continued).

$$\begin{aligned}\|f(x + h) - f(x) - Lh\|_\infty &= \max_{i=1,\dots,m} \left| f_i(x + h) - f_i(x) - \sum_{j=1}^n L_{ij} h_j \right| \\ &= \max_{i=1,\dots,m} \left| \sum_{j=1}^n h_j (\partial_j f_i(u^j) - \partial_j f_i(x)) \right| \\ &\leq \|h\|_\infty \underbrace{\sum_{j=1}^n \max_{i=1,\dots,m} |\partial_j f_i(u^j) - \partial_j f_i(x)|}_{\rightarrow 0 \text{ as } h \rightarrow 0} \\ &= o(h) \quad \text{as } h \rightarrow 0.\end{aligned}$$

Observe that we use the assumption that $\partial_j f_i(x)$ is continuous at x . This proves that f is differentiable, $L = Df|_x$ and $Df|_x$ depends continuously on x .



The Jacobian

2.2.19. Remark. Let $\Omega \subset \mathbb{R}^n$ be an open set. Then

$$C^1(\Omega, \mathbb{R}^m) = \{f: \Omega \rightarrow \mathbb{R}^m: \partial_j f_i \text{ is continuous for } j = 1, \dots, n \\ \text{and } i = 1, \dots, m\}.$$

If $m = 1$, we write $C^1(\Omega) := C^1(\Omega, \mathbb{R})$ for short.

We will next establish the product and chain rules for differentiation.

Generalized Products

To avoid having to re-prove the product rule for various types of products that we will encounter, we first define a generalized product through precisely those properties that we shall need.

2.2.20. **Definition.** Let X_1, X_2, V be normed vector spaces. A map

$\odot: X_1 \times X_2 \rightarrow V$ is called a **(generalized) product** if

1. \odot is bilinear, i.e., linear in each entry and
2. $\|u \odot v\|_V \leq \|u\|_{X_1} \|v\|_{X_2}$ for all $u \in X_1, v \in X_2$.

2.2.21. **Examples.**

1. The scalar product in \mathbb{R}^n ;
2. The cross product $\times: \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$;
3. For a compact non-empty set $K \subset \mathbb{R}^n$ and $f, g \in C(K, \mathbb{R})$ the pointwise product $f \cdot g \in C(K, \mathbb{R})$, defined by

$$(f \cdot g)(x) = f(x)g(x)$$

The Product Rule

2.2.22. **Product Rule.** Let U, X_1, X_2, V be finite-dimensional vector spaces and $\Omega \subset U$ an open set. Let $f: \Omega \rightarrow X_1$ and $g: \Omega \rightarrow X_2$ be differentiable maps and $\odot: X_1 \times X_2 \rightarrow V$ a generalized product. Then $f \odot g: \Omega \rightarrow V$ is also differentiable and

$$D(f \odot g) = (Df) \odot g + f \odot (Dg). \quad (2.2.2)$$

At $x \in \Omega$ the right-hand side is interpreted as a linear map $U \rightarrow V$

$$u \mapsto D(f \odot g)|_x u = (Df|_x u) \odot g(x) + f(x) \odot (Dg|_x u). \quad (2.2.3)$$

The Product Rule

Proof.

The proof is similar to that for the product rule for functions of one variable. We telescope the difference,

$$\begin{aligned} & f(x+h) \odot g(x+h) - f(x) \odot g(x) \\ &= f(x+h) \odot (g(x+h) - g(x)) + (f(x+h) - f(x)) \odot g(x) \\ &= (f(x) + O(h)) \odot (Dg|_x h + o(h)) + (Df|_x h + o(h)) \odot g(x) \end{aligned}$$

as $h \rightarrow 0$. Extending the relevant limit theorems from the pointwise product to the generalized product, we have

$$\begin{aligned} & f(x+h) \odot g(x+h) - f(x) \odot g(x) \\ &= f(x) \odot (Dg|_x h) + O(\|h\|^2) + o(h) + (Df|_x h) \odot g(x) + o(h) \\ &= f(x) \odot (Dg|_x h) + (Df|_x h) \odot g(x) + o(h) \end{aligned}$$

□

The Chain Rule

2.2.23. **Chain Rule.** Let U, X, V be finite-dimensional vector spaces and $\Omega \subset U$, $\Sigma \subset X$ open sets. Let $g: \Omega \rightarrow \Sigma$ and $f: \Sigma \rightarrow V$ be differentiable maps. Then the composition $f \circ g: \Omega \rightarrow V$ is also differentiable and for all $x \in \Omega$

$$D(f \circ g)|_x = Df|_{g(x)} \circ Dg|_x, \quad (2.2.4)$$

where the right-hand side is a composition of linear maps.

The proof is basically identical to that of 186 Theorem 3.1.12, the chain rule for functions of one real variable. You are encouraged to revisit that proof and apply it to the general chain rule here.

2.2.24. **Example.** Consider the polar coordinates $(r, \phi) \in (0, \infty) \times [0, 2\pi]$, defined through the map

$$\Phi(r, \phi) = \begin{pmatrix} r \cos \phi \\ r \sin \phi \end{pmatrix}.$$

The Chain Rule

Then

$$D\Phi|_{(r,\phi)} = \begin{pmatrix} \frac{\partial\Phi_1}{\partial r} & \frac{\partial\Phi_1}{\partial\phi} \\ \frac{\partial\Phi_2}{\partial r} & \frac{\partial\Phi_2}{\partial\phi} \end{pmatrix} = \begin{pmatrix} \cos\phi & -r\sin\phi \\ \sin\phi & r\cos\phi \end{pmatrix}.$$

Next, consider the map $U: \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x_1, x_2) \mapsto x_1^2 + x_2^2$. The derivative is

$$DU|_x = \left(\frac{\partial U}{\partial x_1}, \frac{\partial U}{\partial x_2} \right) = (2x_1, 2x_2)$$

Now $U \circ \Phi = (r\cos\phi)^2 + (r\sin\phi)^2 = r^2$. Clearly, $D(U \circ \Phi)|_{(r,\phi)} = (2r, 0)$. We can also apply the chain rule:

$$\begin{aligned} D(U \circ \Phi)|_{(r,\phi)} &= DU|_{(r\cos\phi, r\sin\phi)} D\Phi|_{(r,\phi)} \\ &= (2r\cos\phi, 2r\sin\phi) \begin{pmatrix} \cos\phi & -r\sin\phi \\ \sin\phi & r\cos\phi \end{pmatrix} \\ &= (2r\cos^2\phi + 2r\sin^2\phi, -2r^2\cos\phi\sin\phi + 2r^2\sin\phi\cos\phi) \\ &= (2r, 0) \end{aligned}$$

Integrals of Vector-Space-Valued Functions

The following important result will require the integral of a function of a single variable, albeit with values in a vector space V . In other words, we need to assign a meaning to

$$\int_a^b f(x) dx,$$
 where $f: [a, b] \rightarrow V.$

Fortunately, the procedure is completely analogous to that of functions $f: \mathbb{R} \rightarrow \mathbb{R}$, at least for the regulated integral: we define step functions on $[a, b]$ with respect to a partition P by setting them constant on sub-intervals of the partition:

2.2.25. Definition. Let V be a real or complex vector space. A function $f: [a, b] \rightarrow V$ is called a **step function with respect to a partition** $P = (a_0, \dots, a_n)$ if there exist elements $y_i \in V$ such that $f(t) = y_i$ whenever $a_{i-1} < t < a_i$, $i = 1, \dots, n$. We denote the set of all step functions by $\text{Step}([a, b], V)$.

Step Functions

2.2.26. Example. The map

$$f: [0, 1] \rightarrow \mathbb{R}^2, \quad f(x) = \begin{cases} \begin{pmatrix} 0 \\ 1/2 \end{pmatrix} & 0 \leq x < 1/2 \\ \begin{pmatrix} 1 \\ 1 \end{pmatrix} & x = 1/2 \\ \begin{pmatrix} 2 \\ 0 \end{pmatrix} & 1/2 < x \leq 1 \end{cases}$$

is a step function.

We then follow through with analogous definitions to the ones for real functions, replacing the modulus in \mathbb{R} by the norm in a vector space:

2.2.27. Definition. Let $I \subset \mathbb{R}$ be an interval and $(V, \|\cdot\|_V)$ a normed vector space. We say that a function $f: I \rightarrow V$ is **bounded** if

$$\|f\|_\infty := \sup_{x \in I} \|f(x)\|_V < \infty. \quad (2.2.5)$$

The set of all bounded functions $f: I \rightarrow V$ is denoted $L^\infty(I, V)$.

Bounded Functions

2.2.28. Example. The map

$$f: \mathbb{R} \rightarrow \mathbb{R}^2, \quad f(t) = \begin{pmatrix} \sin t \\ e^{-t^2} \end{pmatrix}$$

is a bounded map. To see this, we endow \mathbb{R}^2 with the norm

$\|x\|_1 := |x_1| + |x_2|$. (Since all norms in \mathbb{R}^n are equivalent, it doesn't matter which norm we take.) Then

$$\begin{aligned} \|f\|_\infty &:= \sup_{t \in \mathbb{R}} \|f(t)\|_1 = \sup_{t \in \mathbb{R}} (|\sin t| + |e^{-t^2}|) \\ &\leq \sup_{t \in \mathbb{R}} |\sin t| + \sup_{t \in \mathbb{R}} |e^{-t^2}| = 2 < \infty. \end{aligned}$$

Integrals of Step Functions

We then define the integral of a step function as before:

2.2.29. Theorem. Let $f: [a, b] \rightarrow V$ be a step function with respect to some partition P . Then

$$I_P(f) := (a_1 - a_0)y_1 + \cdots + (a_n - a_{n-1})y_n \in V$$

is independent of the choice of the partition P and is called the **integral** of f .

Note that if $f: [a, b] \rightarrow V$, then $\int_a^b f(x) dx \in V$, the integral of f is an element of the vector space V .

(This makes it impossible to define the Riemann integral for functions $f: I \rightarrow V$, because it relies comparing the size of upper and lower step functions.)

Integrals of Step Functions

The main ingredient is again uniform convergence, where we now say that a sequence of functions (f_n) , $f_n: I \rightarrow V$, $I \subset \mathbb{R}$, converges uniformly to $f: I \rightarrow V$ in a normed vector space $(V, \|\cdot\|_V)$ if

$$\|f_n - f\|_{\infty} := \sup_{x \in I} \|f_n(x) - f(x)\|_V \xrightarrow{n \rightarrow \infty} 0.$$

A function f is then said to be **regulated** if it is the uniform limit of a sequence of step functions. We can then define the integral of f as the limit of the integrals of these step functions.

You are invited to check that everything in fact works just as in the regulated integral for scalar real functions!

Integrals of Step Functions

The upshot is the following: if $f: [a, b] \rightarrow \mathbb{R}^n$ is piecewise continuous, then f is regulated and

$$\int_a^b f(x) dx = \int_a^b \begin{pmatrix} f_1(x) \\ \vdots \\ f_n(x) \end{pmatrix} dx = \begin{pmatrix} \int_a^b f_1(x) dx \\ \vdots \\ \int_a^b f_n(x) dx \end{pmatrix}$$

(This follows because a sequence of step functions converging uniformly to f will converge uniformly in each component; the individual components are then equal to the “usual” regulated integrals of real-valued functions.)

Furthermore, we have the standard estimate

$$\left\| \int_a^b f(x) dx \right\|_V \leq \int_a^b \|f(x)\|_V dx \leq |b - a| \cdot \sup_{x \in [a, b]} \|f(x)\|_V.$$

The Mean Value Theorem

In general, the mean value theorem (186 Theorem 3.2.7) that we proved last term does not survive here. However, we can formulate an “integrated version:”

2.2.30. Mean Value Theorem. Let X, V be finite-dimensional vector spaces, $\Omega \subset X$ open and $f \in C^1(\Omega, V)$. Let $x, y \in \Omega$ and assume that the line segment $x + ty$, $0 \leq t \leq 1$, is wholly contained in Ω . Then

$$f(x + y) - f(x) = \int_0^1 Df|_{x+ty} y \, dt = \left(\int_0^1 Df|_{x+ty} \, dt \right) y. \quad (2.2.6)$$

2.2.31. Remark. The integrals in (2.2.6) are integrals of elements of V (the integrand $Df|_{x+ty} y$) and $\mathcal{L}(X, V)$ (the integrand $Df|_{x+ty}$). The definition of the regulated integral in Vv186 can easily be extended to such vector-space valued functions. However, in the case of $V = \mathbb{R}^m$ and $\mathcal{L}(X, V) = \text{Mat}(m \times n; \mathbb{R})$ the vectors and matrices are simply integrated in each component to yield again vectors and matrices, respectively.

The Mean Value Theorem

The Mean Value Theorem 2.2.30 can also be understood as a generalization of the fundamental theorem of calculus: For single-variable functions, the fundamental theorem of calculus can be expressed as

$$f(x+y) - f(x) = \int_x^{x+y} f'(\xi) d\xi.$$

Substituting $t = (\xi - x)/y$ in the integral, we have the equivalent identity

$$f(x+y) - f(x) = \int_0^1 f'(x+yt)y dt,$$

a special case of (2.2.6).

The Mean Value Theorem

Proof of Theorem 2.2.30.

Define the auxiliary function $g \in C^1([0, 1], V)$ by $g(t) := f(x + ty)$. Thus (by 186 Lemma 4.2.3) we have

$$f(x + y) - f(x) = g(1) - g(0) = \int_0^1 g'(t) dt.$$

For $\gamma(t) = x + ty$ we have $\gamma'(t) = y$. Applying the chain rule,

$$g'(t) = D(f \circ \gamma)|_t = Df|_{\gamma(t)} D\gamma|_t = Df|_{x+ty} y.$$

Thus we obtain

$$f(x + y) - f(x) = \int_0^1 Df|_{x+ty} y dt,$$

proving the first equality.

The Mean Value Theorem

Proof of Theorem 2.2.30 (continued).

We now prove that y may be “taken out” of the integral. Let us abbreviate $L(t) = Df|_{x+ty}$. For $z \in (0, 1)$ we have

$$\frac{d}{dz} \int_0^z L(t)y \, dt = L(z)y = \frac{d}{dz} \left\{ \left(\int_0^z L(t) \, dt \right) y \right\}.$$

Furthermore, setting $z = 0$ we have

$$\int_0^0 L(t)y \, dt = 0 = \left(\int_0^0 L(t) \, dt \right) y.$$

Therefore,

$$\int_0^z L(t)y \, dt = \left(\int_0^z L(t) \, dt \right) y$$

for all $z \in [0, 1]$, in particular also for $z = 1$.



The Mean Value Theorem

2.2.32. Corollary. From the standard estimate

$$\left\| \int_a^b f(t) dt \right\|_V \leq |b - a| \cdot \sup_{t \in [a,b]} \|f(t)\|_V$$

Theorem 2.2.30 yields

$$\|f(x + y) - f(x)\|_V \leq \|y\|_X \cdot \sup_{0 \leq t \leq 1} \|Df|_{x+ty}\|,$$

where $\|Df|_{x+ty}\|$ denotes the operator norm of $Df|_{x+ty} \in \mathcal{L}(X, V)$.

Differentiating Under an Integral

We close this section with a useful result concerning the interchanging of differentiation and integration.

2.2.33. Theorem. Let X, V be finite-dimensional vector spaces, $I = [a, b] \subset \mathbb{R}$ an interval and $\Omega \subset X$ an open set. Let $f: I \times \Omega \rightarrow V$ be a continuous function such that $Df(t, \cdot)|_x$ exists and is continuous at every $(t, x) \in I \times \Omega$. Then

$$g(x) = \int_a^b f(t, x) dt$$

is differentiable in Ω and

$$Dg(x) = \int_a^b Df(t, \cdot)|_x dt$$

Differentiating Under an Integral

Proof.

Fix $x \in \Omega$ and choose h small enough such that $x + h \in \Omega$. In any case, we assume $\|h\| < 1$. We need to prove

$$g(x + h) - g(x) - \underbrace{\left(\int_a^b Df(t, \cdot)|_x dt \right) h}_{=:L} = o(h). \quad (2.2.7)$$

By the Mean Value Theorem, the left-hand side equals

$$\begin{aligned} & \int_a^b (f(t, x + h) - f(t, x) - Df(t, \cdot)|_x h) dt \\ &= \int_a^b \left(\int_0^1 Df(t, \cdot)|_{x+sh} ds h - Df(t, \cdot)|_x h \right) dt \\ &= \int_a^b \left(\int_0^1 (Df(t, \cdot)|_{x+sh} - Df(t, \cdot)|_x) ds \right) h dt. \end{aligned}$$

Differentiating Under an Integral

Proof (continued).

Taking the norm, we have

$$\begin{aligned} & \|g(x + h) - g(x) - Lh\|_V \\ & \leq (b - a) \sup_{t \in [a, b]} \left\| \int_0^1 (Df(t, \cdot)|_{x+sh} - Df(t, \cdot)|_x) ds \right\| \cdot \|h\|_X \\ & \leq (b - a) \sup_{t \in [a, b]} \sup_{s \in [0, 1]} \|Df(t, \cdot)|_{x+sh} - Df(t, \cdot)|_x\| \cdot \|h\|_X \end{aligned}$$

Here $\|\cdot\|$ denotes the operator norm. We now need to show that

$$\sup_{t \in [a, b]} \sup_{s \in [0, 1]} \|Df(t, \cdot)|_{x+sh} - Df(t, \cdot)|_x\|$$

vanishes when $h \rightarrow 0$. However, this requires a subtle argument, because s and t are free to vary independent of h

Differentiating Under an Integral

Proof (continued).

Consider the function $Df(t, \cdot)|_y$ where y varies in the closed and bounded set $\overline{B_1(x)}$. Since V is finite-dimensional, this set is compact and so is $[a, b] \times \overline{B_1(x)}$.

Since $Df(t, \cdot)|_y$ is continuous in the compact set $[a, b] \times \overline{B_1(x)}$, it is also uniformly continuous by Theorem 2.1.38. Hence, if $\varepsilon > 0$ is given we can find $\delta > 0$ such that

$$\|h\|_X < \delta \quad \Rightarrow \quad \|Df(t, \cdot)|_{x+sh} - Df(t, \cdot)|_x\| < \varepsilon,$$

independent of s and t . This implies that

$$\|h\|_X < \delta \quad \Rightarrow \quad \frac{\|g(x+h) - g(x) - Lh\|_V}{\|h\|_X} < (b-a)\varepsilon,$$

proving (2.2.7). □

Differentiating Under an Integral

Differentiating an integral with respect to a parameter can be very useful for calculating integrals that are otherwise difficult to evaluate directly. For example, by differentiating

$$g(x) := \int_0^{\infty} \frac{\sin t}{t} e^{-xt} dt$$

with respect to x you will show in the assignments that

$$g(0) = \int_0^{\infty} \frac{\sin t}{t} dt = \frac{\pi}{2}.$$

(Compare with the discussion of the Dirichlet integral last term, see 186 Example 4.2.14.)

Convergence and Continuity

Functions and Derivatives

Curves in Vector Spaces

Potential Functions

The Second Derivative

Extrema of Potential Functions

Constrained Extrema

Curves

An important case of a map between vector space is a map $\mathbb{R} \rightarrow V$, where $(V, \|\cdot\|)$ is a normed vector space.

2.3.1. Definition. Let V be a finite-dimensional vector space and $I \subset \mathbb{R}$ an interval.

- ▶ A set $\mathcal{C} \subset V$ for which there exists a continuous, surjective and locally injective map $\gamma: I \rightarrow \mathcal{C}$ is called a **curve**.
- ▶ The map γ is called a **parametrization** of \mathcal{C} .
- ▶ A curve \mathcal{C} together with a parametrization is called a **parametrized curve**.

2.3.2. Remark. Here **locally injective** means that in the neighborhood $B_\varepsilon(x) \cap I$ of any point $x \in I$ the parametrization is injective.

More generally, let f be a function of a single real variable. We say that a property holds **locally at a point** $p \in \mathbb{R}$ if this property holds in some ε -neighborhood $B_\varepsilon(p)$.

Local Properties of Real Functions

2.3.3. Examples.

- ▶ If $f \in C(\mathbb{R})$ and $f(0) > 0$, then f is **locally positive at 0**.
- ▶ If $f \in C^1(\mathbb{R})$ and $f'(0) > 0$, then f is **locally increasing at 0**.
- ▶ If $f \in C^1(\mathbb{R})$ and $f'(0) \neq 0$, then f is **locally invertible at 0**.

We simply say that a property holds **locally** if this property holds locally at every point $p \in \mathbb{R}$.

2.3.4. Example.

The sequence of functions (f_n) given by

$f_n(x) = (x - 1/n)^2$ converges to $f(x) = x^2$ locally uniformly, because at every point $p \in \mathbb{R}$ there is an ε -neighborhood such that the convergence is uniform in this neighborhood.

Curves

2.3.5. Example. The set

$$S^1 := \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1\}$$

is a curve in \mathbb{R}^2 because we can find a parametrization, e.g.,

$$\gamma: [0, 2\pi] \rightarrow S^1, \quad \gamma(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}.$$

It is clear that γ is continuous. Furthermore, $\text{ran } \gamma \subset S^1$ since

$$\cos^2 t + \sin^2 t = 1 \quad \text{for all } t \in [0, 2\pi].$$

The map γ is not injective, since $\gamma(0) = \gamma(2\pi) = (1, 0)$, but it is injective on $(0, 2\pi)$: If $\gamma(t_1) = \gamma(t_2) \neq (1, 0)$, then $\cos t_1 = \cos t_2$ and $\sin t_1 = \sin t_2$. The second equation implies $t_1, t_2 \in (0, \pi]$ or $t_1, t_2 \in (\pi, 2\pi)$. However, since the cosine function is injective on $(0, \pi]$ and on $(\pi, 2\pi)$ we obtain $t_1 = t_2$. Hence γ is locally injective.

Curves

Now suppose $(x_1, x_2) \in S^1$ is given. Then taking

$$t_0 = \begin{cases} \arctan \frac{x_2}{x_1} & \text{if } x_2 > 0, x_1 \neq 0, \\ \pi/2 & \text{if } (x_1, x_2) = (0, 1), \\ \pi + \arctan \frac{x_2}{x_1} & \text{if } x_2 < 0, x_1 \neq 0, \\ 3\pi/2 & \text{if } (x_1, x_2) = (0, -1), \end{cases} \quad (2.3.1)$$

(where we use a suitable branch of the inverse tangent) gives $\gamma(t_0) = (x_1, x_2)$ for some $t_0 \in [0, 2\pi]$. Thus, γ is surjective and therefore a parametrization.

Another parametrization is

$$\tilde{\gamma}: [0, 1] \rightarrow S^1, \quad \tilde{\gamma}(t) = \begin{pmatrix} \cos(2\pi t) \\ -\sin(2\pi t) \end{pmatrix}.$$

Both (\mathcal{C}, γ) and $(\mathcal{C}, \tilde{\gamma})$ are parametrized curves.

Parametrizations of Curves

It is clear that a curve will have an infinite number of parametrizations. Physically, a curve \mathcal{C} might be considered to be the path of a particle, while the parametrization γ gives the position of the particle at each time t . Hence, in Example 2.3.5, γ describes the counter-clockwise movement of a particle around the unit circle, while $\tilde{\gamma}$ describes a clockwise movement around the same path. The parametrization γ corresponds to completing the path in time 2π , while $\tilde{\gamma}$ corresponds to completing the path in time 1. Hence, the parametrization $\tilde{\gamma}$ can be said to correspond to a greater velocity of the particle.

Simple, Open and Closed Curves

2.3.6. Definition. Let $\mathcal{C} \subset V$ be a curve possessing a parametrization $\gamma: I \rightarrow \mathcal{C}$ with $\text{int } I = (a, b)$ for $-\infty \leq a < b \leq \infty$.

- (i) If γ is (globally) injective parametrization we say that \mathcal{C} is a **simple curve**.
- (ii) If

$$\lim_{t \rightarrow a} \gamma(t) = \lim_{t \rightarrow b} \gamma(t),$$

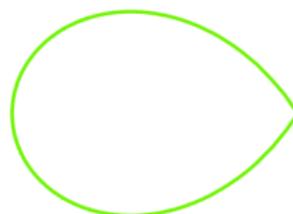
the curve \mathcal{C} is said to be **closed**.

- (iii) If a curve is not closed, it is said to be **open**. The points

$$x := \lim_{t \rightarrow a} \gamma(t) \quad \text{and} \quad y := \lim_{t \rightarrow b} \gamma(t)$$

are called the **initial point** and the **final point** of the parametrized curve (\mathcal{C}, γ) . The open curve is said to join x and y .

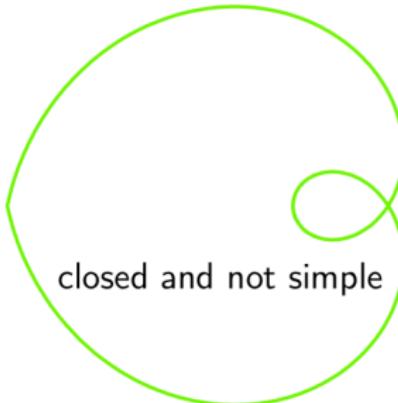
Sketches of Simple, Open and Closed Curves



closed and simple



simple



closed and not simple



not simple

Initial and Final Points of Open Curves

2.3.7. Remark. Whether a point is an initial or final point of an open curve depends on the parametrization. We will explore this a little further.

2.3.8. Example. The simple open curve

$$\mathcal{C} = \{(x_1, x_2) \in \mathbb{R}^2 : 0 \leq x_1 \leq 1, x_2 = x_1^2\}$$

joins the points $x = (0, 0)$ and $y = (1, 1)$. Either may be considered the initial point or the final point. Possible parametrizations are

$$\gamma(t) = \begin{pmatrix} t \\ t^2 \end{pmatrix}, \quad \tilde{\gamma}(t) = \begin{pmatrix} 1-t \\ (1-t)^2 \end{pmatrix}$$

where both $\gamma, \tilde{\gamma}: [0, 1] \rightarrow \mathcal{C}$.

Reparametrization of Curves

2.3.9. Definition. Let $\mathcal{C} \subset V$ be a curve with parametrization $\gamma: I \rightarrow \mathcal{C}$.

- (i) Let $J \subset \mathbb{R}$ be an interval. A continuous, bijective map $r: J \rightarrow I$ is called a **reparametrization** of the parametrized curve (\mathcal{C}, γ) .
- (ii) If r is **increasing** the reparametrization is said to be **orientation-preserving**.
- (iii) If r is **decreasing** the reparametrization is said to be **orientation-reversing**.

2.3.10. Remarks.

- (i) Given any two parametrizations $\gamma, \tilde{\gamma}$ of an open curve, one can always find a reparametrization by setting $r = \gamma^{-1} \circ \tilde{\gamma}$ (the continuity and local injectivity is enough for this definition to make sense).
- (ii) Since every bijective map in \mathbb{R} is either decreasing or increasing (see 186 Theorem 2.5.20), it follows that a reparametrization is either orientation-preserving or orientation-reversing.

Reparametrization of Curves

2.3.11. Example. Consider the unit circle S^1 of Example 2.3.5 with parametrizations

$$\gamma: [0, 2\pi] \rightarrow S^1, \quad \gamma(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix},$$

$$\tilde{\gamma}: [0, 1] \rightarrow S^1, \quad \tilde{\gamma}(t) = \begin{pmatrix} \cos(2\pi t) \\ -\sin(2\pi t) \end{pmatrix}.$$

Then $r: [0, 1] \rightarrow [0, 2\pi]$, $r(t) = -2\pi t$, is a reparametrization of the parametrized curve (\mathcal{C}, γ) . In fact,

$$\tilde{\gamma} = \gamma \circ r.$$

The reparametrization is not orientation-preserving since $r'(t) = -2\pi < 0$.

Orientation of Curves

A reparametrization of a parametrized curve (\mathcal{C}, γ) yields a new parametrized curve $(\mathcal{C}, \tilde{\gamma})$ where $\tilde{\gamma} = \gamma \circ r$.

It is easy to see that an orientation-preserving reparametrization of an open curve (\mathcal{C}, γ) yields an open parametrized curve $(\mathcal{C}, \tilde{\gamma})$ with the same initial and final points.

2.3.12. Definition. Let (\mathcal{C}, γ) be a parametrized curve and r a reparametrization of (\mathcal{C}, γ) .

The curve $(\mathcal{C}, \tilde{\gamma})$ with $\tilde{\gamma} = \gamma \circ r$ is said to have the **same orientation** as (\mathcal{C}, γ) if r is orientation-preserving. Otherwise it is said to have **reverse orientation**.

2.3.13. Remark. The orientation of an open curve can be fixed by selecting the initial and final points. The orientation of a closed curve can be fixed by splitting the curve into two disjoint simple curves and fixing appropriate orientations for them.

Orientation of Curves

Hence a curve can have two possible orientations. If we want to fix a curve \mathcal{C} together with an orientation (but not necessarily a concrete parametrization), we denote it by \mathcal{C}^* and if necessary give a single parametrization γ so that (\mathcal{C}, γ) has the desired orientation. The same curve with opposite orientation is denoted by $-\mathcal{C}^*$. (This will be quite important when we discuss integration later.)

There is in general no natural way to select a “proper” or positive orientation of a curve; rather both possible orientations have equal validity. There is a single exception, however:

2.3.14. Definition. Let (\mathcal{C}, γ) be a parametrized, simple, closed curve in \mathbb{R}^2 . Then \mathcal{C} is said to have **positive orientation** if γ traverses \mathcal{C} in a **counter-clockwise** direction.

Curves in Polar Coordinates

When we previously introduced polar coordinates in \mathbb{C} , we remarked that there is a one-to-one correspondence

$$\mathbb{C} \setminus \{0\} \ni x + iy \leftrightarrow (r, \varphi) \in \mathbb{R}_+ \times [0, 2\pi)$$

We want to adapt this to \mathbb{R}^2 instead of \mathbb{C} , i.e., associate an angle φ and a distance r to every point $(x_1, x_2) \in \mathbb{R}^2$.

One of the main difficulties stems from the fact that we can not associate an angle φ to $x = 0$. However, if we do not focus on associating an angle φ to every point $x \in \mathbb{R}^2$, but only on finding a cartesian point $(x_1, x_2) \in \mathbb{R}^2$ given (r, φ) , we can be a bit more flexible.

Curves in Polar Coordinates

We will allow $(r, \varphi) \in \mathbb{R}^2$, and associate to them a point $x \in \mathbb{R}^2$ as follows:

$$x = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix}$$

Of course this association is not injective, but this will not matter for our present purposes. We consider a particular type of curve, defined through the map

$$\gamma(t) = \begin{pmatrix} f(t) \cos t \\ f(t) \sin t \end{pmatrix}, \quad (2.3.2)$$

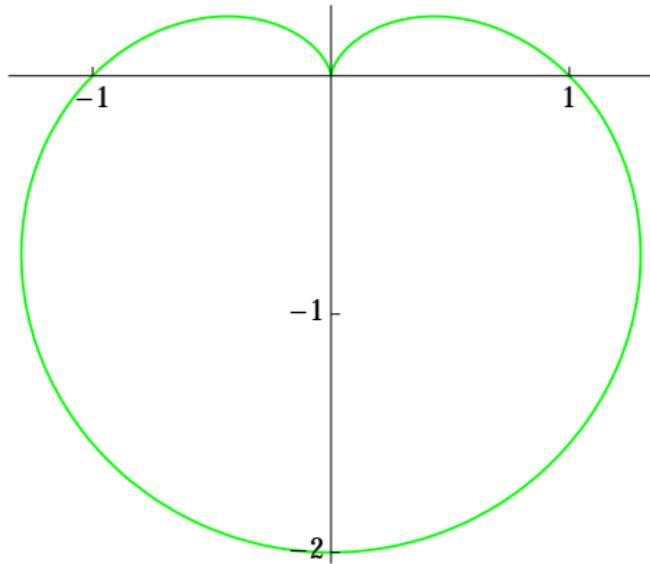
where $f: \mathbb{R} \rightarrow \mathbb{R}$ is some function. For short, such a curve is sometimes written as

$$r = f(t) \quad (2.3.3)$$

A curve given (2.3.3) is known as a **curve in polar coordinates**. The equation (2.3.3) is to be interpreted in the sense of (2.3.2).

Curves in Polar Coordinates

2.3.15. Example. The **cardioid** is given by $r = 1 - \sin t$:



Smooth Curves

2.3.16. Definition. A curve $\mathcal{C} \subset V$ is said to be **smooth** if there exists a parametrization $\gamma: I \rightarrow \mathcal{C}$ such that

- (i) γ is continuously differentiable on $\text{int } I$ and
- (ii) $D\gamma|_t \neq 0$ for all $t \in \text{int } I$.

A **smooth reparametrization** is a reparametrization that is continuously differentiable with non-vanishing derivative in the interior of its domain.

If $V = \mathbb{R}^n$, the Jacobian $D\gamma = \gamma'$ of a smooth curve $\gamma: I \rightarrow \mathbb{R}^n$ is given by

$$\gamma'(t) = \begin{pmatrix} \gamma'_1(t) \\ \vdots \\ \gamma'_n(t) \end{pmatrix}, \quad t \in \text{int } I.$$

Graphs of Functions as Curves

Let us consider the case of the graph Γ of a function $f: I \rightarrow \mathbb{R}$, $I \subset \mathbb{R}$ an interval: it is defined as the set

$$\Gamma = \{(x, y) \in \mathbb{R}^2 : x \in I, y = f(x)\}.$$

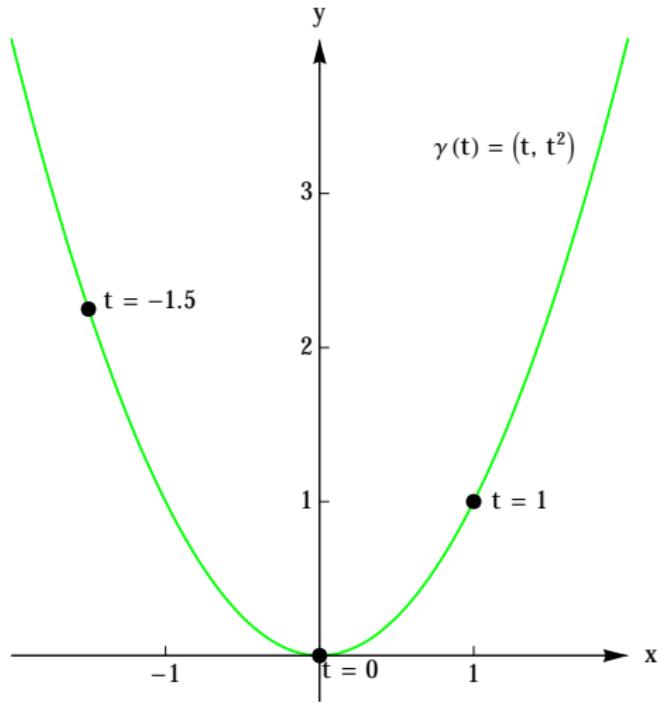
This set can be regarded as a curve with parametrization

$$\gamma: I \rightarrow \mathbb{R}^2, \quad t \mapsto \begin{pmatrix} t \\ f(t) \end{pmatrix}.$$

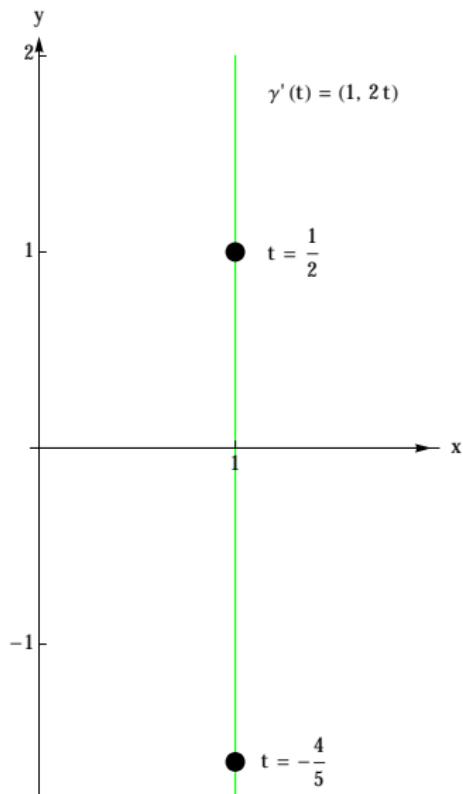
2.3.17. Example. Consider the function $f: \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x^2$. Its graph is just the curve parametrized by

$$\gamma(t) = \begin{pmatrix} \gamma_1(t) \\ \gamma_2(t) \end{pmatrix} = \begin{pmatrix} t \\ t^2 \end{pmatrix}.$$

Curves and Graphs of Functions



Curves and Graphs of Functions



By our previous considerations, γ is differentiable and

$$\gamma'(t) = \begin{pmatrix} \gamma'_1(t) \\ \gamma'_2(t) \end{pmatrix} = \begin{pmatrix} 1 \\ 2t \end{pmatrix}.$$

The graph of γ' is quite unspectacular.

Tangent Lines of Curves

So what is the interpretation of γ' ? If $\gamma = (t, f(t))$ parametrizes the graph function of some function f , then $\gamma'(t) = (1, f'(t))$. Recall that the derivative satisfies

$$\gamma(t_0 + t) = \gamma(t_0) + \gamma'(t_0)t + o(t),$$

so $\gamma(t_0) + \gamma'(t_0)t$ is a linear approximation to the parametrization γ at a point t_0 .

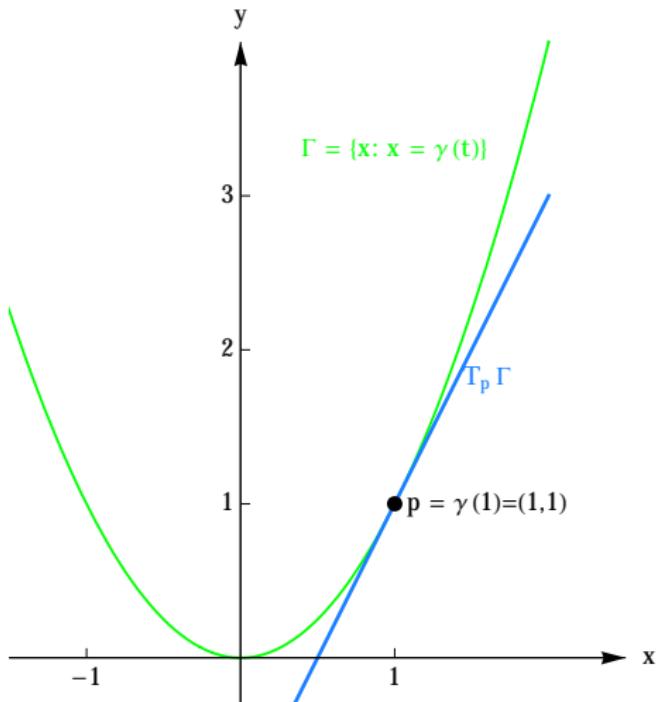
In fact, if $\mathcal{C} \subset \mathbb{R}^n$ is a curve and $p = \gamma(t_0) \in \mathcal{C}$, then

$$T_p \mathcal{C} = \{x \in \mathbb{R}^n : x = \gamma(t_0) + \gamma'(t_0)t : t \in \mathbb{R}\}$$

gives the **tangent line** to Γ at p .

Tangent Lines of Curves

Continuing from Example 2.3.17, we have the following tangent line $T_p \Gamma$ for $p = (1, 1)$:

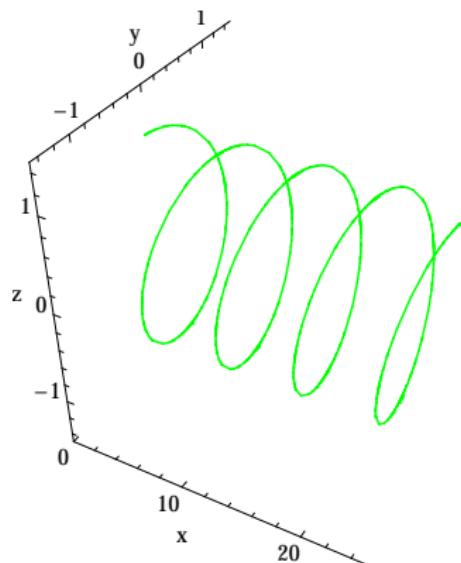


Tangent Lines of Curves

2.3.18. Example. Consider the curve given by

$$\gamma: [0, 8\pi] \rightarrow \mathbb{R}^3, \quad \gamma(t) = \begin{pmatrix} t \\ \cos t \\ \sin t \end{pmatrix}.$$

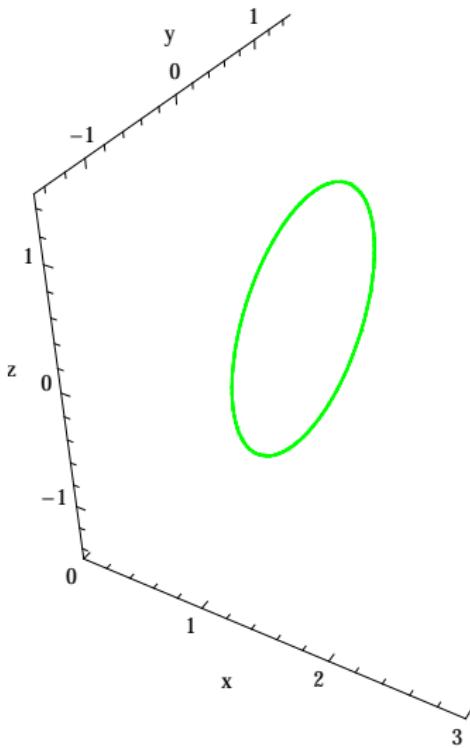
This curve is called a **helix**.



Tangent Lines of Curves

This is the graph of

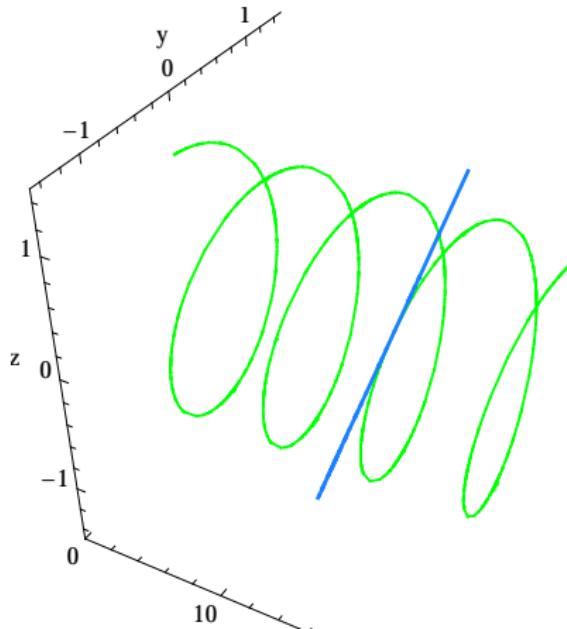
$$\gamma'(t) = \begin{pmatrix} 1 \\ -\sin t \\ \cos t \end{pmatrix}$$



Tangent Lines of Curves

The tangent line makes sense as a linear approximation

$$\gamma(t_0 + t) = \begin{pmatrix} t_0 \\ \cos t_0 \\ \sin t_0 \end{pmatrix} + \begin{pmatrix} 1 \\ -\sin t_0 \\ \cos t_0 \end{pmatrix} t + o(t)$$



The Tangent Vector to a Curve

2.3.19. Definition. Let $\mathcal{C}^* \subset \mathbb{R}^n$ be an oriented smooth curve and $p \in \mathcal{C}^*$. Let $\gamma: I \rightarrow \mathbb{R}^n$ be a parametrization of \mathcal{C}^* . Then we define the **unit tangent vector** to \mathcal{C}^* at $p = \gamma(t)$ by

$$T \circ \gamma(t) := \frac{\gamma'(t)}{\|\gamma'(t)\|}, \quad t \in \text{int } I. \quad (2.3.4)$$

This defines the **tangent vector field** $T: \mathcal{C}^* \rightarrow \mathbb{R}^n$ on \mathcal{C} .

Note that (2.3.4) does not depend on the parametrization γ , as long as γ corresponds to the orientation of \mathcal{C}^* .

In fact, suppose $\gamma: I \rightarrow \mathcal{C}$, $\tilde{\gamma}: J \rightarrow \mathcal{C}$ are two smooth parametrizations connected by a reparametrization $r: J \rightarrow I$ so that $\tilde{\gamma} = \gamma \circ r$. Let $p \in \mathcal{C}$ satisfy $p = \gamma(t) = \tilde{\gamma}(\tau)$, $t = r(\tau)$.

Then

$$\tilde{\gamma}'(\tau) = \frac{d}{d\tau} \gamma(r(\tau)) = \gamma'(r(\tau))r'(\tau) = \gamma'(t)r'(\tau).$$

The Tangent Vector to a Curve

Hence,

$$T \circ \tilde{\gamma}(\tau) = \frac{\tilde{\gamma}'(\tau)}{\|\tilde{\gamma}'(\tau)\|} = \frac{r'(\tau)}{|r'(\tau)|} \frac{\gamma'(t)}{\|\gamma'(t)\|} = \frac{r'(\tau)}{|r'(\tau)|} T \circ \gamma(t). \quad (2.3.5)$$

We see that if r is orientation-preserving, then $r'(t) > 0$ and the tangent vector is the same when calculated using γ as when using $\tilde{\gamma}$. If r is orientation-reversing, the tangent vector reverses direction. This proves that (2.3.4) defines a unique unit tangent vector for an oriented curve.

The Tangent Vector to a Curve

2.3.20. Example. Consider the circle of radius R in \mathbb{R}^2 ,

$$\mathcal{C} := \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 = R^2\}.$$

By choosing the parametrization

$$\gamma: [0, 2\pi) \rightarrow \mathcal{C}, \quad \gamma(t) = \begin{pmatrix} R \cos t \\ R \sin t \end{pmatrix}$$

we endow \mathcal{C} with a positive (counter-clockwise) orientation. The unit tangent vector at $\gamma(t)$ is then given by

$$T \circ \gamma(t) = \frac{\gamma'(t)}{\|\gamma'(t)\|} = \frac{1}{R} \begin{pmatrix} -R \sin t \\ R \cos t \end{pmatrix} = \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix}.$$

Thus, if $p = (p_1, p_2) \in \mathcal{C}$, then

$$T(p) = \frac{1}{R} \begin{pmatrix} -p_2 \\ p_1 \end{pmatrix}.$$

Hence $T(p) \perp p$.

Rate of Change of the Tangent Vector

In order to gain more insight into the geometric properties of a curve, we will study the rate of change of the tangent vector as it “travels” along a curve. We assume from now on that

- ▶ V is a real inner product space and
- ▶ $\mathcal{C} \subset V$ has a parametrization γ such that γ'' exists and $\gamma'' \neq 0$.

We call such a \mathcal{C} a **smooth C^2 -curve** and γ a C^2 -parametrization.

We are interested in

$$\frac{d}{dt}(T \circ \gamma(t)). \quad (2.3.6)$$

Now $T \circ \gamma$ itself parametrizes a curve \mathcal{T} , so (2.3.6) gives the tangent vector of \mathcal{T} at $T \circ \gamma(t)$. Moreover, since $\|T\| = 1$ we see that

$$\mathcal{T} \subset S = \{x \in V : \|x\| = 1\}.$$

The Normal Vector of a Curve

Just as in Example 2.3.20, this implies that (2.3.6) is perpendicular to $T \circ \gamma(t)$:

$$\begin{aligned} 1 &= \|T \circ \gamma(t)\|^2 = \langle T \circ \gamma(t), T \circ \gamma(t) \rangle \\ \Rightarrow 0 &= \frac{d}{dt} \langle T \circ \gamma(t), T \circ \gamma(t) \rangle = 2 \langle (T \circ \gamma)'(t), T \circ \gamma \rangle \end{aligned} \quad (2.3.7)$$

2.3.21. Definition. Let $\mathcal{C} \subset V$ be a smooth C^2 -curve. Let $\gamma: I \rightarrow V$ be a smooth C^2 -parametrization of \mathcal{C} . Then the unit normal vector $N: \mathcal{C} \rightarrow \mathbb{R}$ is defined by

$$N \circ \gamma(t) := \frac{(T \circ \gamma)'(t)}{\|(T \circ \gamma)'(t)\|}, \quad t \in \text{int } I. \quad (2.3.8)$$

The Normal Vector of a Curve

The unit normal vector does not depend on γ , not even up to orientation: suppose $\gamma: I \rightarrow \mathcal{C}$, $\tilde{\gamma}: J \rightarrow \mathcal{C}$ are two C^2 -parametrizations connected by a C^2 -reparametrization $r: J \rightarrow I$ so that $\tilde{\gamma} = \gamma \circ r$. Let $p \in \mathcal{C}$ satisfy $p = \gamma(t) = \tilde{\gamma}(\tau)$, $t = r(\tau)$.

By (2.3.5),

$$T \circ \tilde{\gamma}(\tau) = \frac{r'(\tau)}{|r'(\tau)|} T \circ \gamma(t).$$

Suppose that r is orientation-reversing. Then

$$\frac{d}{d\tau} T \circ \tilde{\gamma}(\tau) = -\frac{d}{d\tau} T \circ \gamma(r(\tau)) = -(T \circ \gamma)'(r(\tau))r'(\tau)$$

and

$$N \circ \tilde{\gamma}(\tau) = -\frac{r'(\tau)}{|r'(\tau)|} \frac{(T \circ \gamma)'(t)}{\|(T \circ \gamma)'(t)\|} = \frac{(T \circ \gamma)'(t)}{\|(T \circ \gamma)'(t)\|}.$$

Of course, if r is orientation-preserving, the proof is similar.

The Normal Vector of a Curve

2.3.22. Example. We return to Example 2.3.20 and study the circle of radius R in \mathbb{R}^2 with parametrization

$$\gamma: [0, 2\pi) \rightarrow \mathcal{C}, \quad \gamma(t) = \begin{pmatrix} R \cos t \\ R \sin t \end{pmatrix}.$$

The unit tangent vector at $\gamma(t)$ is given by

$$T \circ \gamma(t) = \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} \quad \Rightarrow \quad (T \circ \gamma)'(t) = \begin{pmatrix} -\cos t \\ -\sin t \end{pmatrix}$$

and $\|(T \circ \gamma)'(t)\| = 1$. Then

$$N \circ \gamma(t) = \frac{(T \circ \gamma)'(t)}{\|(T \circ \gamma)'(t)\|} = \begin{pmatrix} -\cos t \\ -\sin t \end{pmatrix}$$

Thus, if $p = (p_1, p_2) \in \mathcal{C}$, then

$$N(p) = -\frac{1}{R}p$$

and we see $N(p) \perp T(p)$.

Physical interpretation of Tangent and Normal Vectors

If the parametrization $\gamma(t)$ describes the trajectory of a particle (point mass) and the curve $\mathcal{C} \subset \mathbb{R}^3$ the travelled path in space, then $\gamma'(t)$ gives the velocity of the particle.

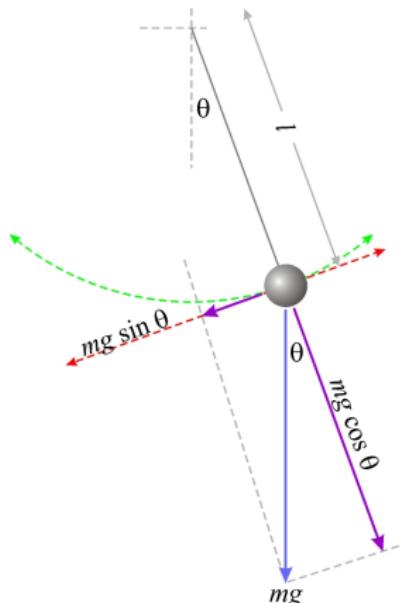
Hence, the unit tangent vector T gives the direction of the velocity and the unit normal vector N gives the direction of the change in T . By Newton's laws, a force F acting on a particle causes an acceleration, i.e., a change in $\gamma'(t)$. The acceleration will have a component parallel to T and a component orthogonal to T , given by N . Then F , proportional to the acceleration, can be decomposed as

$$F = F_T + F_N \quad \text{where} \quad F_T = \langle F, T \rangle T, \quad F_N = \langle F, N \rangle N.$$

The plane spanned by N and T is called the **osculating plane** (the latin word **osculum** means "to kiss"). It lies geometrically closest to the curve at the point where T and N are calculated.

Physical interpretation of Tangent and Normal Vectors

2.3.23. Example. Consider a mathematical pendulum, where a mass point swings at the length of a string under the influence of gravity $F_g = mg$.



The trajectory is a circle (segment) centered at the fulcrum of the pendulum. If we place the origin of our coordinate system at this fulcrum, the tangent vector will be orthogonal to the position of the mass point, pointing in the direction of movement. We obtain

$$F_T := \langle F, T \rangle T = m\langle g, T \rangle T = mg \sin \theta T,$$
$$F_N := \langle F, N \rangle N = m\langle g, N \rangle N = mg \cos \theta N,$$

where θ is shown in the image. The normal component F_N induces tension in the string, while F_T causes acceleration of the mass point along its trajectory.

Curve Length

Consider a simple curve $\mathcal{C} \subset V$ parametrized by $\gamma: [a, b] \rightarrow V$, where $(V, \|\cdot\|)$ is a normed vector space. Then a natural approximation to the length of γ , which **will** depend on the norm used, is found by taking a partition $\mathcal{P} = (a_0, \dots, a_n)$ of $[a, b]$ and considering the lengths of the straight line segments joining $(a_{i-1}, \gamma(a_{i-1}))$ to $(a_i, \gamma(a_i))$, $i = 1, \dots, n$.

The sum of the lengths of these line segments is given by

$$\ell_{\mathcal{P}}(\mathcal{C}) = \sum_{i=1}^n \|\gamma(a_i) - \gamma(a_{i-1})\|.$$

We will say that a curve has a length if there exists an upper bound to these lengths. Note that $\ell_{\mathcal{P}}(\mathcal{C})$ is of course independent of the parametrization γ , since only the actual points $\gamma(a_i) \in \mathcal{C}$ are used in this definition.

Curve Length

2.3.24. Definition. Let $(V, \|\cdot\|)$ be a normed vector space and $\mathcal{C} \subset V$ an open curve. Then we say that \mathcal{C} is **rectifiable** if

$$\ell(\mathcal{C}) := \sup_{\text{partitions } \mathcal{P}} \ell_{\mathcal{P}}(\mathcal{C})$$

exists. We then call $\ell(f)$ the **curve length** or **arc length** of \mathcal{C} .

2.3.25. Theorem. Let $\mathcal{C} \subset V$ be a smooth and open curve with parametrization $\gamma: [a, b] \rightarrow \mathcal{C}$. Then \mathcal{C} is rectifiable if and only if

$$\int_a^b \|\gamma'(t)\| dt < \infty.$$

Furthermore, if \mathcal{C} is rectifiable,

$$\ell(\mathcal{C}) = \int_a^b \|\gamma'(t)\| dt,$$

where the right-hand side is independent of γ .

Curve Length

2.3.26. Example. Consider the helix segment \mathcal{C} given by the graph of

$$\gamma: [0, 2\pi] \rightarrow \mathbb{R}^3, \quad \gamma(t) = \begin{pmatrix} \alpha t \\ R \cos t \\ R \sin t \end{pmatrix}, \quad \alpha, R > 0.$$

The length of $\mathcal{C} = \gamma([0, 2\pi])$ is given by

$$\begin{aligned} \ell(\mathcal{C}) &= \int_0^{2\pi} \|\gamma'(t)\| dt = \int_0^{2\pi} \sqrt{\alpha^2 + (-R \sin t)^2 + R^2 \cos^2 t} dt \\ &= 2\pi\sqrt{\alpha^2 + R^2}. \end{aligned}$$

2.3.27. Remark. Definition 2.3.24 and Theorem 2.3.25 refer to open curves. To find the length of a closed curve, we express it as the disjoint union of two simple curves and find their lengths separately.

Curve Length

Proof of Theorem 2.3.25.

We first show that the value of the integral

$$\int_a^b \|\gamma'(t)\| dt$$

does not depend on the parametrization γ . Let $\mathcal{C} \subset V$ be a smooth curve and $\gamma: [a, b] \rightarrow \mathcal{C}$ a parametrization of \mathcal{C} . Let $\tilde{\gamma}: [\alpha, \beta] \rightarrow \mathcal{C}$ be some other parametrization and let $r: [\alpha, \beta] \rightarrow [a, b]$ be a smooth reparametrization so that

$$\tilde{\gamma}(\tau) = \gamma(r(\tau))$$

Suppose that r is orientation-preserving. Then $r(\alpha) = a$ and $r(\beta) = b$.

Curve Length

Proof (continued).

Let $t = r(\tau)$ so $dt = r'(\tau) d\tau$. Furthermore,

$$\tilde{\gamma}'(\tau) = (\gamma \circ r(\tau))' = \gamma'(r(\tau))r'(\tau) \quad \Leftrightarrow \quad \gamma'(r(\tau)) = \frac{\tilde{\gamma}'(\tau)}{r'(\tau)}$$

Thus, substituting $t = r(\tau)$,

$$\begin{aligned} \int_a^b \|\gamma'(t)\| dt &= \int_{\alpha}^{\beta} \|\gamma'(r(\tau))\| r'(\tau) d\tau = \int_{\alpha}^{\beta} \left\| \frac{\tilde{\gamma}'(\tau)}{r'(\tau)} \right\| r'(\tau) d\tau \\ &= \int_{\alpha}^{\beta} \frac{\|\tilde{\gamma}'(\tau)\|}{|r'(\tau)|} r'(\tau) d\tau = \int_{\alpha}^{\beta} \|\tilde{\gamma}'(\tau)\| d\tau, \end{aligned}$$

where we have used that r is increasing, i.e., $r'(\tau) > 0$. This proves that $\int_a^b \|\gamma'(t)\| dt$ is independent of the parametrization γ .

Curve Length

Proof (continued).

Now, for any partition \mathcal{P} of $[a, b]$ and any parametrization γ we have

$$\begin{aligned}\ell_{\mathcal{P}}(\mathcal{C}) &= \sum_{i=1}^n \|\gamma(a_i) - \gamma(a_{i-1})\| \\ &= \sum_{i=1}^n \left\| \int_{a_{i-1}}^{a_i} \gamma'(t) dt \right\| \\ &\leq \sum_{i=1}^n \int_{a_{i-1}}^{a_i} \|\gamma'(t)\| dt \\ &= \int_a^b \|\gamma'(t)\| dt.\end{aligned}$$

Hence, $\ell(\mathcal{C}) \leq \int_a^b \|\gamma'(t)\| dt.$

Curve Length

Proof (continued).

Proving the converse inequality is slightly more difficult. We first establish three preliminary estimates. We will use the fact that since \mathcal{C} is smooth, $\gamma': [a, b] \rightarrow V$ is continuous and hence uniformly continuous on $[a, b]$ (see Theorem 2.1.38). Fix $\varepsilon > 0$.

(i) Choose a $\delta > 0$ such that

$$|t - \tau| < \delta \quad \Rightarrow \quad \|\gamma'(t) - \gamma'(\tau)\| < \frac{\varepsilon}{2(b-a)}$$

for all $t, \tau \in [a, b]$.

Curve Length

Proof (continued).

- (ii) Consider the function $f: [a, b] \rightarrow \mathbb{R}$, $f(t) = \|\gamma'(t)\|$. Since γ is smooth, f is continuous and we can find a step function that uniformly approximates f . In fact, there is a $0 < \delta_1 < \delta$ and a partition $\mathcal{P} = (a_0, \dots, a_n)$ with $a_i - a_{i-1} < \delta_1$, $i = 1, \dots, n$, of $[a, b]$ such that

$$\left| \int_a^b \|\gamma'(t)\| dt - \sum_{i=1}^n (a_i - a_{i-1}) \|\gamma'(a_{i-1})\| \right| < \varepsilon/2$$

Curve Length

Proof (continued).

(iii) For any $t \in (a, b)$ and $h < \delta$ with $t + h \in [a, b]$,

$$\begin{aligned}\left\| \frac{\gamma(t+h) - \gamma(t)}{h} - \gamma'(t) \right\| &= \left\| \frac{1}{h} \int_t^{t+h} \gamma'(\tau) d\tau - \gamma'(t) \right\| \\ &= \left\| \frac{1}{h} \int_t^{t+h} (\gamma'(\tau) - \gamma'(t)) d\tau \right\| \\ &\leq \frac{1}{h} \int_t^{t+h} \|\gamma'(\tau) - \gamma'(t)\| d\tau \\ &\leq \frac{h}{h} \sup_{\tau \in [t, t+h]} \|\gamma'(\tau) - \gamma'(t)\| < \frac{\varepsilon}{2(b-a)}.\end{aligned}$$

This implies $\|\gamma'(t)\| \leq \left\| \frac{\gamma(t+h) - \gamma(t)}{h} \right\| + \frac{\varepsilon}{2(b-a)}$.

Curve Length

Proof (continued).

We then have

$$\begin{aligned} \int_a^b \|\gamma'(t)\| dt &\leq \sum_{i=1}^n (a_i - a_{i-1}) \|\gamma'(a_{i-1})\| + \frac{\varepsilon}{2} \\ &\leq \sum_{i=1}^n (a_i - a_{i-1}) \left\| \frac{\gamma(a_i) - \gamma(a_{i-1})}{a_i - a_{i-1}} \right\| + \frac{\varepsilon}{2(b-a)}(b-a) + \frac{\varepsilon}{2} \\ &\leq \sum_{i=1}^n (a_i - a_{i-1}) \frac{\|\gamma(a_i) - \gamma(a_{i-1})\|}{|a_i - a_{i-1}|} + \varepsilon \\ &= \sum_{i=1}^n \|\gamma(a_i) - \gamma(a_{i-1})\| + \varepsilon = \ell_{\mathcal{P}}(\mathcal{C}) + \varepsilon \\ &\leq \ell(\mathcal{C}) + \varepsilon \end{aligned}$$

Letting $\varepsilon \rightarrow 0$, we obtain the desired inequality. □

Curve Length

We can now express the total curve length by

$$\ell(\mathcal{C}) = \int_a^b \|\gamma'(t)\| dt.$$

More generally, we can define a **length function**

$$(\ell \circ \gamma)(t) = \int_a^t \|\gamma'(\tau)\| d\tau \tag{2.3.9}$$

so that $(\ell \circ \gamma)(b) = \ell(\mathcal{C})$.

The function

$$\ell \circ \gamma: [a, b] \rightarrow [0, \infty)$$

associates to each value of t the length of the curve at $\gamma(t)$. Since the integral (2.3.9) is strictly monotonic, $\ell \circ \gamma$ is strictly increasing and hence bijective.

Curve Length

2.3.28. Example. We return to Example 2.3.20 and study the circle of radius R in \mathbb{R}^2 with parametrization

$$\gamma: [0, 2\pi) \rightarrow \mathcal{C}, \quad \gamma(t) = \begin{pmatrix} R \cos t \\ R \sin t \end{pmatrix}.$$

The curve length is given by

$$(\ell \circ \gamma)(t) = \int_0^t \|\gamma'(\tau)\| d\tau = \int_0^t R d\tau = Rt.$$

Thus, for $p = (p_1, p_2) \in \mathcal{C}$, we can read off

$$\ell(p) = R \cdot \arctan \frac{p_2}{p_1}.$$

where the arctangent is understood in the sense of (2.3.1); i.e., the appropriate branch is chosen depending on the signs of p_1 and p_2 .

Parametrization Through Curve Length

The map $\ell: \mathcal{C}^* \rightarrow [0, \infty)$ is bijective, since its inverse is given by

$$\ell^{-1} = \gamma \circ (\ell \circ \gamma)^{-1}.$$

If $s = \ell(p)$ is the curve length at $p \in \mathcal{C}^*$, then $p = \ell^{-1}(s)$ is the unique point in \mathcal{C} associated to this curve length.

In other words, once we fix an orientation of a simple curve \mathcal{C} , the curve length determines all other points of \mathcal{C}^* uniquely. This means that we can use the curve length as a **natural parametrization** of \mathcal{C} , i.e., we can parametrize \mathcal{C} using

$$\gamma = \ell^{-1}: I \rightarrow \mathcal{C}, \quad \text{int } I = (0, \ell(\mathcal{C})).$$

If we want to parametrize closed curves through curve length, we must fix an “initial point” in some fashion.

Curvature

We are interested in the rate of change of the direction of the tangent vector to a curve. However, while T does not depend on γ , if we simply differentiate $T \circ \gamma(t)$, the derivative **will** depend on γ . In order to obtain a purely geometric measure for the rate of change of T , we need to settle on a “canonical” parametrization. Luckily, we have just developed one: parametrization using curve length. This parametrization takes into account only the specific geometric properties of the curve.

2.3.29. Definition. The **curvature** of a smooth C^2 -curve $\mathcal{C} \subset V$ is

$$\kappa: \mathcal{C} \rightarrow \mathbb{R}, \quad \kappa \circ \ell^{-1}(s) := \left\| \frac{d}{ds} (T \circ \ell^{-1}(s)) \right\|$$

where T is the unit tangent vector and $\ell^{-1}: I \rightarrow \mathcal{C}$ is the curve length parametrization of \mathcal{C} .

Note that, like the unit normal vector N , κ also does not depend on the orientation of \mathcal{C} .

Curvature in Arbitrary Parametrization

Given a parametrization $\gamma: I \rightarrow \mathcal{C}$ of \mathcal{C} (which is not the curve length), then by the chain rule

$$\frac{d(T \circ \gamma)}{dt} \Big|_t = \frac{d(T \circ \ell^{-1})}{ds} \Big|_{s=\ell \circ \gamma(t)} \frac{d(\ell \circ \gamma)}{dt} \Big|_t.$$

Using (2.3.9),

$$\frac{d(T \circ \ell^{-1})}{ds} \Big|_{s=\ell \circ \gamma(t)} = \frac{1}{\|\gamma'(t)\|} \frac{d(T \circ \gamma)}{dt} \Big|_t$$

and so we obtain for the curvature at $p = \gamma(t) = \ell^{-1}(s)$

$$\kappa \circ \gamma(t) = \kappa \circ \ell^{-1}(s) \Big|_{s=\ell \circ \gamma(t)} = \frac{\|(T \circ \gamma)'(t)\|}{\|\gamma'(t)\|}. \quad (2.3.10)$$

Curvature

2.3.30. Example. We return to Example 2.3.20 and study the circle of radius R in \mathbb{R}^2 with parametrization

$$\gamma: [0, 2\pi) \rightarrow \mathcal{C}, \quad \gamma(t) = \begin{pmatrix} R \cos t \\ R \sin t \end{pmatrix}.$$

We have

$$T \circ \gamma(t) = \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix},$$

so by (2.3.10)

$$\kappa \circ \gamma(t) = \frac{\|(T \circ \gamma)'(t)\|}{\|\gamma'(t)\|} = \frac{1}{R} \left\| \begin{pmatrix} -\cos t \\ -\sin t \end{pmatrix} \right\| = \frac{1}{R}.$$

Thus, the curvature of circle is constant and equal to the inverse of its radius.

Curvature in \mathbb{R}^3

2.3.31. Lemma. Let $\mathcal{C} \subset \mathbb{R}^3$ be a smooth C^2 -curve with parametrization $\gamma: I \rightarrow \mathcal{C}$. Then

$$\kappa \circ \gamma(t) = \frac{\|\gamma'(t) \times \gamma''(t)\|}{\|\gamma'(t)\|^3} \quad (2.3.11)$$

Proof.

From (2.3.4) we have $\gamma'(t) = \|\gamma'(t)\|(T \circ \gamma)(t)$. The product rule yields

$$\gamma''(t) = \frac{d\|\gamma'(t)\|}{dt}(T \circ \gamma)(t) + \|\gamma'(t)\|(T \circ \gamma)'(t).$$

Now, since $T \times T = 0$ and $T \perp T'$ we have

$$\begin{aligned} \|\gamma'(t) \times \gamma''(t)\| &= \|\gamma'(t)\|^2 \cdot \|(T \circ \gamma)(t) \times (T \circ \gamma)'(t)\| \\ &= \|\gamma'(t)\|^2 \cdot \|(T \circ \gamma)'(t)\| \end{aligned} \quad (2.3.12)$$

Curvature in \mathbb{R}^3

Proof (continued).

From

$$\|\gamma'(t) \times \gamma''(t)\| = \|\gamma'(t)\|^2 \cdot \|(T \circ \gamma)'(t)\|$$

we obtain

$$\kappa \circ \gamma(t) = \frac{\|(T \circ \gamma)'(t)\|}{\|\gamma'(t)\|} = \frac{\|\gamma'(t) \times \gamma''(t)\|}{\|\gamma'(t)\|^3}$$

□

2.3.32. Remark. For technical reasons we have defined curvature only for simple open curves. Of course, the curvature is a local property, and if we are interested in the curvature at point, we may simply look at a simple and open subsection of the curve that includes this point.

Convergence and Continuity

Functions and Derivatives

Curves in Vector Spaces

Potential Functions

The Second Derivative

Extrema of Potential Functions

Constrained Extrema

Potentials

A map $f: \Omega \rightarrow \mathbb{R}$ where $\Omega \subset \mathbb{R}^n$ is called a scalar function or a **potential**.

Physically, if $n = 3$, a potential assigns to each point in space a scalar value. Examples include temperature, pressure or height.

The Jacobian of a differentiable potential is given by

$$Df|_x = \begin{pmatrix} \frac{\partial f}{\partial x_1}|_x & \cdots & \frac{\partial f}{\partial x_n}|_x \end{pmatrix}.$$

The row vector $Df|_x$ may be regarded as a linear map $Df|_x: \mathbb{R}^n \rightarrow \mathbb{R}$,

$$Df|_x y = \begin{pmatrix} \frac{\partial f}{\partial x_1}|_x & \cdots & \frac{\partial f}{\partial x_n}|_x \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i=1}^n y_i \left. \frac{\partial f}{\partial x_i} \right|_x$$

Thus $Df|_x \in (\mathbb{R}^n)^*$, the dual space of \mathbb{R}^n (see Examples 1.4.6 ii)).

Coordinate Maps

Classically, we considered x_j ($1 \leq j \leq n$) as a coordinate of the vector $x = (x_1, \dots, x_n) \in \mathbb{R}^n$. We now introduce a different interpretation. Define the map

$$\mathbb{R}^n \rightarrow \mathbb{R}, \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto x_j.$$

This map is clearly linear; it is the **coordinate map** that assigns to $x \in \mathbb{R}^n$ its coordinate x_j . We denote this map by x_j also; hence $x_j: x \mapsto x_j$ or $x_j(x) = x_j$. The dual meaning of x_j as a map and the value of this map is very convenient. In fact, the entire discipline of **differential geometry** hinges on exploiting this ambiguity.

The derivative of the map x_j is given by

$$dx_j = (0, \dots, 0, \underset{j}{\uparrow}, 1, 0, \dots, 0) \tag{2.4.1}$$

The Differential and the Gradient

Note that we have written dx_j instead of Dx_j ; this is traditional. The derivative of a scalar map f is also called a **differential** and written df instead of Df . Note also that $dx_j|_x$ does not depend on x . Therefore, we have

$$df|_x = \left(\frac{\partial f}{\partial x_1} \Big|_x \quad \dots \quad \frac{\partial f}{\partial x_n} \Big|_x \right) = \frac{\partial f}{\partial x_1} \Big|_x dx_1 + \dots + \frac{\partial f}{\partial x_n} \Big|_x dx_n$$

Each differential $dx_j = e_j^*$ is the dual basis vector to the standard basis vector e_j with respect to the euclidean scalar product.

The transpose of the Jacobian is the **gradient**,

$$\nabla f(x) := (J_f(x))^T = \begin{pmatrix} \frac{\partial f}{\partial x_1} \Big|_x \\ \vdots \\ \frac{\partial f}{\partial x_n} \Big|_x \end{pmatrix}$$

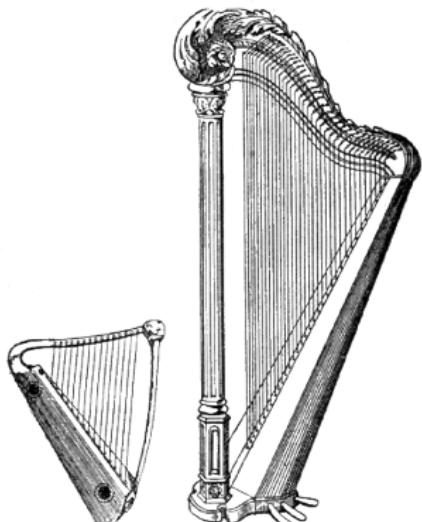
The triangle symbol is called **nabla**.

About Nabla

The term “nabla” comes from the greek word for a Phoenician harp, whose shape the nabla triangle ∇ is supposed to resemble.

It was used by the physicists James Maxwell and Peter Tait (the latter developed much of the modern mathematics of the nabla operator) in their private correspondence. There is evidence that this was a private joke between them and Maxwell did not use the term in serious publications. However, it became popular nevertheless, being used by William Thomson (Lord Kelvin) at the end of the 19th century.

Another proposal has been to call the symbol “del”, but it seems that “nabla” is the most common term today.



Harps, p. 984.

Harps from 1911 Webster's Dictionary. Wikimedia Commons. Wikimedia Foundation. Web. 18 July 2018

The Directional Derivative

2.4.1. Definition. Let $\Omega \subset \mathbb{R}^n$ be an open set, $f: \Omega \rightarrow \mathbb{R}$ continuous and $h \in \mathbb{R}^n$, $\|h\| = 1$, be a unit vector. Then the **directional derivative** $D_h f$ in the direction h is defined by

$$D_h f|_x := \frac{d}{dt} f(x + th) \Big|_{t=0}. \quad (2.4.2)$$

if the right-hand side exists.

The directional derivative has the following interpretation: if $\gamma(t) = x + th$, $t \in [0, 1]$, parametrizes the straight line segment joining x and $x + h$, then $D_h f$ is simply the derivative of $f \circ \gamma$.

Hence,

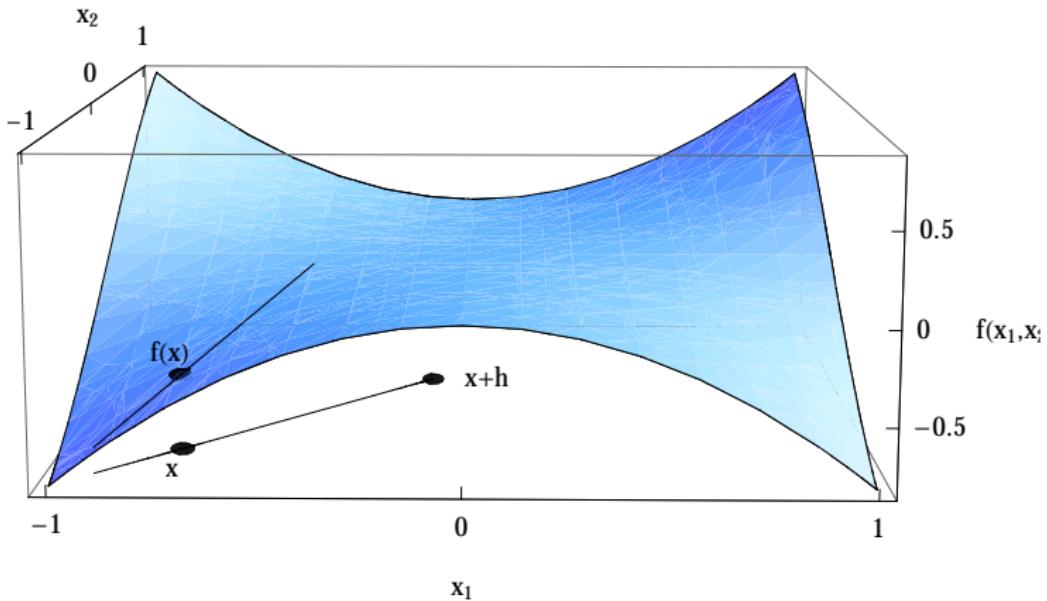
The directional derivative $D_h f|_x$ is the derivative of f at x along the line segment joining x and $x + h$.

Another way of stating this is

The directional derivative $D_h f|_x$ gives the slope of the tangent line of f at x in the direction of h .

The Directional Derivative

It is essential that $\|h\| = 1$, otherwise the slope will not be scaled correctly!



The Directional Derivative

We note that the tangent line of $f: \mathbb{R}^n \rightarrow \mathbb{R}$ at x in the direction h is given by

$$t_{f,x;h}(s) = \begin{pmatrix} x + sh \\ f(x) + D_h f|_x s \end{pmatrix}, \quad s \in \mathbb{R}, \quad (2.4.3)$$

where $h \in \mathbb{R}^n$ (so the above vector is a “block vector” with $n + 1$ entries).

For functions $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ the directional derivative is sometimes specified through the angle θ . This understood to mean that

$$h = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}.$$

Note that the directional derivative is a **number**, in contradistinction to the derivative. Thus it should perhaps be more properly known as the “directional slope.”

The Directional Derivative

2.4.2. Example. Let $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x_1, x_2) = x_1^2 - 4x_2$. Then the directional derivative of f at x in the direction is

$$\begin{aligned} D_h f|_x &= \frac{d}{dt} f(x + th) \Big|_{t=0} = \frac{d}{dt} ((x_1 + th_1)^2 - 4(x_2 + th_2)) \Big|_{t=0} \\ &= 2h_1(x_1 + th_1) - 4h_2 \Big|_{t=0} \\ &= 2h_1x_1 - 4h_2. \end{aligned}$$

For $h = (1/\sqrt{2}, 1/\sqrt{2})$ (or $\theta = \pi/4$) we would have

$$D_h f|_x = \sqrt{2}x_1 - 2\sqrt{2}.$$

At $x = (0, 0)$, the directional derivative in direction h is

$$D_h f|_{x=0} = -2\sqrt{2}$$

The Directional Derivative for Smooth Functions

Suppose that f is differentiable. If $\gamma(t) = x + th$, $t \in [0, 1]$, parametrizes the straight line segment joining x and $x + h$, then by the chain rule

$$D_h f|_x = \frac{d}{dt} f(x + th) \Big|_{t=0} = Df|_{x+th} h \Big|_{t=0} = Df|_x h$$

so

$$D_h f|_x = Df|_x h = \langle \nabla f(x), h \rangle. \quad (2.4.4)$$

This is a useful expression for calculating the directional derivative, but it supposes that f is differentiable. In practice, (2.4.4) will be valid if the partial derivatives of f exist and are continuous at x .

The Directional Derivative for Smooth Functions

2.4.3. Example. Returning to Example 2.4.2, we have

$$\nabla f(x) = \begin{pmatrix} 2x_1 \\ -4 \end{pmatrix}$$

Since the partial derivatives are continuous,

$$D_h f|_x = \langle \nabla f(x), h \rangle = \left\langle \begin{pmatrix} 2x_1 \\ -4 \end{pmatrix}, \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \right\rangle = 2x_1 h_1 - 4h_2.$$

This coincides with the result obtained previously.

The Normal Derivative in \mathbb{R}^2

An important special case of the directional derivative is the **normal derivative in \mathbb{R}^2** .

2.4.4. Definition. Let $\Omega \subset \mathbb{R}^2$ be an open set, $f: \Omega \rightarrow \mathbb{R}$ and \mathcal{C} a simple smooth C^2 curve in Ω . Let $p \in \mathcal{C}$ and $N(p)$ denote the normal vector at p . Then

$$\left. \frac{\partial f}{\partial n} \right|_p := D_{N(p)} f|_p$$

is called the **normal derivative of f at p** (with respect to the curve \mathcal{C}).

2.4.5. Example. Let $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x_1, x_2) = x_1^2 - 4x_2$, and

$$\mathcal{C} = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 = x_1^2, x_1 \in \mathbb{R}\}.$$

Then \mathcal{C} is parametrized by $\gamma(t) = (t, t^2)$, $t \in \mathbb{R}$, and

$$T \circ \gamma(t) = \frac{1}{\sqrt{1+4t^2}} \begin{pmatrix} 1 \\ 2t \end{pmatrix}.$$

The Normal Derivative in \mathbb{R}^2

The normal vector is then found from

$$\begin{aligned}(T \circ \gamma)'(t) &= \frac{-4t}{(1+4t^2)^{3/2}} \begin{pmatrix} 1 \\ 2t \end{pmatrix} + \frac{1}{\sqrt{1+4t^2}} \begin{pmatrix} 0 \\ 2 \end{pmatrix} \\ &= \begin{pmatrix} \frac{-4t}{(1+4t^2)^{3/2}} \\ \frac{-8t^2+2(1+4t^2)}{(1+4t^2)^{3/2}} \end{pmatrix} \\ &= \frac{2}{(1+4t^2)^{3/2}} \begin{pmatrix} -2t \\ 1 \end{pmatrix}\end{aligned}$$

The unit normal vector is found by normalizing $(T \circ \gamma)'$, so we have

$$N \circ \gamma(t) = \frac{1}{\sqrt{1+4t^2}} \begin{pmatrix} -2t \\ 1 \end{pmatrix}.$$

The Normal Derivative in \mathbb{R}^2

At a point $p = \gamma(t)$ on \mathcal{C} the normal derivative is hence

$$\begin{aligned}\frac{\partial f}{\partial n}\Big|_{\gamma(t)} &= \langle \nabla f(\gamma(t)), N \circ \gamma(t) \rangle \\ &= \frac{1}{\sqrt{1+4t^2}} \left\langle \begin{pmatrix} 2t \\ -4 \end{pmatrix}, \begin{pmatrix} -2t \\ 1 \end{pmatrix} \right\rangle \\ &= -\frac{4(t^2 + 1)}{\sqrt{1+4t^2}}\end{aligned}$$

The Gradient

The gradient vector of f at x , $\nabla f(x)$, has some interesting properties:

- ▶ $\nabla f(x)$ **points in the direction of the greatest directional derivative of f at x .**

This follows from

$$D_h f(x) = \langle \nabla f(x), h \rangle = |\nabla f(x)| \cos \angle(\nabla f(x), h),$$

which becomes maximal if $\angle(\nabla f(x), h) = 0$.

- ▶ $\nabla f(x)$ **is perpendicular to the contour line of f at x .**

More precisely, it is perpendicular to the tangent line of the contour line at x . This is due to the fact that the tangent line to the contour is parallel to the direction h_0 in which $D_{h_0} f(x) = 0$, so

$$\langle \nabla f(x), h_0 \rangle = 0.$$

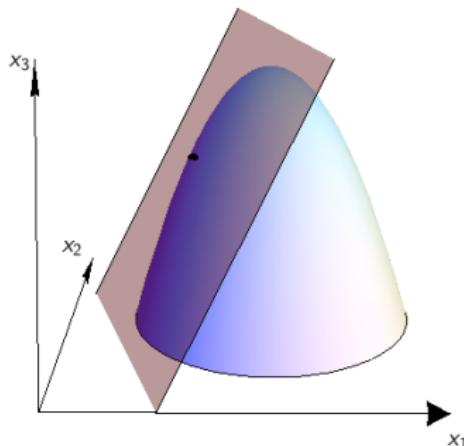
The Tangent Plane to the Graph of a Function

Consider the graph of a function $f: \Omega \rightarrow \mathbb{R}$, $\Omega \subset \mathbb{R}^n$ an open set. The graph of f is

$$\Gamma(f) := \{(x_1, \dots, x_n, x_{n+1}) \in \mathbb{R}^{n+1} : x_{n+1} = f(x_1, \dots, x_n), (x_1, \dots, x_n) \in \Omega\}.$$

It is an example of a **surface** in \mathbb{R}^{n+1} . (We will study general surfaces in a later section.)

We are interested in finding the **tangent plane to the graph of f** .



The Tangent Plane to the Graph of a Function

Recall that for $f: \mathbb{R}^n \rightarrow \mathbb{R}$ and $x_0 \in \mathbb{R}^n$ we have

$$f(x_0 + h) = f(x_0) + Df|_{x_0} h + o(h) \quad \text{as } h \rightarrow 0.$$

Setting $x = x_0 + h$, we can rewrite this as

$$\begin{aligned} f(x) &= f(x_0) + Df|_{x_0}(x - x_0) + o(x - x_0) \quad \text{as } x \rightarrow x_0. \\ &\approx f(x_0) + Df|_{x_0}(x - x_0). \end{aligned}$$

We hence define the function

$$Tf(\cdot; x_0) := f(x_0) + Df|_{x_0}(\cdot - x_0)$$

which gives the **best linear approximation** to f at x_0 .

The Tangent Plane to the Graph of a Function

2.4.6. Definition. Let $\Omega \subset \mathbb{R}^n$ be open and $f: \Omega \rightarrow \mathbb{R}$ differentiable at $x_0 \in \Omega$. Then the equation

$$x_{n+1} = Tf(x; x_0), \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n,$$

defines the **tangent plane** to the graph $\Gamma(f) \in \mathbb{R}^n \times \mathbb{R}$ of f at the point $(x_0, f(x_0)) \in \mathbb{R}^{n+1}$.

It is easy to see that the tangent plane actually is a plane. To illustrate, let us discuss the case $n = 2$ in more detail.

The tangent plane of a function $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ at (x_0, y_0) is given by

$$\begin{aligned} z &= f(x_0, y_0) + \left(\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0) \right) \left(\begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \right) \\ &= f(x_0, y_0) + (x - x_0) \frac{\partial f}{\partial x}(x_0, y_0) + (y - y_0) \frac{\partial f}{\partial y}(x_0, y_0) \end{aligned} \tag{2.4.5}$$

Tangent Vectors to the Graph of a Function

The equation (2.4.5) for the tangent plane is a single linear equation in the three unknowns $(x, y, z) \in \mathbb{R}^3$. The general solution can be found to be

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \\ f(x_0, y_0) \end{pmatrix} + s \begin{pmatrix} 1 \\ 0 \\ \frac{\partial f}{\partial x}(x_0, y_0) \end{pmatrix} + t \begin{pmatrix} 0 \\ 1 \\ \frac{\partial f}{\partial y}(x_0, y_0) \end{pmatrix}, \quad (2.4.6)$$

which defines a plane in \mathbb{R}^3 . The vectors

$$t_1 := \begin{pmatrix} 1 \\ 0 \\ \frac{\partial f}{\partial x}(x_0, y_0) \end{pmatrix} \quad \text{and} \quad t_2 := \begin{pmatrix} 0 \\ 1 \\ \frac{\partial f}{\partial y}(x_0, y_0) \end{pmatrix}$$

are called **tangent vectors** to the graph $\Gamma(f)$ at $(x_0, y_0, f(x_0, y_0))$.

The Normal Vector to the Graph of a Function From (2.4.6)

we can find a vector orthogonal to the tangent plane by taking the vector product

$$\begin{aligned} n &= \begin{pmatrix} 1 \\ 0 \\ \frac{\partial f}{\partial x} \end{pmatrix} \times \begin{pmatrix} 0 \\ 1 \\ \frac{\partial f}{\partial y} \end{pmatrix} = \det \begin{pmatrix} e_1 & e_2 & e_3 \\ 1 & 0 & \frac{\partial f}{\partial x} \\ 0 & 1 & \frac{\partial f}{\partial y} \end{pmatrix} \\ &= \begin{pmatrix} -\frac{\partial f}{\partial x} \\ -\frac{\partial f}{\partial y} \\ 1 \end{pmatrix} \end{aligned}$$

This gives a (non-normalized) **normal vector to the graph of f .**

The above discussion carries over naturally to the case of general $n \geq 2$. Note that a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ will have n tangent and a single normal vector to its graph at all points.

In a later chapter we will study general surfaces and their tangent and normal vectors.

Convergence and Continuity

Functions and Derivatives

Curves in Vector Spaces

Potential Functions

The Second Derivative

Extrema of Potential Functions

Constrained Extrema

The Second Derivative

In the next section, we wish to discuss maxima and minima of potential functions on \mathbb{R}^n . This requires us to analyze the concept of the second derivative of a function a little more closely.

2.5.1. Definition. Let X, V be finite-dimensional normed vector spaces and $\Omega \subset X$ an open set. A function $f: \Omega \rightarrow V$ is said to be **twice differentiable at $x \in \Omega$** if

- ▶ f is differentiable in an open ball $B_\varepsilon(x)$ around x and
- ▶ the function $Df: B_\varepsilon(x) \rightarrow \mathcal{L}(X, V)$ is differentiable at x .

We say that f is twice differentiable on Ω if f is twice differentiable at every $x \in \Omega$.

The derivative of Df (if it exists) is a map

$$D(Df) =: D^2f: \Omega \rightarrow \mathcal{L}(X, \mathcal{L}(X, V)). \quad (2.5.1)$$

We call (2.5.1) the **second derivative of f** . If the map $x \mapsto D^2f|_x$ is continuous on Ω we say that $f \in C^2(\Omega, V)$.

The Second Derivative for Potential Functions

2.5.2. Example. Consider a differentiable potential function $f: \mathbb{R}^n \rightarrow \mathbb{R}$. Then the derivative is given by the Jacobian

$$Df|_x = \begin{pmatrix} \frac{\partial f}{\partial x_1}|_x & \dots & \frac{\partial f}{\partial x_n}|_x \end{pmatrix}.$$

The second derivative is the derivative of the map $Df: \mathbb{R}^n \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R})$,

$$Df: x \mapsto Df|_x = \begin{pmatrix} \frac{\partial f}{\partial x_1}|_x & \dots & \frac{\partial f}{\partial x_n}|_x \end{pmatrix}.$$

The map Df is of course in general non-linear. The derivative is found by taking

$$Df|_{x+h} = Df|_x + D^2f|_x h + o(h) \quad \text{as } h \rightarrow 0.$$

Here $Df|_x$ and $Df|_{x+h} \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ are linear maps from $\mathbb{R}^n \rightarrow \mathbb{R}$, while

$$D^2f|_x \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}))$$

so that $D^2f|_x h \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$.

The Second Derivative for Potential Functions

Now what shape does $D^2f|_x$ take? A “column vector” $h \in \mathbb{R}^n$ is transformed into a linear map in $\mathcal{L}(\mathbb{R}^n, \mathbb{R}) \simeq \text{Mat}(n \times 1; \mathbb{R})$, which we can regard as the space of “row vectors”.

Recall that in this case $Df|_x = (\nabla f(x))^T$ and that the transposition is a linear map. Then we can write

$$Df: x \mapsto (\nabla f(x))^T = Df|_x.$$

Hence, $Df = (\cdot)^T \circ \nabla f$ and we can differentiate Df by the chain rule. The derivative of the map

$$\nabla f: \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad \nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1}|_x \\ \vdots \\ \frac{\partial f}{\partial x_n}|_x \end{pmatrix}$$

can be easily calculated: it is just the Jacobian of ∇f .

The Hessian

We hence have

$$D(\nabla f)|_x = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} \Big|_x & \frac{\partial^2 f}{\partial x_2 \partial x_1} \Big|_x & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1} \Big|_x \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} \Big|_x & \frac{\partial^2 f}{\partial x_2 \partial x_n} \Big|_x & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n} \Big|_x \end{pmatrix} \in \text{Mat}(n \times n; \mathbb{R}). \quad (2.5.2)$$

where

$$\frac{\partial^2 f}{\partial x_i \partial x_j} := \frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_j}$$

is the second partial derivative of f with respect to x_j (first) and x_i (second). The matrix in (2.5.2) is important enough to warrant a special name: It is called the **Hessian** of f and denoted by

$$\text{Hess } f(x).$$

The Hessian as a Bilinear Map

Recall that the transposition is a linear map, so its derivative is again the transposition (see Example 2.2.10). Hence,

$$\begin{aligned} D^2f|_x h &= D((\cdot)^T \circ \nabla f(x))h = D(\cdot)^T|_{\nabla f(x)} \circ D(\nabla f)|_x h \\ &= (\cdot)^T \circ D(\nabla f)|_x h = (\text{Hess } f(x)h)^T. \end{aligned} \quad (2.5.3)$$

As required, $D^2f|_x h$ is a “row vector”, i.e., an element of $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$. We see that if $\tilde{h} \in \mathbb{R}^n$ is some other vector, $D^2f|_x h$ acts on \tilde{h} via

$$(D^2f|_x h)\tilde{h} = (\text{Hess } f(x)h)^T \tilde{h} = \langle \text{Hess } f(x)h, \tilde{h} \rangle \in \mathbb{R}.$$

Note that the expression $\langle \text{Hess } f(x)h, \tilde{h} \rangle$ is linear in both h and \tilde{h} ; hence we can also regard the second derivative as a **bilinear map**

$$D^2f|_x: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}, \quad (h, \tilde{h}) \mapsto \langle \text{Hess } f(x)h, \tilde{h} \rangle.$$

The Second Derivative for General Functions

The preceding extended example already includes all relevant ideas for the general case, which we now discuss.

Let X, V be normed vector spaces, $\Omega \subset X$ open and $f: \Omega \rightarrow V$ a differentiable function. Then the derivative of f is a map

$$Df: \Omega \rightarrow \mathcal{L}(X, V), \quad x \mapsto Df|_x. \quad (2.5.4)$$

The derivative of Df (if it exists) is a map

$$D(Df) =: D^2f: \Omega \rightarrow \mathcal{L}(X, \mathcal{L}(X, V)).$$

We will investigate the space $\mathcal{L}(X, \mathcal{L}(X, V))$ a little more closely. Let $x_1, x_2 \in X$ and $L \in \mathcal{L}(X, \mathcal{L}(X, V))$. Then $Lx_1 \in \mathcal{L}(X, V)$ and

$$(Lx_1)(x_2) \in V.$$

The Second Derivative as a Bilinear Map

To $L \in \mathcal{L}(X, \mathcal{L}(X, V))$ we can associate a map $\tilde{L}: X \times X \rightarrow V$ given by

$$\tilde{L}(x_1, x_2) := (Lx_1)(x_2) \quad (2.5.5)$$

Obviously, for $x_1, x_2, x'_2 \in X$ and $\lambda \in \mathbb{F}$,

$$\begin{aligned} \tilde{L}(x_1, x_2 + x'_2) &= (Lx_1)(x_2 + x'_2) = (Lx_1)(x_2) + (Lx_1)(x'_2) \\ &= \tilde{L}(x_1, x_2) + \tilde{L}(x_1, x'_2), \end{aligned}$$

$$L(x_1, \lambda x_2) = (Lx_1)(\lambda x_2) = \lambda(Lx_1)(x_2) = \lambda \tilde{L}(x_1, x_2)$$

because $Lx_1 \in \mathcal{L}(X, V)$ is linear. Furthermore, since $L \in \mathcal{L}(X, \mathcal{L}(X, V))$,

$$\begin{aligned} \tilde{L}(x_1 + x'_1, x_2) &= (L(x_1 + x'_1))(x_2) = (Lx_1 + Lx'_1)(x_2) \\ &= (Lx_1)(x_2) + (Lx'_1)(x_2) = \tilde{L}(x_1, x_2) + \tilde{L}(x'_1, x_2), \end{aligned}$$

$$\tilde{L}(\lambda x_1, x_2) = (\lambda Lx_1)(x_2) = \lambda(Lx_1)(x_2) = \lambda \tilde{L}(x_1, x_2).$$

We thus see that \tilde{L} is a bilinear map, i.e., linear in both components.

Multilinear Maps

2.5.3. Definition. Let X, V be finite-dimensional normed vector spaces. The set of multilinear maps from X to V is denoted by

$$\mathcal{L}^{(n)}(X, V) := \left\{ L: X \times \cdots \times X \rightarrow V : L \text{ linear in each component} \right\}.$$

In the special case $V = \mathbb{R}$ an element of $\mathcal{L}^{(n)}(X, V)$ is called a **multilinear form**.

From the previous discussion we see that there is a canonical isomorphism

$$\mathcal{L}(X, \mathcal{L}(X, V)) \cong \mathcal{L}^{(2)}(X, V)$$

given by (2.5.5).

From now on, we will make no difference between these two spaces, and in fact drop the tilde in (2.5.5), treating L either as a bilinear map $X \times X \rightarrow V$ or as a linear map $X \rightarrow \mathcal{L}(X, V)$.

Bilinear Forms on \mathbb{R}^n

2.5.4. Example. Let $X = \mathbb{R}^n$ and $V = \mathbb{R}$. Then we have seen that

$$\mathcal{L}^{(2)}(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}) \cong \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R})).$$

We know that $\mathcal{L}(\mathbb{R}^n, \mathbb{R}) = (\mathbb{R}^n)^* \cong \mathbb{R}^n$, so we have

$$\mathcal{L}^{(2)}(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}) \cong \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) \cong \text{Mat}(n \times n, \mathbb{R}). \quad (2.5.6)$$

Thus the space of bilinear maps on \mathbb{R}^n is isomorphic to the set of square $n \times n$ matrices. How does this work in practice? Every linear map $L \in (\mathbb{R}^n)^*$ has the form

$$L = \langle z, \cdot \rangle \quad \text{for some } z \in \mathbb{R}^n.$$

Bilinear Forms on \mathbb{R}^n

We thus interpret an element of $A \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}))$ as a linear map that associates

$$A: y \mapsto L_y := \langle z_y, \cdot \rangle \quad (2.5.7)$$

Equivalently, we associate to every y some $z_y = A(y)$; this is realized through a matrix which we also call A :

$$A: y \mapsto z_y. \quad (2.5.8)$$

Hence, for every $y \in \mathbb{R}^n$ we obtain a linear map $\langle Ay, \cdot \rangle \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$. Letting this linear map act on some $x \in \mathbb{R}^n$ we get

$$L_y x = \langle Ay, x \rangle = L(x, y).$$

Often, one prefers to write $\langle x, Ay \rangle$ instead of $\langle Ay, x \rangle$. We thus see that

$$\text{Mat}(n \times n, \mathbb{R}) \cong \mathcal{L}^{(2)}(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}) \quad \text{via} \quad A \leftrightarrow \langle \cdot, A(\cdot) \rangle.$$

Schwarz's Theorem

If $f: \mathbb{R}^n \rightarrow \mathbb{R}$ is twice differentiable, we can represent D^2f in the form of a square $n \times n$ matrix; this is just the Hessian we have introduced in (2.5.2).

However, in general we can not represent the second derivative of a function $\mathbb{R}^n \rightarrow \mathbb{R}^m$ as a matrix; furthermore, even in the case of potential functions ($m = 1$) higher order derivatives can not be expressed as matrices.

We now introduce a fundamental result governing the second derivative:

2.5.5. Schwarz's Theorem. Let X, V be normed vector spaces and $\Omega \subset X$ an open set. Let $f \in C^2(\Omega, V)$. Then $D^2f|_x \in \mathcal{L}^{(2)}(X \times X, V)$ is symmetric for all $x \in \Omega$, i.e.,

$$D^2f(u, v) = D^2f(v, u), \quad \text{for all } u, v \in X.$$

Schwarz's Theorem

Proof.

Fix $x \in \Omega$ and choose $r > 0$ sufficiently small that $B_r(x) \subset \Omega$. Choose $u, v \in \Omega$ such that $\|u\|, \|v\| < r/2$. Set $g(x) := f(x + v) - f(x)$. Then the Mean Value Theorem 2.2.30 we have

$$\begin{aligned} & f(x + v + u) - f(x + u) - f(x + v) + f(x) \\ &= g(x + u) - g(x) = \int_0^1 Dg|_{x+tu} u \, dt \\ &= \int_0^1 (Df(x + tu + v) - Df(x + tu)) u \, dt \\ &= \int_0^1 \left(\int_0^1 D^2 f|_{x+sv+tu} v \, ds \right) u \, dt \end{aligned}$$

Schwarz's Theorem

Proof (continued).

Now the continuity of D^2f implies that

$$D^2f|_{x+sv+tu} - D^2f|_x = o(1) \quad \text{as } \|u\| + \|v\| \rightarrow 0.$$

for any $0 \leq s, t \leq 1$. In fact, the convergence is even uniform in s and t , i.e.,

$$\sup_{0 \leq s, t \leq 1} \|D^2f|_{x+sv+tu} - D^2f|_x\| = o(1) \quad \text{as } \|u\| + \|v\| \rightarrow 0$$

where we use the operator norm. This implies that

$$\int_0^1 \left(\int_0^1 D^2f|_{x+sv+tu} v \, ds \right) u \, dt = \int_0^1 \left(\int_0^1 D^2f|_x v \, ds \right) u \, dt + \|u\| \|v\| o(1).$$

as $\|u\| + \|v\| \rightarrow 0$.

Schwarz's Theorem

Proof (continued).

Again from the Mean Value Theorem 2.2.30 we have

$$\int_0^1 \left(\int_0^1 D^2 f|_x v \, ds \right) u \, dt = \int_0^1 \int_0^1 (D^2 f|_x v) u \, ds \, dt$$

where the two-fold integration on the right is known as an **iterated integral**. If we regard $D^2 f|_x$ as a bilinear map, we have, as $\|u\| + \|v\| \rightarrow 0$,

$$\begin{aligned} g(x+u) - g(x) &= \int_0^1 \int_0^1 D^2 f|_x(v, u) \, ds \, dt + \|u\|\|v\|o(1) \\ &= D^2 f|_x(v, u) + \|u\|\|v\|o(1) \end{aligned} \tag{2.5.9}$$

since the integrand does not depend on s or t .

Schwarz's Theorem

Proof (continued).

We may repeat this entire calculation, using $\tilde{g}(x) := f(x + u) - f(x)$ instead of g . We then obtain the same result, but with u and v interchanged:

$$g(x + v) - g(x) = D^2f|_x(u, v) + \|u\|\|v\|o(1) \quad (2.5.10)$$

as $\|u\| + \|v\| \rightarrow 0$. Now both (2.5.9) and (2.5.10) are equal to $f(x + v + u) - f(x + u) - f(x + v) + f(x)$, so taking the difference yields

$$D^2f|_x(v, u) - D^2f|_x(u, v) = \|u\|\|v\|o(1)$$

as $\|u\| + \|v\| \rightarrow 0$. Furthermore, we can now use a scaling argument to see that the right-hand side is actually zero. For this, note that the left-hand side may be regarded as a bilinear map $L \in \mathcal{L}(X \times X, V)$.

Schwarz's Theorem

Proof (continued).

We will show that if $L \in \mathcal{L}(X \times X, V)$ and

$$L(u, v) = \|u\|\|v\|o(1) \quad \text{as } \|u\| + \|v\| \rightarrow 0$$

then $L = 0$. Choose $s, t \in \mathbb{R} \setminus \{0\}$. Then

$$L(u, v) = \frac{1}{st} L(su, tv).$$

So for $|s| + |t| \rightarrow 0$ we have

$$\|L(u, v)\| = \frac{1}{|st|} \|L(su, tv)\| = o(1) \frac{1}{|st|} \|su\|\|tv\| = o(1) \|u\|\|v\|$$

as $|s| + |t| \rightarrow 0$. Since the left-hand side does not depend on s or t , we see that $L(u, v) = 0$. □

Symmetry of the Hessian

In the case of potential functions ($X = \mathbb{R}^n$, $V = \mathbb{R}$), Theorem 2.5.5 implies that

$$\langle \text{Hess } f(x)y, z \rangle = \langle \text{Hess } f(x)z, y \rangle, \quad x, y, z \in \mathbb{R}^n.$$

which means $\text{Hess } f(x) = (\text{Hess } f(x))^T$, i.e., the Hessian of f at x is a symmetric matrix. Writing out the components of $\text{Hess } f(x)$, this means that

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

In other words, if f is twice continuously differentiable, the order of differentiation in the second-order partial derivatives does not matter. (This will be the case if all second-order partial derivatives are continuous. Why?)

Convergence and Continuity

Functions and Derivatives

Curves in Vector Spaces

Potential Functions

The Second Derivative

Extrema of Potential Functions

Constrained Extrema

Extrema of Real Functions

In this section we will focus on extrema of functions. Recall that a twice continuously differentiable real function $f: \mathbb{R} \rightarrow \mathbb{R}$ satisfies

$$f(x+h) = f(x) + f'(x)h + \frac{f''(x)}{2}h^2 + o(h^2) \quad \text{as } h \rightarrow 0. \quad (2.6.1)$$

The stationary points of f are given by $f'(x) = 0$ and their nature is determined by the sign of $f''(x)$. We now aim to develop an analogous theory for functions $f: \mathbb{R}^n \rightarrow \mathbb{R}$. Our first goal will be to extend the formula

$$f(x+h) = f(x) + Df|_x h + o(h) \quad \text{as } h \rightarrow 0$$

into an expression analogous to (2.6.1).

Quadratic Approximation of Potential Functions

2.6.1. Lemma. Let $\Omega \subset \mathbb{R}^n$ be an open set and $f \in C^2(\Omega, \mathbb{R})$. Then for any $h \in \mathbb{R}^n$ small enough that $x + h \in \Omega$,

$$f(x + h) = f(x) + \langle \nabla f(x), h \rangle + \int_0^1 (1 - t) \langle \text{Hess } f(x + th)h, h \rangle dt. \quad (2.6.2)$$

Proof.

By the Mean Value Theorem 2.2.30,

$$f(x + h) - f(x) = \int_0^1 Df|_{x+th} h dt = \int_0^1 1 \cdot Df|_{x+th} h dt$$

We now want to integrate by parts, differentiating $Df|_{x+th} h$ and integrating 1 in the integrand. As a primitive for 1 we can take $t + c$ for any $c \in \mathbb{R}$; we choose $t - 1$.

Quadratic Approximation of Potential Functions

Proof (continued).

By the chain rule and (2.5.3),

$$\frac{d}{dt} Df|_{x+th} = D^2 f|_{x+th} \frac{d}{dt}(x + th) = (\text{Hess } f(x + th)h)^T.$$

Hence

$$\frac{d}{dt} Df|_{x+th} h = \langle \text{Hess } f(x + th)h, h \rangle.$$

Then

$$\begin{aligned} f(x + h) - f(x) &= (t - 1)Df|_{x+th} h \Big|_0^1 - \int_0^1 (t - 1) \cdot \langle \text{Hess } f(x + th)h, h \rangle dt \\ &= Df|_x h + \int_0^1 (1 - t) \cdot \langle \text{Hess } f(x + th)h, h \rangle dt \\ &= \langle \nabla f(x), h \rangle + \int_0^1 (1 - t) \langle \text{Hess } f(x + th)h, h \rangle dt. \end{aligned}$$
□

Quadratic Approximation of Potential Functions

2.6.2. Corollary. Let $\Omega \subset \mathbb{R}^n$ be an open set and $f \in C^2(\Omega, \mathbb{R})$. Then, as $h \rightarrow 0$,

$$f(x + h) = f(x) + \langle \nabla f(x), h \rangle + \frac{1}{2} \langle \text{Hess } f(x)h, h \rangle + o(h^2). \quad (2.6.3)$$

Proof.

In view of (2.6.2), we just need to show that

$$\int_0^1 (1-t) \langle \text{Hess } f(x+th)h, h \rangle dt = \frac{1}{2} \langle \text{Hess } f(x)h, h \rangle + o(h^2),$$

as $h \rightarrow 0$.

Quadratic Approximation of Potential Functions

Proof (continued).

We have

$$\begin{aligned} & \int_0^1 (1-t) \langle \text{Hess } f(x + th)h, h \rangle dt \\ &= \int_0^1 (1-t) \langle (\text{Hess } f(x + th) - \text{Hess } f(x))h, h \rangle dt \\ &\quad + \int_0^1 (1-t) \langle \text{Hess } f(x)h, h \rangle dt \\ &= \frac{1}{2} \langle \text{Hess } f(x)h, h \rangle + \int_0^1 (1-t) \langle (\text{Hess } f(x + th) - \text{Hess } f(x))h, h \rangle dt, \end{aligned}$$

so it remains to show

$$\int_0^1 (1-t) \langle (\text{Hess } f(x + th) - \text{Hess } f(x))h, h \rangle dt = o(h^2).$$

Quadratic Approximation of Potential Functions

Proof (continued).

We have

$$\begin{aligned} & \left| \int_0^1 (1-t) \langle (\text{Hess } f(x + th) - \text{Hess } f(x))h, h \rangle dt \right| \\ & \leq 1 \cdot \sup_{t \in [0,1]} |1-t| \cdot \sup_{t \in [0,1]} |\langle (\text{Hess } f(x + th) - \text{Hess } f(x))h, h \rangle| \\ & \leq \sup_{t \in [0,1]} \|(\text{Hess } f(x + th) - \text{Hess } f(x))\| \cdot \|h\|^2, \end{aligned}$$

where we have used the operator norm for the Hessian. In order to show the desired estimate, we need to establish

$$\lim_{h \rightarrow 0} \sup_{t \in [0,1]} \|\text{Hess } f(x + th) - \text{Hess } f(x)\| = 0.$$

Quadratic Approximation of Potential Functions

Proof (continued).

Since the second derivative of f is continuous, we have

$$\lim_{h \rightarrow 0} \|\text{Hess } f(x + th) - \text{Hess } f(x)\| = 0 \quad \text{for fixed } t.$$

For some $\delta > 0$ we have $B_\delta(x) \subset \Omega$ and for $\|h\| < \delta$ we have

$x + th \in \overline{B_\delta(x)}$ for all $t \in [0, 1]$. Now $\text{Hess } f$ is continuous on the compact set $\overline{B_\delta(x)}$ and hence also uniformly continuous there. This implies that for any $\varepsilon > 0$ one can find a $\delta > 0$ such that

$$\|x + th - x\| < \delta \quad \Rightarrow \quad \|\text{Hess } f(x + th) - \text{Hess } f(x)\| < \varepsilon.$$

Hence, choosing a δ for $t = 1$ guarantees that the same δ will work for any t , and the supremum converges to zero. □

Quadratic Forms

2.6.3. Definition. Let $A \in \text{Mat}(n \times n, \mathbb{R})$. Then the ***quadratic form induced by A*** is defined as the map

$$Q_A := \langle \cdot, A(\cdot) \rangle, \quad x \mapsto \langle x, Ax \rangle = \sum_{j,k=1}^n a_{jk} x_j x_k, \quad x \in \mathbb{R}^n.$$

Clearly, $Q_A(\lambda x) = \lambda^2 Q_A(x)$ for any $\lambda \in \mathbb{R}$. Note also that Q_A is continuous, because it is a polynomial in x_1, \dots, x_n .

Quadratic Forms

2.6.4. Definition. A quadratic form Q_A induced by a matrix $A \in \text{Mat}(n \times n, \mathbb{R})$ is called

- ▶ **positive definite** if $Q_A(x) > 0$ for all $x \neq 0$,
- ▶ **negative definite** if $Q_A(x) < 0$ for all $x \neq 0$,
- ▶ **indefinite** if $Q_A(x_0) > 0$ for some $x_0 \in \mathbb{R}^n$ and $Q_A(y_0) < 0$ for some $y_0 \in \mathbb{R}^n$.

A matrix A is said to be negative definite / positive definite / indefinite if the induced quadratic form Q_A has the corresponding property.

2.6.5. Remarks.

- ▶ It is easy to see that not all quadratic forms fall into one of the above three categories.
- ▶ If A is positive definite, then $-A$ is negative definite.

Quadratic Forms

2.6.6. Example. The matrix

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

is positive definite, since

$$\begin{aligned} Q_A(x) &= \left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\rangle \\ &= x_1(x_1 - 2x_2) + x_2(x_1 + x_2) \\ &= x_1^2 + x_2^2 - x_1x_2 = \frac{1}{2}(x_1^2 + x_2^2 + (x_1 - x_2)^2). \end{aligned}$$

This expression is strictly positive when $x \neq 0$.

Criteria for Definiteness

We will be mainly interested in the case $n = 2$, and here we can give an explicit criterion:

2.6.7. Lemma. Let $A \in \text{Mat}(2 \times 2, \mathbb{R})$ be symmetric, i.e.,

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}.$$

Let $\Delta = \det A$. Then

- (i) A positive definite $\Leftrightarrow a > 0$ and $\Delta > 0$
- (ii) A negative definite $\Leftrightarrow a < 0$ and $\Delta > 0$
- (iii) A indefinite $\Leftrightarrow \Delta < 0$

The proof will be part of the assignments.

Criteria for Positive Definiteness

For our analysis of the extrema of a function, the following criteria are essential:

2.6.8. Lemma. The matrix $A \in \text{Mat}(n \times n, \mathbb{R})$ is

$$\text{positive definite} \Leftrightarrow \exists_{\alpha > 0} \forall_{x \in \mathbb{R}^n} Q_A(x) \geq \alpha \|x\|^2$$

$$\text{negative definite} \Leftrightarrow \exists_{\alpha > 0} \forall_{x \in \mathbb{R}^n} Q_A(x) \leq -\alpha \|x\|^2$$

Proof.

If there exists some $\alpha > 0$ such that $Q_A(x) \geq \alpha \|x\|^2$ for all $x \in \mathbb{R}^n$, it is clear that $Q_A(x) > 0$ if $x \neq 0$, so A is positive definite.

Now let A be positive definite. In particular, $Q_A(x) > 0$ for $x \in S^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}$. Since S^{n-1} is closed and bounded, it is compact by Theorem 2.1.33.

Criteria for Definiteness

Proof (continued).

By Theorem 2.1.36, the minimum

$$\alpha := \min_{x \in S^{n-1}} Q_A(x)$$

exists and is strictly positive. Thus, for $x \neq 0$,

$$\frac{1}{\|x\|^2} Q_A(x) = Q_A\left(\frac{x}{\|x\|}\right) \geq \alpha$$

and so $Q_A(x) \geq \alpha \|x\|^2$. This is also trivially true for $x = 0$, so we have proven the statement for positive definite matrices.

The matrix A will be negative definite if and only if $-A$ is positive definite. Since $Q_{-A}(x) = -Q_A(x)$, the statement for negative definite matrices follows. □

Extrema of Real Functions

We state the obvious definitions:

2.6.9. Definition. Let $\Omega \subset \mathbb{R}^n$ and $f: \Omega \rightarrow \mathbb{R}$.

- (i) f is said to have a **(global) maximum** at $\xi \in \Omega$ if

$$x \in \Omega \quad \Rightarrow \quad f(x) \leq f(\xi).$$

- (ii) f is said to have a **strict (global) maximum** at $\xi \in \Omega$ if

$$x \in \Omega \setminus \{\xi\} \quad \Rightarrow \quad f(x) < f(\xi).$$

The function f is said to have a (strict) global minimum at ξ if $-f$ has a (strict) global maximum at ξ .

Extrema of Real Functions

2.6.10. Definition. Let $\Omega \subset \mathbb{R}^n$ and $f: \Omega \rightarrow \mathbb{R}$.

- (i) f is said to have a **local maximum** at $\xi \in \Omega$ if there exists a $\delta > 0$ such that

$$x \in \Omega \cap B_\delta(\xi) \quad \Rightarrow \quad f(x) \leq f(\xi).$$

- (ii) f is said to have a **strict local maximum** at $\xi \in \Omega$ if there exists a $\delta > 0$ such that

$$x \in \Omega \cap B_\delta(\xi) \setminus \{\xi\} \quad \Rightarrow \quad f(x) < f(\xi).$$

The function f is said to have a (strict) local minimum at ξ if $-f$ has a (strict) local maximum at ξ .

As usual, we will be able to deal with extrema at interior points of Ω using methods based on differentiation, while the boundary points of Ω must be considered separately.

Extrema of Real Functions

2.6.11. Theorem. Let $\Omega \subset \mathbb{R}^n$, $f: \Omega \rightarrow \mathbb{R}$ and $\xi \in \text{int } \Omega$. Assume that all partial derivatives of f exist at ξ and that f has a local extremum (maximum or minimum) in ξ . Then

$$\nabla f(\xi) = 0.$$

If f is differentiable in ξ , this implies $Df|_{\xi} = 0$.

Proof.

Let $\xi = (\xi_1, \dots, \xi_n)$. Define

$$g_1(x_1) := f(x_1, \xi_2, \dots, \xi_n).$$

Then g_1 has a local extremum at $x_1 = \xi_1$ and

$$0 = \frac{dg_1}{dx_1}\Big|_{x_1=\xi_1} = \frac{\partial f(x_1, \xi_2, \dots, \xi_n)}{\partial x_1}\Big|_{x_1=\xi_1} = \frac{\partial f}{\partial x_1}\Big|_{x=\xi}.$$

In the same way we see that all partial derivatives of f vanish at ξ . □

Extrema of Real Functions

2.6.12. Theorem. Let $\Omega \subset \mathbb{R}^n$ be open, $f \in C^2(\Omega)$ and $\xi \in \Omega$. Let $\nabla f(\xi) = 0$ (i.e., $Df|_{\xi} = 0$).

- (i) If $\text{Hess } f|_{\xi}$ is positive definite, f has a strict local minimum at ξ .
- (ii) If $\text{Hess } f|_{\xi}$ is negative definite, f has a strict local maximum at ξ .
- (iii) If $\text{Hess } f|_{\xi}$ is indefinite, f has no extremum at ξ .

Proof.

Denote by Q the quadratic form induced by $\text{Hess } f|_{\xi}$. Since $Df|_{\xi} = 0$, we see from (2.6.3) that

$$\frac{f(\xi + h) - f(\xi)}{\|h\|^2} = \frac{1}{2} Q\left(\frac{h}{\|h\|}\right) + \varrho(h) \quad \text{with } \lim_{h \rightarrow 0} \varrho(h) = 0. \quad (2.6.4)$$

Now let $\text{Hess } f|_{\xi}$ be positive definite. By Lemma 2.6.8 we can find $\alpha > 0$ such that $Q(h/\|h\|) \geq \alpha$ for all $h \neq 0$.

Extrema of Real Functions

Proof (continued).

For this α , we can find a $\delta > 0$ such that

1. $|\varrho(h)| < \alpha/2$ whenever $\|h\| < \delta$ and
2. $B_\delta(\xi) \subset \Omega$.

Now every $x \in B_\delta(\xi) \setminus \{\xi\}$ can be expressed as $x = \xi + h$ with $h := x - \xi \neq 0$, so we have

$$\begin{aligned} f(x) - f(\xi) &= \|h\|^2 \frac{f(\xi + h) - f(\xi)}{\|h\|^2} = \|h\|^2 \left(\frac{1}{2} Q \left(\frac{h}{\|h\|} \right) + \varrho(h) \right) \\ &\geq \|h\|^2 \left(\frac{\alpha}{2} - |\varrho(h)| \right) > 0 \end{aligned}$$

for all $x \in B_\delta(\xi) \setminus \{\xi\}$. Thus $f(\xi)$ is a strict local minimum.

In the same way one sees that $f(\xi)$ is a strict local maximum if $\text{Hess } f|_{\xi}$ is negative definite.

Extrema of Real Functions

Proof (continued).

Now assume that $\text{Hess } f|_{\xi}$ is indefinite. Then there exist $h_0, k_0 \in \mathbb{R}^n$ such that

$$Q(h_0) > 0 \quad \text{and} \quad Q(k_0) < 0.$$

For all $\lambda \neq 0$ we then have

$$Q\left(\frac{\lambda h_0}{\|\lambda h_0\|}\right) = \frac{\lambda^2}{|\lambda|^2} Q\left(\frac{h_0}{\|h_0\|}\right) = \frac{1}{\|h_0\|^2} Q(h_0) =: \varepsilon_1 > 0$$

and similarly

$$Q\left(\frac{\lambda k_0}{\|\lambda k_0\|}\right) = \frac{1}{\|k_0\|^2} Q(k_0) =: -\varepsilon_2 < 0$$

Then we see from (2.6.4) that for sufficiently small $\lambda \neq 0$ we have $f(\xi + \lambda h_0) > f(\xi)$ and $f(\xi + \lambda k_0) < f(\xi)$, so f can not be a local extremum. □

Extrema of Real Functions

Applying Lemma 2.6.7, we have the following result:

2.6.13. Corollary. Let $\Omega \subset \mathbb{R}^2$ be open, $f \in C^2(\Omega)$ and $\xi \in \Omega$ with $\nabla f(\xi) = 0$. Set

$$\Delta := \det \text{Hess } f|_{\xi} = \left. \frac{\partial^2 f}{\partial x_1^2} \right|_{\xi} \left. \frac{\partial^2 f}{\partial x_2^2} \right|_{\xi} - \left(\left. \frac{\partial^2 f}{\partial x_1 \partial x_2} \right|_{\xi} \right)^2$$

Then $f(\xi)$ is

- ▶ a local minimum if $\left. \frac{\partial^2 f}{\partial x_1^2} \right|_{\xi} > 0$ and $\Delta > 0$,
- ▶ a local maximum if $\left. \frac{\partial^2 f}{\partial x_1^2} \right|_{\xi} < 0$ and $\Delta > 0$,
- ▶ no extremum if $\Delta < 0$.

Note that if $\Delta = 0$, Corollary 2.6.13 yields no information.

Finding Extrema

In searching for extrema of functions $f \in C^2(\Omega, \mathbb{R})$, we follow a four-step process:

1. Check for critical points $\xi \in \text{int } \Omega$, i.e., those where $Df|_{\xi} = 0$.
2. Use Theorem 2.6.12 or Corollary 2.6.13 to check which of the critical points is an extremum.
3. Check the boundary $\partial\Omega$ separately for local extrema.
4. Identify the global extrema. Any finite global extremum must also be a local extremum, so it will be included among those found above.

2.6.14. Example. Let $f(x, y) = x^3 + y^3 - 3xy$ on \mathbb{R}^2 . Then $\nabla f = 0$ gives

$$\frac{\partial f}{\partial x} = 3x^2 - 3y = 0, \quad \frac{\partial f}{\partial y} = 3y^2 - 3x = 0$$

or $x^2 = y$ and $y^2 = x$. The only two solutions to these equations are $(0, 0)$ and $(1, 1)$.

Finding Extrema

We have

$$\Delta(x, y) = f_{xx}f_{yy} - f_{xy}^2 = (6x)(6y) - (-3)^2 = 36xy - 9.$$

At $(0, 0)$, $\Delta < 0$ so there is no extremum at this point. At $(1, 1)$, $\Delta > 0$ and $f_{xx} = 6 > 0$, so this point corresponds to local minimum. Since $\Omega = \mathbb{R}^2$ is open, there are no other extrema.

Convergence and Continuity

Functions and Derivatives

Curves in Vector Spaces

Potential Functions

The Second Derivative

Extrema of Potential Functions

Constrained Extrema

Constrained Extrema

We are now interested in studying the extrema of a real function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ under constraints. This means that we want to find maxima and minima of f not in \mathbb{R}^n but in some subset $\Omega \subset \mathbb{R}^n$. This subset will be defined by one or more equations that satisfy suitable conditions.

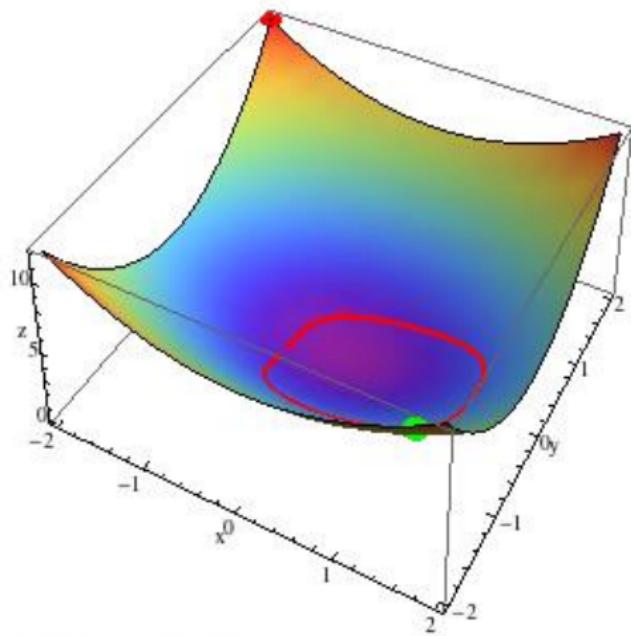
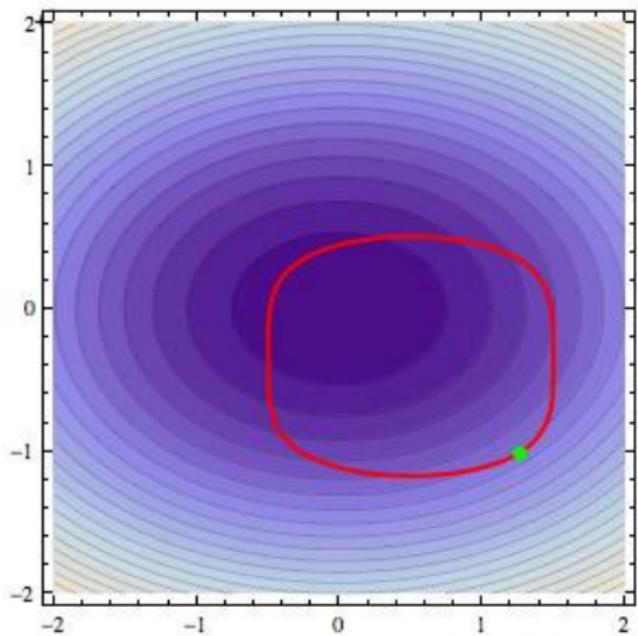
Before going into the technical details, let us consider the function $f(x_1, x_2) = x_1^2 + 2x_2^2$. It obviously has a global minimum at the origin and no local or global maximum. We will look for extrema of f when restricted to

$$E = \left\{ (x_1, x_2) \in \mathbb{R}^2 : \left(x_1 - \frac{1}{2} \right)^2 + 2 \left(x_2 + \frac{1}{3} \right)^4 = 1 \right\}$$

Since E is compact, we know that f will assume a maximum and a minimum value on E . From the following plot it will become clear that the extrema of f must lie on points of E that are tangent to the contours of f ,

Constrained Extrema

Below, the set E is represented by the red curve and the maximum of f is indicated by the green dot.



Constraint Sets and Solving Equations

In our example, E is defined through the equation $F: \mathbb{R}^2 \rightarrow \mathbb{R}$,

$$F(x_1, x_2) = \left(x_1 - \frac{1}{2}\right)^2 + 2\left(x_2 + \frac{1}{3}\right)^4 - 1 = 0.$$

Given a set

$$E = \{x \in \mathbb{R}^2 : F(x) = 0\}$$

we would like to find an expression for the tangent vector at a point ξ on E based on the function F .

A crucial approach is to represent E through the equation

$$x_2 = g(x_1)$$

for a suitable function $g: I \rightarrow \mathbb{R}$, $I \subset \mathbb{R}$ an interval, at least in a neighborhood of $\xi \in E$. Geometrically, this means representing E locally near ξ as the graph of a function. Algebraically, it means “solving” the equation $F(x_1, x_2) = 0$ for x_2 as a function of x_1 .

A First Example

2.7.1. Example. Consider the unit circle in \mathbb{R}^2 ,

$$S^1 = \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1\}.$$

When can points on this set be represented as the graph of a function g , i.e., in the form $x_2 = g(x_1)$? To be more precise, we are interested only in **local** representations.

Given $(\xi_1, \xi_2) \in S^1$, do there exist neighborhoods $B_\varepsilon(\xi_1)$, $B_\delta(\xi_2)$ and $g: B_\varepsilon(\xi_1) \rightarrow B_\delta(\xi_2)$ such that

$$x \in S^1 \cap (B_\varepsilon(\xi_1) \times B_\delta(\xi_2)) \Leftrightarrow x_2 = g(x_1)?$$

For the unit circle, we see that this is the case everywhere except at the points $(-1, 0)$ and $(1, 0)$, and the reason is geometrically clear: at these points, the tangent is perpendicular to the x_1 -axis.

Properties of the Implicitly Defined Function

Let us suppose that $F(x_1, x_2) = 0$, $x_1, x_2 \in \mathbb{R}$, is given and that in some x_1 -interval this equation can be solved for x_2 , yielding $x_2 = g(x_1)$, i.e.,

$$F(x_1, g(x_1)) = 0. \quad (2.7.1)$$

We might try to find the derivative of g ; it can be calculated by differentiating (2.7.1) with respect to x as follows:

$$0 = \frac{d}{dx_1} F(x_1, g(x_1)) = \partial_1 F(x_1, g(x_1)) + \partial_2 F(x_1, g(x_1)) \cdot g'(x_1)$$

so

$$g'(x_1) = -\frac{\partial_1 F(x_1, g(x_1))}{\partial_2 F(x_1, g(x_1))}$$

if F and g are differentiable and if $\partial_2 F(x) \neq 0$. It turns out that this condition is actually sufficient for the existence of the function g .

Returning to the Unit Circle

2.7.2. Example. Let $F(x_1, x_2) = x_1^2 + x_2^2 - 1$. Then

$$\frac{\partial F}{\partial x_2} = 2x_2$$

which is non-zero for $y \neq 0$. The function g satisfying $F(x_2, g(x_1)) = 0$ (if it exists) will then have a derivative given by

$$g'(x_1) = -\frac{\frac{\partial F}{\partial x_1}}{\frac{\partial F}{\partial x_2}} = -\frac{2x_1}{2x_2} = -\frac{x_1}{x_2}.$$

Of course, this can be verified explicitly by solving for x_2 to obtain

$$x_2 = g(x_1) = \pm\sqrt{1 - x_1^2}$$

and then differentiating.

The Implicit Function Theorem

What is surprising is that, given a continuously differentiable function $F: \mathbb{R}^2 \rightarrow \mathbb{R}$, the condition that

$$\frac{\partial F}{\partial x_2} \neq 0 \quad (2.7.2)$$

implies that

- (i) there exists a function $g: \mathbb{R} \rightarrow \mathbb{R}$ such that $F(x_1, g(x_1)) = 0$;
- (ii) the function g is continuously differentiable and

$$g'(x_1) = -\frac{\partial_1 F(x_1, g(x_1))}{\partial_2 F(x_1, g(x_1))}. \quad (2.7.3)$$

(Proving (2.7.3) if g is known to be differentiable is easy; establishing that g is differentiable in the first place is hard.)

These statements are known as the **Implicit Function Theorem**.

Constrained Extrema in \mathbb{R}^2

Let us return to the problem of constrained extrema. Let us consider first the case of a function $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ for which we wish to find extrema under the condition $F(x) = 0$, $F: \mathbb{R}^2 \rightarrow \mathbb{R}$. Suppose that $F(x_1, x_2) = 0$ can be solved for $x_2 = g(x_1)$ in some neighborhood of $\xi \in E$, i.e., we can write $F(x_1, g(x_1)) = 0$ for $x_1 \in I \subset \mathbb{R}$. Then E may be parametrized locally near ξ by

$$\gamma(x_1) = \begin{pmatrix} x_1 \\ g(x_1) \end{pmatrix}, \quad x_1 \in I,$$

and has tangent vector

$$\gamma'(x_1) = \begin{pmatrix} 1 \\ g'(x_1) \end{pmatrix} = \left(-\frac{\partial F}{\partial x_1} / \frac{\partial F}{\partial x_2} \right) = \frac{1}{\frac{\partial F}{\partial x_2}} \begin{pmatrix} \frac{\partial F}{\partial x_2} \\ -\frac{\partial F}{\partial x_1} \end{pmatrix}$$

where we have used (2.7.3).

Constrained Extrema in \mathbb{R}^2

The contour lines of f are given by $f(x) - c = 0$, $c \in \mathbb{R}$, and where they can be explicitly represented as functions of x_1 , their tangent vector is given by

$$\left(-\frac{\partial f}{\partial x_1} / \frac{\partial f}{\partial x_2} \right) = \frac{1}{\frac{\partial f}{\partial x_2}} \left(\begin{array}{c} \frac{\partial f}{\partial x_2} \\ -\frac{\partial f}{\partial x_1} \end{array} \right)$$

If the two tangent vectors are parallel at $\xi \in E$ (a necessary condition for an extremum to occur), there exists a number $\lambda \in \mathbb{R}$ such that

$$\frac{\partial f}{\partial x_1} + \lambda \frac{\partial F}{\partial x_1} = 0, \quad \frac{\partial f}{\partial x_2} + \lambda \frac{\partial F}{\partial x_2} = 0, \quad \lambda \in \mathbb{R}.$$

As a single equation, this may be expressed as

$$Df|_{\xi} + \lambda DF|_{\xi} = 0. \quad (2.7.4)$$

This number λ is called a **Lagrange multiplier**.

Constrained Extrema in \mathbb{R}^2

We can then find possible extremal points of f by solving the three (generally non-linear) equations

$$\begin{aligned}\frac{\partial f}{\partial x_1} + \lambda \frac{\partial F}{\partial x_1} &= 0, \\ \frac{\partial f}{\partial x_2} + \lambda \frac{\partial F}{\partial x_2} &= 0, \\ F(x_1, x_2) &= 0\end{aligned}$$

for x_1, x_2, λ . The number λ is unimportant to us (although there exist interpretations in certain physical contexts) but serves as a tool to find the possible extremal points.

In order for this method to work, (2.7.4) needs to hold. This requires being able to represent the constraint set E locally either in the form $x_2 = g(x_1)$ or $x_1 = g(x_2)$ at every point. This may not always be possible.

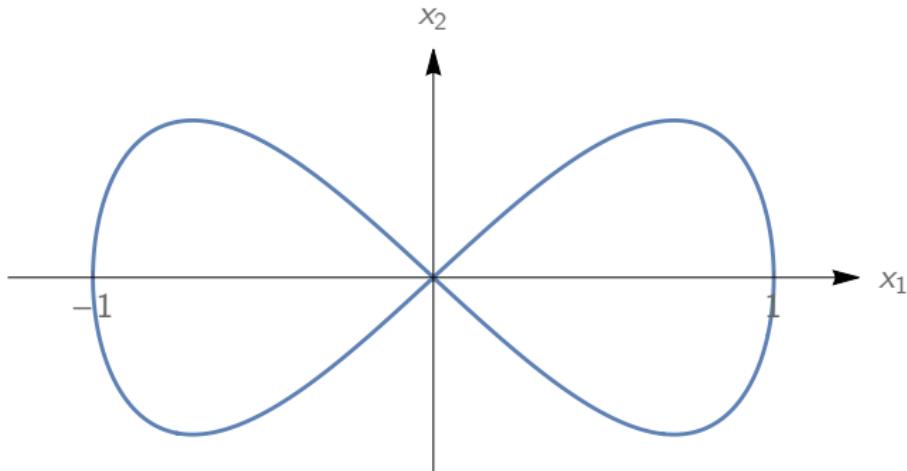
Lemniscate of Gerono

2.7.3. Example. Let $F: \mathbb{R}^2 \rightarrow \mathbb{R}$ be given by $F(x_1, x_2) = x_1^4 - x_1^2 + x_2^2$. Then the equation $F(x) = 0$ defines a curve called the **Lemniscate of Gerono**. We have

$$\frac{\partial F}{\partial x_1} = 4x_1^3 - 2x_1,$$

$$\frac{\partial F}{\partial x_2} = 2x_2,$$

and both derivatives vanish at $x = 0$.



Multiple Variables - One Constraint

Before we illustrate how to find extremal points using (2.7.4), we first generalize the discussion to higher dimensions $n > 2$ and multiple constraints. In such cases, the simple interpretation based on parallel tangent vectors becomes more complicated.

Given a single constraint $F(x_1, x_2) = 0$, we try to characterize the points $(x_1, x_2) \in \mathbb{R}^2$ satisfying this constraint by

$$x_2 = g_2(x_1) \quad (\text{if } \partial_2 F \neq 0) \quad \text{or} \quad x_1 = g_1(x_2) \quad (\text{if } \partial_1 F \neq 0).$$

Suppose that in \mathbb{R}^n a single constraint is given by $F(x_1, \dots, x_n) = 0$. Then we can solve (locally) for x_k as a function of the other variables,

$$x_k = g(x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n) \tag{2.7.5}$$

if $\partial_k F \neq 0$. The equation (2.7.5) determines the $x \in \mathbb{R}^n$ that satisfy the constraint $F(x) = 0$.

Multiple Variables - Multiple Constraints

More generally, we have multiple constraints,

$$F_1, \dots, F_m: \mathbb{R}^n \rightarrow \mathbb{R}, \quad F_1(x) = \dots = F_m(x) = 0.$$

In that case we would like to express m variables in \mathbb{R}^n as a function of $n - m =: k$ variables.

Suppose that these m variables are given by $(x_1, \dots, x_m) \in \mathbb{R}^m$ and set

$$x' = (x_1, \dots, x_m), \quad x'' = (x_{m+1}, \dots, x_n), \quad x = (x', x'').$$

We introduce a vector-valued constraint function

$$F: \mathbb{R}^{n-m} \times \mathbb{R}^m \rightarrow \mathbb{R}^m, \quad F(x', x'') = \begin{pmatrix} F_1(x', x'') \\ \vdots \\ F_m(x', x'') \end{pmatrix}.$$

Multiple Variables - Multiple Constraints

We denote by E the set of all points $(x', x'') \in \mathbb{R}^n$, $x' \in \mathbb{R}^k$, $x'' \in \mathbb{R}^m$, satisfying the m equations

$$F_1(x', x'') = 0, \quad \dots, \quad F_m(x', x'') = 0.$$

We wish to find m functions $g_1, \dots, g_m: \mathbb{R}^k \rightarrow \mathbb{R}$ such that the equations

$$x_1 = g_1(x''), \quad \dots, \quad x_m = g_m(x''),$$

determine E , i.e., we seek a function $g: \mathbb{R}^{n-m} \rightarrow \mathbb{R}^m$ such that $F(g(x''), x'') = 0$. A sufficient condition for this turns out to be

$$\det \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \dots & \frac{\partial F_1}{\partial x_m} \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial x_1} & \dots & \frac{\partial F_m}{\partial x_m} \end{pmatrix} \neq 0. \quad (2.7.6)$$

Multiple Variables - Multiple Constraints

2.7.4. Example. Suppose we have the constraints $F_1, F_2: \mathbb{R}^3 \rightarrow \mathbb{R}$ such that

$$F_1(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - 1, \quad F_2(x_1, x_2, x_3) = x_3 - x_2 - \frac{1}{2}.$$

Then we can define $F: \mathbb{R}^3 \rightarrow \mathbb{R}^2$

$$F(x_1, x_2, x_3) := \begin{pmatrix} x_1^2 + x_2^2 + x_3^2 - 1 \\ x_3 - x_2 - \frac{1}{2} \end{pmatrix}.$$

and attempt to solve $F(x) = 0$ for two variables as a function of the third.
[Why does this make sense?]

Multiple Variables - Multiple Constraints

Set $x' = (x_1, x_2)$, $x'' = x_3$. Then

$$\det DF(\cdot, x'') = \det \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} \\ \frac{\partial F_2}{\partial x_1} & \frac{\partial F_2}{\partial x_2} \end{pmatrix} = \det \begin{pmatrix} 2x_1 & 2x_2 \\ 0 & -1 \end{pmatrix} = -2x_1 \quad (2.7.7)$$

Then everywhere except at the two points

$$\left(0, -\frac{1}{4}(1 + \sqrt{7}), \frac{1}{4}(1 - \sqrt{7})\right) \quad \text{and} \quad \left(0, -\frac{1}{4}(1 - \sqrt{7}), \frac{1}{4}(1 + \sqrt{7})\right)$$

it will be possible to find a function $g(y) = (g_1(y), g_2(y))$ such that

$$F(g_1(x_3), g_2(x_3), x_3) = 0.$$

[Find g_1 and g_2 explicitly through elementary calculations!]

Multiple Variables - Multiple Constraints

Note that the function $F(\cdot, x'')$, for x'' fixed, is a map $\mathbb{R}^m \rightarrow \mathbb{R}^m$ and (2.7.6) can be expressed as $\det DF(\cdot, x'') \neq 0$.

We now finally formulate our result, the proof of which we omit for reasons of space.

2.7.5. Implicit Function Theorem. Let $\Omega \subset \mathbb{R}^n$ be open and $F \in C^p(\Omega, \mathbb{R}^m)$, $p \geq 1$, such that $F(\xi) = 0$ for some $\xi = (\xi', \xi'') \in \Omega$. Assume further that

$$\det DF(\cdot, \xi'')|_{\xi'} \neq 0.$$

Then there exists an open ball $B_\varepsilon(\xi'')$ and a C^p -map $g: B_\varepsilon(\xi'') \rightarrow \mathbb{R}^m$ such that

$$F(g(x''), x'') = 0 \quad \text{for all } x'' \in B_\varepsilon(\xi''). \quad (2.7.8)$$

Derivative of the Implicit Function

2.7.6. Remark. From the chain rule we can obtain the derivative of $g: B_\varepsilon(\xi'') \rightarrow \mathbb{R}^m$. Set

$$H: B_\varepsilon(\xi'') \rightarrow \mathbb{R}^{n-m}, \quad H(x'') := F(x'', g(x'')).$$

Define $h: \mathbb{R}^{n-m} \rightarrow \mathbb{R}^m \times \mathbb{R}^{n-m}$, $h(x) = \begin{pmatrix} g(x'') \\ x'' \end{pmatrix}$. Then $H = F \circ h$ and on $B_\varepsilon(\xi'')$ we have

$$\begin{aligned} 0 &= DH|_{x''} = DF|_{(g(x''), x'')} \cdot Dh|_{x''} \\ &= (DF(\cdot, x'')|_{g(x'')}, DF(g(x''), \cdot)|_{x''}) \begin{pmatrix} Dg|_{x''} \\ \mathbb{1}_{n-m} \end{pmatrix} \\ &= DF(\cdot, x'')|_{g(x'')} Dg|_{x''} + DF(g(x''), \cdot)|_{x''}. \end{aligned}$$

It follows that

$$Dg|_{x''} = -\left(DF(\cdot, x'')|_{g(x'')}\right)^{-1} DF(g(x''), \cdot)|_{x''} \quad (2.7.9)$$

The Implicit Function Theorem

2.7.7. Example. Continuing from the previous example,

$$DF(\cdot, x_3)|_{g(x_3)} = \det \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} \\ \frac{\partial F_2}{\partial x_1} & \frac{\partial F_2}{\partial x_2} \end{pmatrix} \Big|_{g(x_3)} = \begin{pmatrix} 2x_1 & 2x_2 \\ 0 & -1 \end{pmatrix} \Big|_{g(x_3)}$$

and

$$(DF(\cdot, x_3)|_{g(x_3)})^{-1} = \frac{1}{2x_1} \begin{pmatrix} 1 & 2x_2 \\ 0 & -2x_1 \end{pmatrix} \Big|_{g(x_3)}.$$

Furthermore,

$$DF(g(x_1), \cdot)|_{x_3} = \begin{pmatrix} 2x_3 \\ 1 \end{pmatrix}$$

so [verify this!]

$$Dg|_{x_1} = \begin{pmatrix} \frac{dg_1(x_3)}{dx_3} \\ \frac{dg_2(x_3)}{dx_3} \end{pmatrix} = \frac{1}{x_1} \begin{pmatrix} x_2 + x_3 \\ -x_1 \end{pmatrix} \Big|_{(x_1, x_2) = g(x_3)}.$$

The Lagrange Multiplier Rule

2.7.8. Lagrange Multiplier Rule. Let $\Omega \subset \mathbb{R}^n$ be open and $f \in C^1(\Omega, \mathbb{R})$, $F \in C^1(\Omega, \mathbb{R}^m)$, $m < n$. Assume that f has a local extremum on the set $E = \{x \in \mathbb{R}^n : F(x) = 0\}$ at $\xi \in E$. Assume further that in

$$DF|_{\xi} = \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \dots & \frac{\partial F_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial x_1} & \dots & \frac{\partial F_m}{\partial x_n} \end{pmatrix}$$

there exists a submatrix consisting of m columns whose determinant does not vanish. Then there exist m numbers (called **Lagrange multipliers**) $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ such that

$$Df|_{\xi} + \sum_{i=1}^m \lambda_i DF_i|_{\xi} = 0. \quad (2.7.10)$$

The Lagrange Multiplier Rule

Proof.

From our assumption we will assume without loss of generality that

$$\det \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \dots & \frac{\partial F_1}{\partial x_m} \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial x_1} & \dots & \frac{\partial F_m}{\partial x_m} \end{pmatrix} \neq 0$$

(If necessary, we renumber the variables x_1, \dots, x_n accordingly.) We will write

$$x' = (x_1, \dots, x_m), \quad x'' = (x_{m+1}, \dots, x_n), \quad x = (x', x''),$$

$$\xi' = (\xi_1, \dots, \xi_m), \quad \xi'' = (\xi_{m+1}, \dots, \xi_n), \quad \xi = (\xi', \xi''),$$

$$f(x) = f(x', x''), \quad F(x) = F(x', x'').$$

Thus $F(\xi', \xi'') = 0$ and $DF(\cdot, \xi'')|_{\xi'}$ is invertible.

The Lagrange Multiplier Rule

Proof (continued).

By the implicit function theorem, there exists an $\varepsilon > 0$ and a function $g: B_\varepsilon(\xi'') \rightarrow \mathbb{R}^m$ such that

$$g(\xi'') = \xi' \quad \text{and} \quad F(g(x''), x'') = 0 \quad \text{for } x \in B_\varepsilon(\xi'').$$

By (2.7.9) we moreover have

$$\begin{aligned} Dg|_{\xi''} &= -(DF(\cdot, \xi'')|_{g(\xi'')})^{-1} DF(\xi', \cdot)|_{\xi''} \\ &= -(DF(\cdot, \xi'')|_{\xi'})^{-1} DF(\xi', \cdot)|_{\xi''}. \end{aligned} \tag{2.7.11}$$

Now we consider extrema of f on $E = \{x: F(x) = 0\}$. From the above,

$$\varphi: B_\delta(\xi'') \rightarrow \mathbb{R}, \quad \varphi(x'') := f(g(x''), x''),$$

defined for sufficiently small $\delta > 0$, has an extremum at ξ'' if $f(\xi)$ is an extremum on E .

The Lagrange Multiplier Rule

Proof (continued).

Differentiating using the chain rule gives

$$\begin{aligned} D\varphi|_{\xi''} &= Df(\cdot, \xi'')|_{g(\xi'')} Dg|_{\xi''} + Df(g(\xi''), \cdot)|_{\xi''} \\ &= Df(\cdot, \xi'')|_{\xi'} Dg|_{\xi''} + Df(\xi', \cdot)|_{\xi''} \end{aligned}$$

At ξ'' the function φ has an extremum, so the derivative vanishes, yielding

$$0 = Df(\cdot, \xi'')|_{\xi'} Dg|_{\xi''} + Df(\xi', \cdot)|_{\xi''}.$$

Inserting (2.7.11), we have

$$0 = -Df(\cdot, \xi'')|_{\xi'} (DF(\cdot, \xi'')|_{\xi'})^{-1} DF(\xi', \cdot)|_{\xi''} + Df(\xi', \cdot)|_{\xi''} \quad (2.7.12)$$

The Lagrange Multiplier Rule

Proof (continued).

Define the row vector $\Lambda \in (\mathbb{R}^m)^*$ by

$$\Lambda = (\lambda_1, \dots, \lambda_m) := -Df(\cdot, \xi'')|_{\xi'} (DF(\cdot, \xi'')|_{\xi'})^{-1}$$

Then, trivially,

$$Df(\cdot, \xi'')|_{\xi'} + \Lambda \cdot DF(\cdot, \xi'')|_{\xi'} = 0. \quad (2.7.13)$$

On the other hand, Λ can be inserted into (2.7.12) to yield

$$Df(\xi', \cdot)|_{\xi''} + \Lambda \cdot DF(\xi', \cdot)|_{\xi''} = 0. \quad (2.7.14)$$

Together, (2.7.13) and (2.7.14) give (2.7.10). □

Practical Determination of Constrained Extrema

One of the characteristic properties of constrained extremal problems is that it is often not necessary to calculate the second derivative to determine the nature of the extremum. This is the case, because in most applications one of two situations occur:

1. The constraint set $E = \{F(x) = 0\}$ is compact. Then there must exist a maximum and a minimum by Theorem 2.1.36.
2. The problem is one of finding the distance between a compact and a closed set. Then the distance between the sets is found by minimizing the distances between all points and this minimum exists.

In both cases, the strategy is to find all candidates for extremal points and evaluate the values of the function on these points. The largest value will be the maximum, the smallest value will be the minimum.

Practical Determination of Constrained Extrema

A necessary condition for an extremal point of f under $g = 0$ is that it must solve the $m + n$ equations

$$\begin{aligned}\frac{\partial f}{\partial x_i} + \lambda_1 \frac{\partial F_1}{\partial x_i} + \cdots + \lambda_m \frac{\partial F_m}{\partial x_i} &= 0, & i &= 1, \dots, n, \\ F_j(x) &= 0, & j &= 1, \dots, m.\end{aligned}$$

These equations are equivalent to the following: define

$$\Psi(x_1, \dots, x_n, \lambda_1, \dots, \lambda_m) := f(x) + \lambda_1 F_1(x) + \dots + \lambda_m F_m(x).$$

Then at an extremal point all partial derivatives of Ψ will vanish.

Practical Determination of Constrained Extrema

2.7.9. Example. Find the point on the plane given by $z = x + y$ in \mathbb{R}^3 that has the smallest distance from the point $(1, 0, 0)$.

We want to minimize the distance under the constraint of being on the plane. For simplicity, we minimize the **square** of the distance function, i.e.,

$$f(x, y, z) = (x - 1)^2 + y^2 + z^2.$$

The constraint condition is $F(x, y, z) = x + y - z = 0$. We see that in

$$DF|_{(x,y,z)} = (1, 1, -1)$$

there is trivially always a 1×1 submatrix with determinant different from zero, so we can apply the Lagrange multiplier rule. We define

$$\psi(x, y, z, \lambda) = (x - 1)^2 + y^2 + z^2 + \lambda(x + y - z).$$

Practical Determination of Constrained Extrema

We then have the system of equations

$$\frac{\partial \psi}{\partial x} = 2(x - 1) + \lambda = 0,$$

$$\frac{\partial \psi}{\partial y} = 2y + \lambda = 0,$$

$$\frac{\partial \psi}{\partial z} = 2z - \lambda = 0,$$

$$\frac{\partial \psi}{\partial \lambda} = x + y - z = 0.$$

from which we obtain $(x, y, z) = (2/3, -1/3, 1/3)$. Thus only this point can minimize of the distance. Since the plane is a closed set and $\{(1, 0, 0)\}$ is a compact set, we know that a minimum exists. Therefore, $p = (2/3, -1/3, 1/3)$ is this point.

Practical Determination of Constrained Extrema

2.7.10. Example. Find the minimum and the maximum of the function $f(x, y, z) = 5x + y - 3z$ on the intersection of the plane given by $x + y + z = 0$ and the sphere $x^2 + y^2 + z^2 = 1$.

Here we have two constraints, which we formulate as

$$F(x, y, z) = \begin{pmatrix} x + y + z \\ x^2 + y^2 + z^2 - 1 \end{pmatrix} = 0.$$

Then

$$DF|_{(x,y,z)} = \begin{pmatrix} 1 & 1 & 1 \\ 2x & 2y & 2z \end{pmatrix}. \quad (2.7.15)$$

Now for any point (x, y, z) we require that at least one a 2×2 submatrix of Dg will not have determinant zero, i.e., one of

$$\begin{pmatrix} 1 & 1 \\ 2x & 2y \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ 2y & 2z \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ 2x & 2z \end{pmatrix}$$

must be invertible at any point under consideration.

Practical Determination of Constrained Extrema

If all determinants vanish, then $x = y = z$. Our constraint set is a subset of the plane given by $x + y + z = 0$, so this would imply $x = y = z = 0$. However, it is also a subset of the unit sphere, i.e., we are only interested in points where $x^2 + y^2 + z^2 = 1$, so the origin is not included. Thus, for every point under consideration, there will always be an invertible 2×2 submatrix of (2.7.15).

This allows us to apply the Lagrange multiplier rule, and we define

$$\Psi(x, y, z, \lambda, \mu) = 5x + y - 3z + \lambda(x + y + z) + \mu(x^2 + y^2 + z^2 - 1).$$

We then solve

$$\frac{\partial \Psi}{\partial x} = 5 + \lambda + 2\mu x = 0,$$

$$\frac{\partial \Psi}{\partial \lambda} = x + y + z = 0,$$

$$\frac{\partial \Psi}{\partial y} = 1 + \lambda + 2\mu y = 0,$$

$$\frac{\partial \Psi}{\partial \mu} = x^2 + y^2 + z^2 - 1 = 0.$$

$$\frac{\partial \Psi}{\partial z} = -3 - \lambda + 2\mu z = 0,$$

Practical Determination of Constrained Extrema

This system of equations yields two solutions (check for yourself), namely the points

$$p_1 = \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}} \right) \quad \text{and} \quad p_2 = \left(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right).$$

We find that $f(p_1) = 4\sqrt{2}$ and $f(p_2) = -4\sqrt{2}$. Since our constraint set is compact, f must assume a minimum and a maximum there, so $f(p_1)$ is the maximum and $f(p_2)$ the minimum.

Second Midterm Exam

The preceding material completes the first third of the course material. It encompasses everything that will be the subject of the **Second Midterm Exam.**

The exam date will be posted on Canvas.

No calculators or other aids will be permitted during the exam. A sample exam with solutions has been uploaded to Canvas. Please study it carefully, including the instructions on the cover page.

Part III

Vector Fields and Integral Calculus

Vector Fields and Line Integrals

Circulation and Flux

The Riemann Integral and Measurable Sets

Integration in Practice

Surfaces and Surface Integrals

The Theorems of Gauß and Stokes

Vector Fields and Line Integrals

Circulation and Flux

The Riemann Integral and Measurable Sets

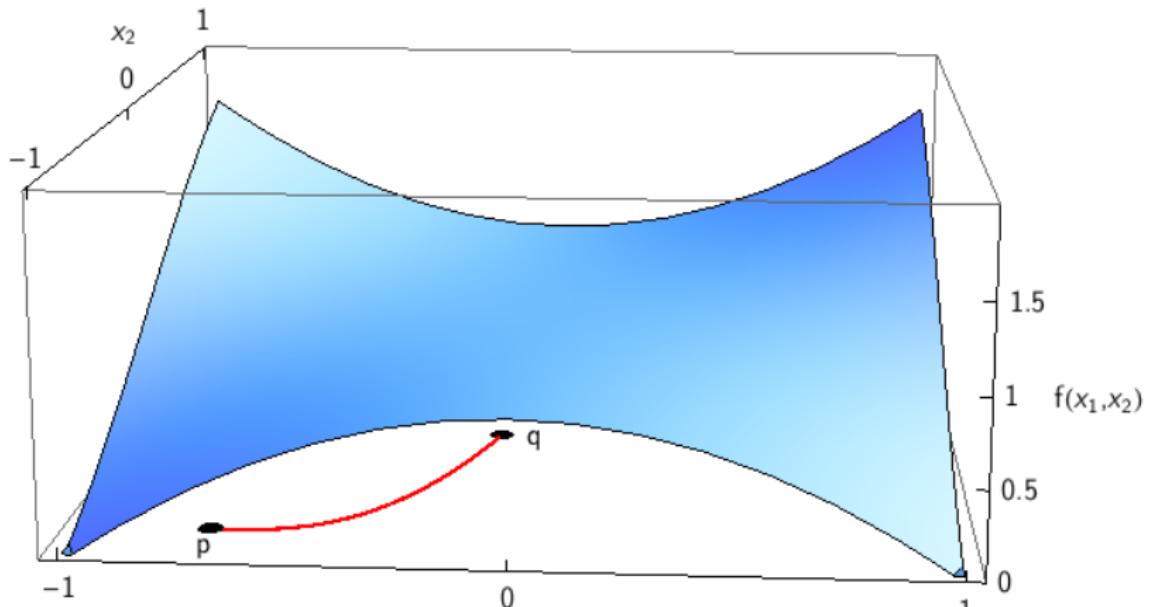
Integration in Practice

Surfaces and Surface Integrals

The Theorems of Gauß and Stokes

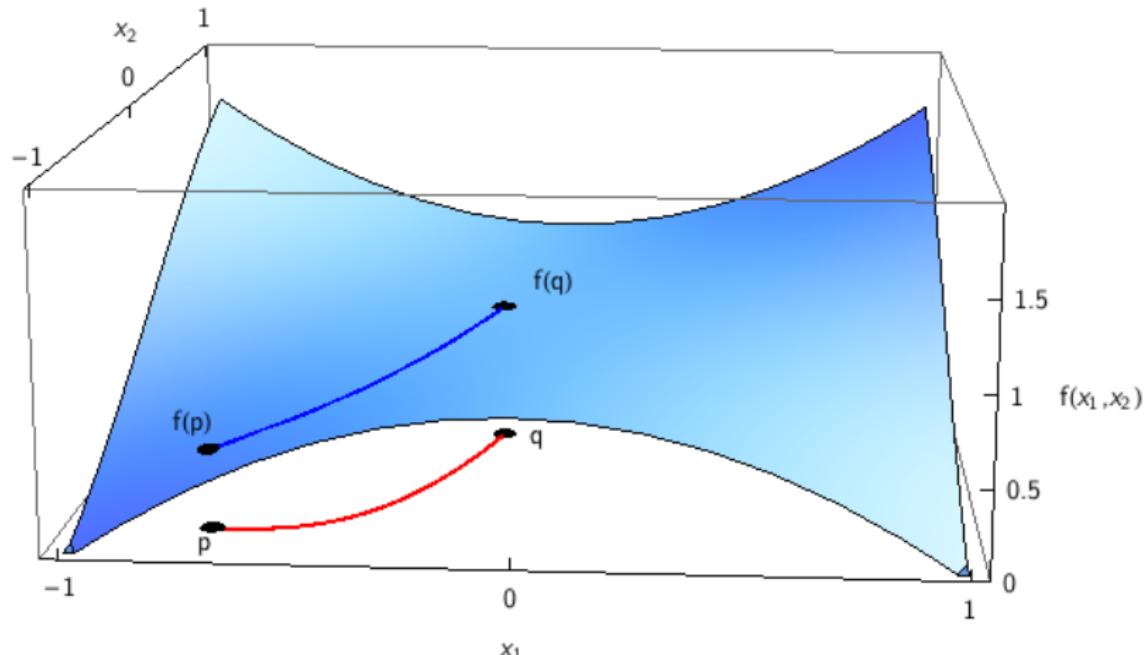
The Line Integral of a Potential Function

Suppose that we are given a simple, open, oriented curve $\mathcal{C}^* \subset \mathbb{R}^2$ and a scalar function $f: \mathbb{R}^2 \rightarrow \mathbb{R}$. In the sketch below, the red curve \mathcal{C}^* joins the points p and q in the x_1 - x_2 -plane, and the function f is given by $f(x_1, x_2) = 4/5 + x_1^2 \sin x_2$.



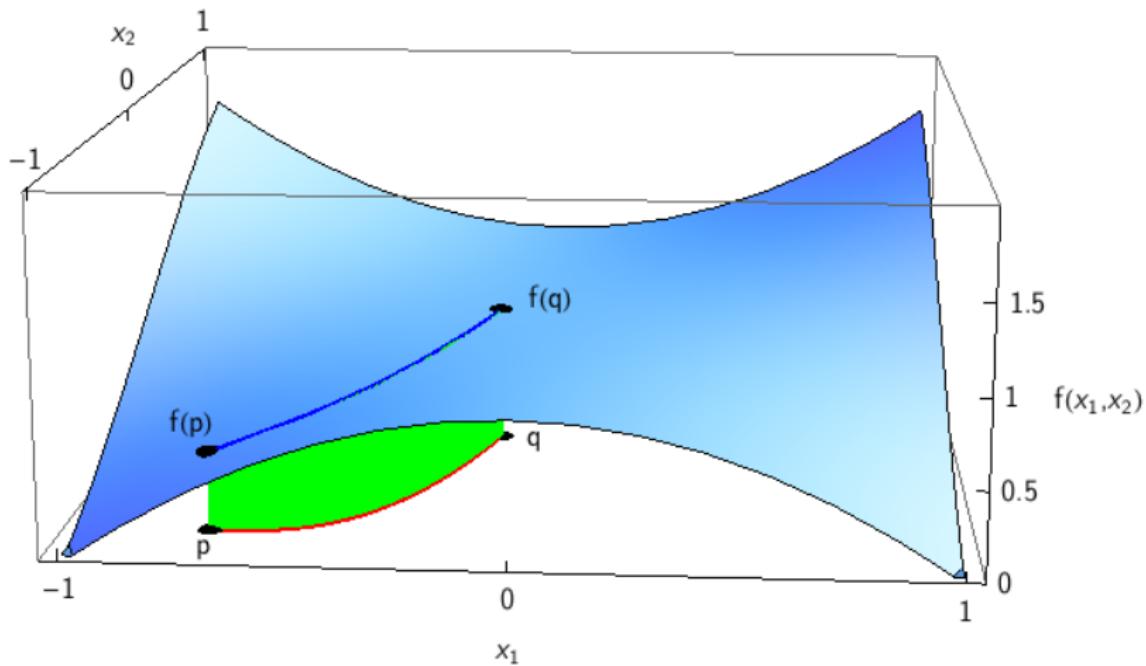
The Line Integral of a Potential Function

Suppose that \mathcal{C}^* is parametrized by a function $\gamma: [a, b] \rightarrow \mathcal{C}$ such that $\gamma(a) = p$ and $\gamma(b) = q$. Then the blue curve below shows the values of $f \circ \gamma: [a, b] \rightarrow \mathbb{R}$.



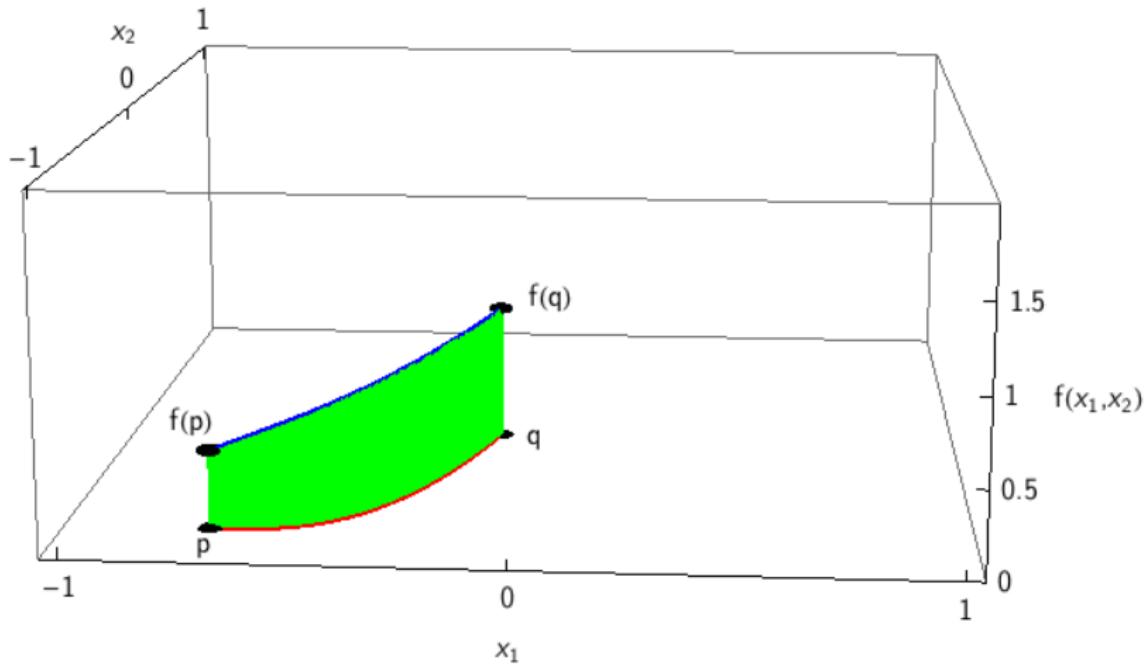
The Line Integral of a Potential Function

We now want to integrate the values of f along the red curve, i.e., we will determine the area of the green surface.



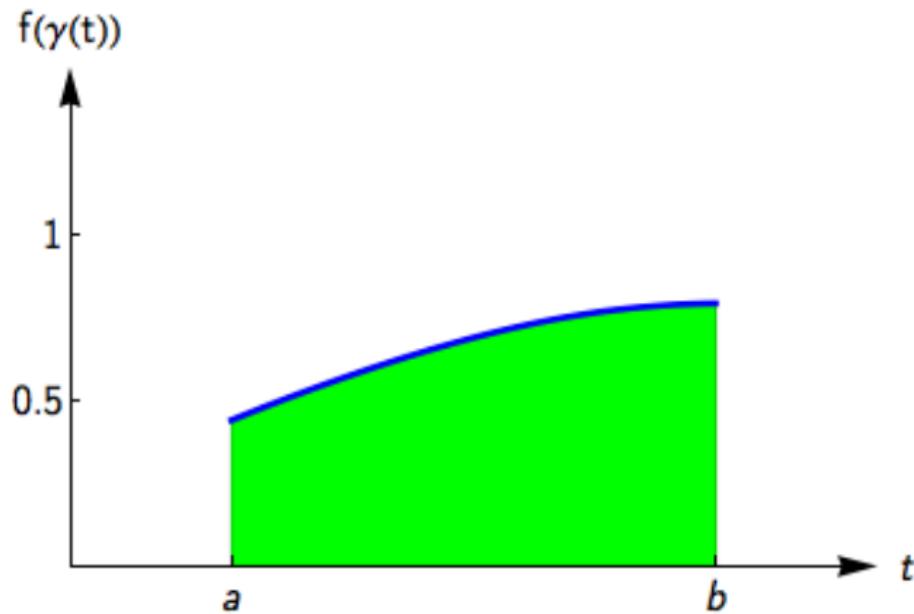
The Line Integral of a Potential Function

For clarity, the graph of f has been removed in the sketch below.



The Line Integral of a Potential Function

By considering the composition $f \circ \gamma$, we are effectively “straightening out” the red curve to the interval $[a, b]$.



The Line Integral of a Potential Function

3.1.1. Definition. Let $\Omega \subset \mathbb{R}^n$, $f: \Omega \rightarrow \mathbb{R}$ be a continuous potential function and $\mathcal{C}^* \subset \Omega$ an oriented smooth curve with parametrization $\gamma: I \rightarrow \mathcal{C}$. We then define the **line integral of the potential f along \mathcal{C}^*** by

$$\int_{\mathcal{C}^*} f \, ds := \int_I (f \circ \gamma)(t) \cdot |\gamma'(t)| \, dt$$

3.1.2. Remarks.

- ▶ Using the chain rule it can easily be seen that this integral is independent of the parametrization of \mathcal{C}^* .
- ▶ The line integral of a piecewise-smooth curve is defined as the sum of the integrals of the individual smooth segments.
- ▶ The symbol “ ds ” is, strictly speaking, unnecessary decoration. However, it can be interpreted geometrically as a **scalar line element** and one often writes

$$ds = |\gamma'(t)| \, dt,$$

but this should not be understood in any strict mathematical sense.

The Line Integral of a Potential Function

The potential doesn't really need to be defined on an open set; it is sufficient for it to be defined on \mathcal{C} .

3.1.3. Example. The mass of a physical wire (interpreted as a curve; i.e., having no thickness) can be obtained by integrating its density along its path. If a wire \mathcal{C} is taken to have variable density ϱ its mass is given by

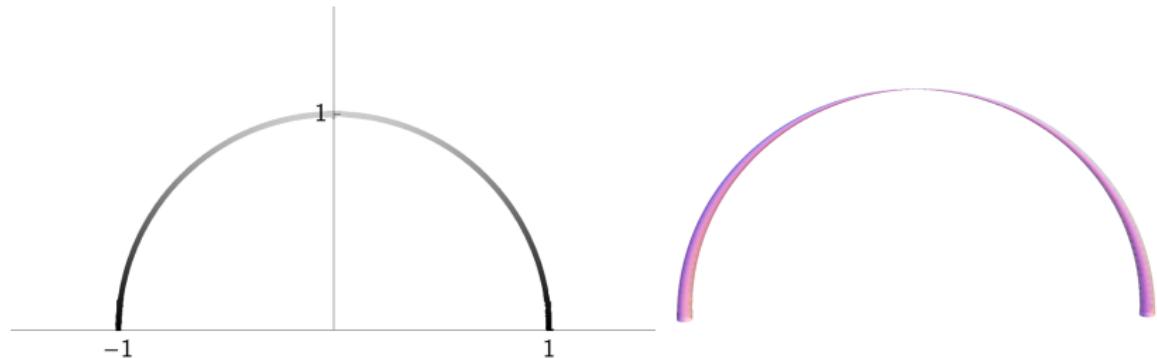
$$m = \left| \int_{\mathcal{C}} \varrho \, ds \right|.$$

As an example, consider a semi-circular wire

$$\mathcal{C} = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1, y \geq 0\}$$

with density $\varrho(x, y) = k(1 - y)$ where $k > 0$ is a constant. (Thus the wire is denser at its base and light at the top. We might alternatively interpret the varying density as varying thickness of a uniformly dense wire.)

Total Mass of a Wire



We choose the parametrization $\gamma(t) = (\cos t, \sin t)$, $I = [0, \pi]$. We have

$$\int_C \varrho \, ds = \int_0^\pi \varrho \circ \gamma(t) \cdot \|\gamma'(t)\| \, dt = \int_0^\pi k(1 - \sin t) \cdot 1 \, dt = k(\pi - 2),$$

so $m = |k(\pi - 2)| = k(\pi - 2)$.

Center of Mass of a Wire

The center of mass of the wire is given by (x_c, y_c) , where

$$x_c := \frac{1}{m} \int_{\mathcal{C}} x \cdot \varrho \, ds, \quad y_c := \frac{1}{m} \int_{\mathcal{C}} y \cdot \varrho \, ds.$$

(of course, an analogous formula holds for objects represented as one-dimensional curves in \mathbb{R}^n).

In our example,

$$x_c = \frac{1}{m} \int_0^1 (x \cdot \varrho) \circ \gamma(t) \, dt = \frac{1}{m} \int_0^\pi \cos t \cdot k(1 - \sin t) \, dt = 0$$

$$y_c = \frac{1}{m} \int_0^\pi \sin t \cdot k(1 - \sin t) \, dt = \frac{4 - \pi}{2(\pi - 2)}$$

Vector Fields

We now turn to a very important type of map, the vector field. Vector fields play an extremely important role in physics and mathematics. Examples include the flow field of a fluid or the electromagnetic field induced by a charge.

3.1.4. Definition. Let $\Omega \subset \mathbb{R}^n$. Then a function $F: \Omega \rightarrow \mathbb{R}^n$,

$$F(x) = \begin{pmatrix} F_1(x) \\ \vdots \\ F_n(x) \end{pmatrix}.$$

is called a **vector field** on Ω .

3.1.5. Example. Let $f: \mathbb{R}^n \rightarrow \mathbb{R}$ be a potential function. Then the **gradient field of f** given by

$$F: \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad F(x) = \nabla f(x),$$

associates to every $x \in \mathbb{R}^n$ the direction of largest slope of f .

Force Fields

3.1.6. Example. A mass M situated at the origin of a coordinate system exerts an attractive force on another mass m at position $x \in \mathbb{R}^3 \setminus \{0\}$. This **force field** is given by

$$F: \mathbb{R}^3 \setminus \{0\} \rightarrow \mathbb{R}^3, \quad F(x) = -G \frac{m \cdot M}{|x|^2} \frac{x}{|x|}, \quad (3.1.1)$$

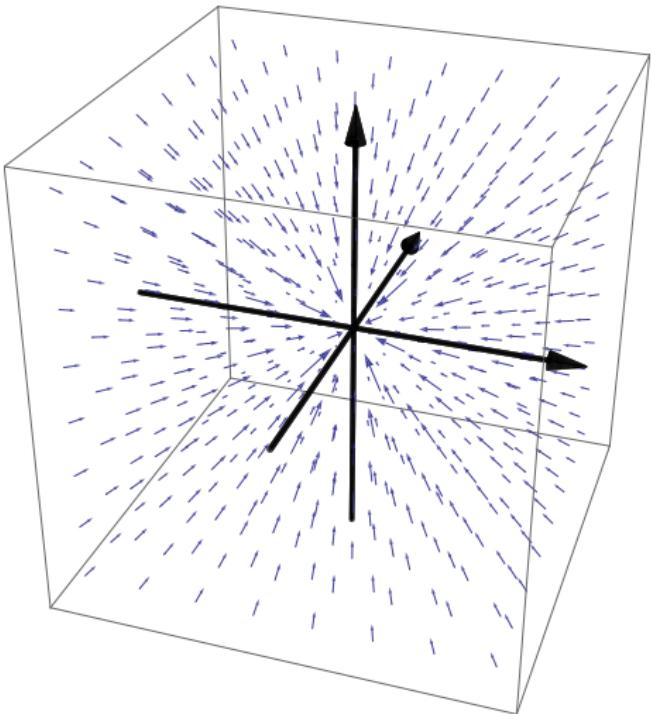
where G is Newton's gravitational constant.

Any vector field that associates to each $x \in \mathbb{R}^n$ a physical force vector is said to be a **force field**. (This term of course has only physical, not mathematical, significance.)

In physics, the concept of **work** arises from the integration of the forces acting along a particle's trajectory (curve), where force that are orthogonal to the trajectory do not contribute to the work. In particular, the work is obtained by integrating only the tangential components of the force field.

Gravitational Force Field

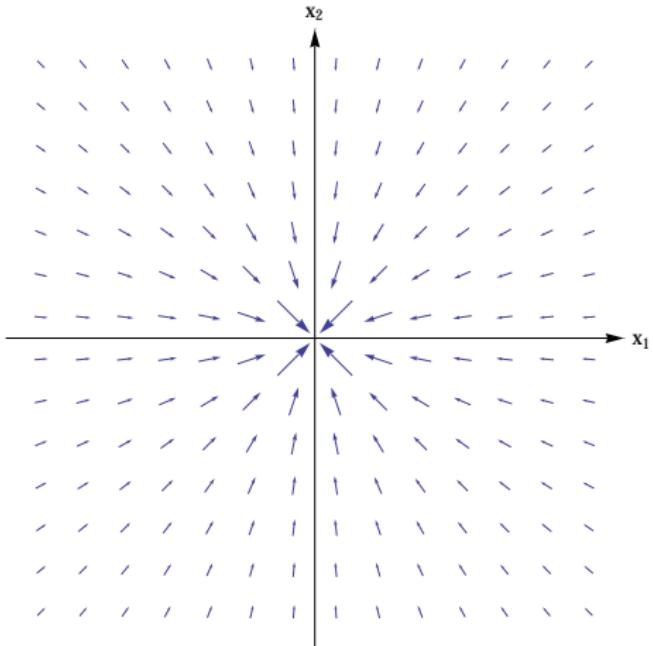
The plot below shows the gravitational force field (3.1.1) by attaching a vector representing $F(x)$ to each $x \in \mathbb{R}^3 \setminus \{0\}$



Gravitational Force Field

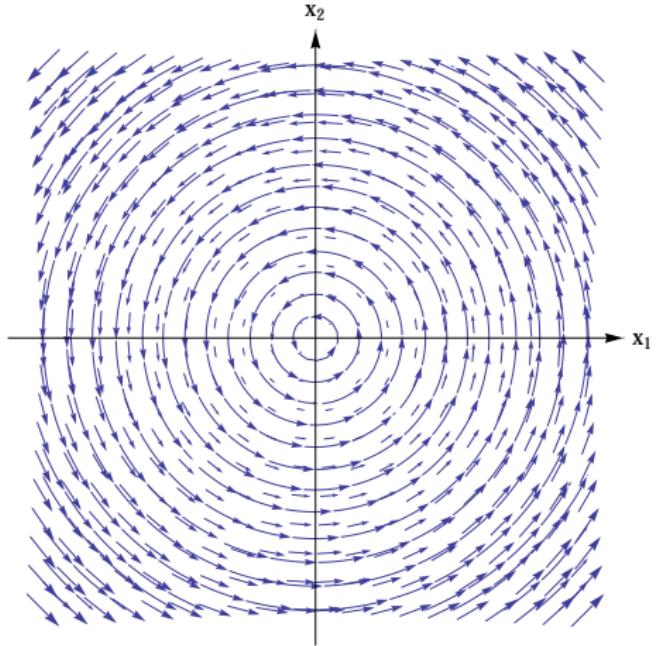
For future examples we will use the two-dimensional version,

$$F: \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}^2, \quad F(x) = -G \frac{m \cdot M}{|x|^2} \frac{x}{|x|}, \quad (3.1.2)$$



Streamlines of Fluid Flow

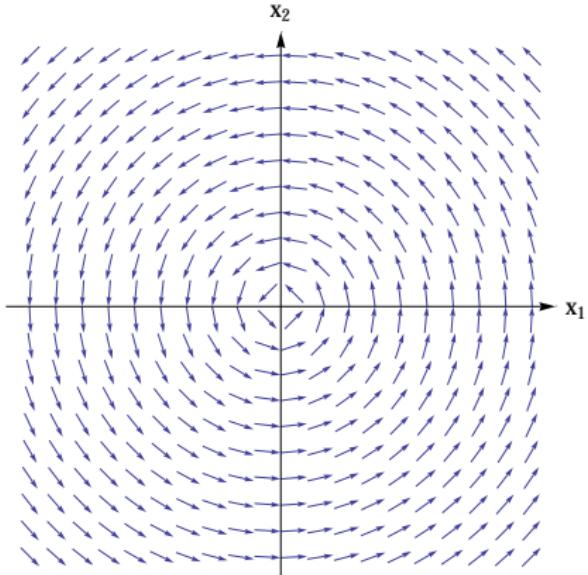
3.1.7. Example. Consider a fluid flow in \mathbb{R}^2 where the fluid rotates about the origin in a counter-clockwise manner. The **streamlines** show the paths of a “fluid particle”:



Direction Field of Fluid Flow

The streamlines are circles and the unit tangent vector field (the **direction field**) of the circles is given by

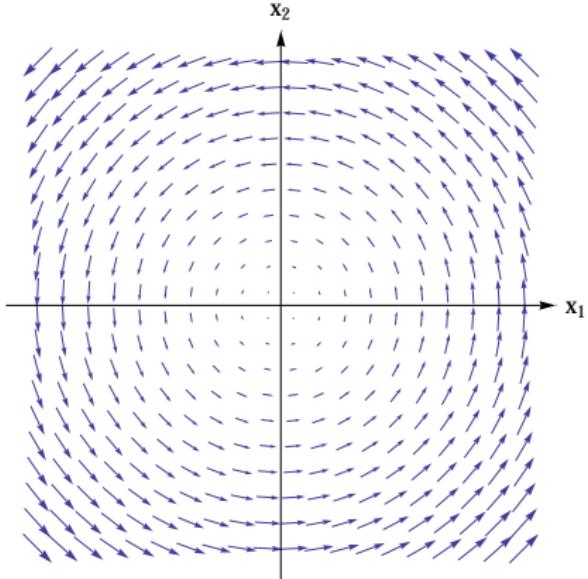
$$F: \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}^2, \quad F(x_1, x_2) = \frac{1}{\sqrt{x_1^2 + x_2^2}} \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}. \quad (3.1.3)$$



Velocity Field of Fluid Flow

The velocity at a distance $r = \sqrt{x_1^2 + x_2^2}$ from the origin is $r \cdot \omega$, where ω is the rotational velocity. Hence, the velocity vector field is given by

$$v: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad v(x_1, x_2) = r\omega F(x_1, x_2) = \omega \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}. \quad (3.1.4)$$



The Line Integral of a Vector Field

3.1.8. Definition. Let $\Omega \subset \mathbb{R}^n$, $F: \Omega \rightarrow \mathbb{R}$ be a continuous vector field and $\mathcal{C}^* \subset \Omega$ an oriented open, smooth curve in \mathbb{R}^n . We then define the **line integral of the vector field F along \mathcal{C}^*** by

$$\int_{\mathcal{C}^*} F d\vec{s} := \int_{\mathcal{C}^*} \langle F, T \rangle ds \quad (3.1.5)$$

3.1.9. Remarks.

- (i) We have defined the line integral of the vector field F as the line integral of the scalar product $\langle F, T \rangle$ on \mathcal{C}^* . Since T does not depend on the parametrization of \mathcal{C}^* and the line integral of a scalar function doesn't either, the line integral of a vector field is independent of parametrization of \mathcal{C}^* .

The Line Integral of a Vector Field

- (ii) The symbol “ $d\vec{s}$ ” can be interpreted geometrically as a ***vectorial line element*** and one often writes

$$d\vec{s} = \gamma'(t) dt.$$

In the same spirit, one sometimes writes

$$\int_{\mathcal{C}^*} F d\vec{s} = \int_{\mathcal{C}^*} \langle F, d\vec{s} \rangle$$

- (iii) Integrals along closed curves are sometimes emphasized by writing

$$\oint_{\mathcal{C}^*} f ds \quad \text{or} \quad \oint_{\mathcal{C}^*} F d\vec{s}$$

if the curve \mathcal{C} is closed.

Integrals of Vector Fields

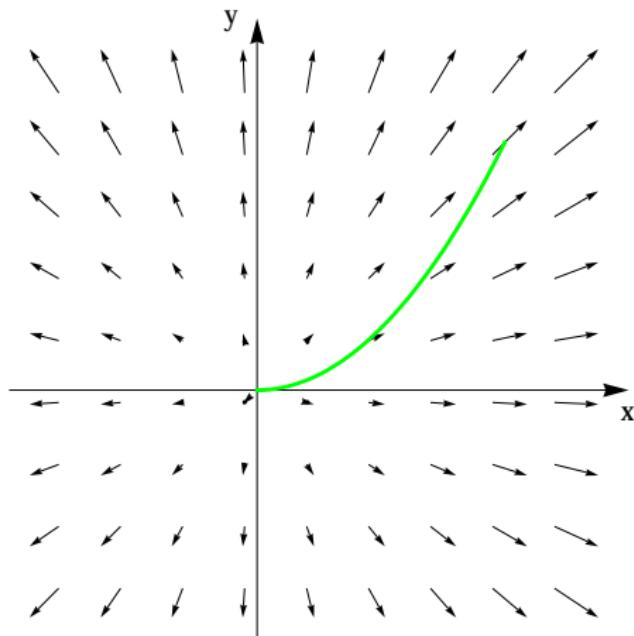
If we calculate the line integral using a concrete parametrization $\gamma: I \rightarrow \mathcal{C}$, we obtain

$$\begin{aligned}
 \int_{\mathcal{C}^*} F d\vec{s} &= \int_{\mathcal{C}^*} \langle F, T \rangle ds = \int_I \langle F \circ \gamma(t), T \circ \gamma(t) \rangle \|\gamma'(t)\| dt \\
 &= \int_I \left\langle F \circ \gamma(t), \frac{\gamma'(t)}{\|\gamma'(t)\|} \right\rangle \|\gamma'(t)\| dt \\
 &= \int_I \langle F \circ \gamma(t), \gamma'(t) \rangle dt
 \end{aligned} \tag{3.1.6}$$

3.1.10. Example. Calculate the work performed when traveling in a force field $F(x, y) = (x, y)$ along the parabola $y = x^2$ in \mathbb{R}^2 from $(0, 0)$ to $(1, 1)$.

$$\begin{aligned}
 W &= \int_{\mathcal{C}^*} F d\vec{s} = \int_0^1 \langle F \circ \gamma(t), \gamma'(t) \rangle dt = \int_0^1 \left\langle \begin{pmatrix} t \\ t^2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2t \end{pmatrix} \right\rangle dt \\
 &= \int_0^1 (t + 2t^3) dt = \frac{1}{2} + \frac{1}{2} = 1
 \end{aligned}$$

Integrals of Vector Fields



Potential Fields

3.1.11. Definition. Let $\Omega \subset \mathbb{R}^n$ be an open set. A vector field $F: \Omega \rightarrow \mathbb{R}^n$ is said to be a **potential field** if there exists a differentiable potential function $U: \Omega \rightarrow \mathbb{R}$ such that

$$F(x) = \nabla U(x).$$

3.1.12. Example. The gravitational force field (3.1.1) introduced in Example 3.1.6 is a potential field, because $F = \nabla U$ for

$$U: \mathbb{R}^3 \setminus \{0\} \rightarrow \mathbb{R}, \quad U(x) = G \frac{m \cdot M}{|x|}, \quad (3.1.7)$$

as is easily checked.

Integrals of Potential Fields

Potential fields are very useful, as the integral along an oriented open curve \mathcal{C}^* depends only on the initial and the final point of the curve. This can be seen from

$$\begin{aligned}\int_I \langle F \circ \gamma(t), \gamma'(t) \rangle dt &= \int_I \langle \nabla U \circ \gamma(t), \gamma'(t) \rangle dt = \int_I DU|_{\gamma(t)}(\gamma'(t)) dt \\ &= \int_I (U \circ \gamma)'(t) dt.\end{aligned}$$

where we have used the chain rule.

Supposing that the initial point of the curve is p_{initial} and the final point is p_{final} , we have from the fundamental theorem of calculus

$$\int_{\mathcal{C}^*} F d\vec{s} = \int_I (U \circ \gamma)'(t) dt = U(p_{\text{final}}) - U(p_{\text{initial}}).$$

We see that for a potential field, the line integral along a simple open curve \mathcal{C}^* depends only on the initial and final points of \mathcal{C} ; the shape of the curve is irrelevant. The potential function U plays the role of a primitive for F .

Conservative Fields

Integrals along closed curves can be easily realized by splitting a closed curve into two open curves. The final point of one curve is the initial point of the other curve.

3.1.13. **Lemma.** Let $\Omega \subset \mathbb{R}^n$ be open, $F: \Omega \rightarrow \mathbb{R}^n$ a potential field and $\mathcal{C} \subset \Omega$ a closed curve. Then

$$\oint_{\mathcal{C}} F d\vec{s} = 0.$$

The proof is obvious from the preceding discussion.

3.1.14. **Definition.** Let $\Omega \subset \mathbb{R}^n$ be open and $F: \Omega \rightarrow \mathbb{R}^n$ a vector field. If the integral along any open curve \mathcal{C}^* depends only on the initial and final points or, equivalently,

$$\oint_{\mathcal{C}} F d\vec{s} = 0 \quad \text{for any closed curve } \mathcal{C},$$

then F is called **conservative**.

Potential Fields are Conservative

In physical terms, a conservative force field has the property that the work required to move a particle from one point to another does not depend on the path taken. Therefore, energy is conserved.

3.1.15. Remark. We note explicitly that every potential field is a conservative field.

In fact, under certain conditions a conservative field is also a potential field.

3.1.16. Definition. Let $\Omega \subset \mathbb{R}^n$. Then Ω is said to be **(pathwise) connected** if for any two points in Ω there exists an open curve within Ω joining the two points.

Conservative Fields are Potential Fields

3.1.17. Theorem. Let $\Omega \subset \mathbb{R}^n$ be a connected open set and suppose that $F: \Omega \rightarrow \mathbb{R}^n$ is a continuous, conservative field. Then F is a potential field.

Proof.

We need to show that there exists a function U such that $F = \nabla U$ on Ω . In fact, we fix an arbitrary point $x_0 \in \Omega$ and define

$$U(x) := \int_{\mathcal{C}^*} F \, d\vec{s}$$

for any path \mathcal{C}^* joining x_0 and x . (The path exists because Ω is connected; the integral does not depend on which path is chosen since F is conservative.) We will show that

$$\frac{\partial U}{\partial x_i} = F_i, \quad i = 1, \dots, n. \quad (3.1.8)$$

Conservative Fields are Potential Fields

Proof (continued).

Let e_i be the i th unit vector and h small enough to ensure that $x + he_i \in \Omega$. A path joining x_0 to $x + he_i$ can be found by taking a path \mathcal{C}^* joining x_0 and x and a straight line segment \mathcal{C}_h^* parametrized by $\gamma(t) = x + the_i$, $0 \leq t \leq 1$. We then have

$$\begin{aligned} U(x + he_i) &= \int_{x_0}^{x+he_i} \mathbf{F} d\vec{s} = \int_{\mathcal{C}^*} \mathbf{F} d\vec{s} + \int_{\mathcal{C}_h^*} \mathbf{F} d\vec{s} \\ &= U(x) + \int_0^1 \langle \mathbf{F}(x + the_i), he_i \rangle dt \\ &= U(x) + h \int_0^1 F_i(x + the_i) dt \end{aligned}$$

Since F_i is continuous, we have $F_i(x + the_i) = F_i(x) + o(1)$ for fixed t .

Conservative Fields are Potential Fields

Proof (continued).

The usual argument (see the proof of Corollary 2.6.2) shows the uniform continuity of F in a small compact neighborhood to ensure that the estimate holds uniformly in t . Then

$$U(x + he_i) = U(x) + h \int_0^1 (F_i(x) + o(1)) dt = U(x) + F_i(x)h + o(h)$$

which establishes (3.1.8). □

Criteria for Potential Fields

3.1.18. Lemma. Let $\Omega \subset \mathbb{R}^n$ be a connected open set and suppose that $F: \Omega \rightarrow \mathbb{R}^n$ is continuously differentiable. Then F is a potential field only if for all $i, j = 1, \dots, n$

$$\frac{\partial F_i}{\partial x_j} = \frac{\partial F_j}{\partial x_i}. \quad (3.1.9)$$

Proof.

If F is a potential field, then $F = \nabla U$ for some potential function U , i.e., $F_i = \frac{\partial U}{\partial x_i}$. Since the second derivative of U is continuous, we have

$$\frac{\partial F_i}{\partial x_j} = \frac{\partial^2 U}{\partial x_j \partial x_i} = \frac{\partial^2 U}{\partial x_i \partial x_j} = \frac{\partial F_j}{\partial x_i}$$

by Schwarz's Theorem 2.5.5. □

Criteria for Potential Fields

3.1.19. Example. The velocity field (3.1.4) introduced in Example 3.1.7 is not a potential field, since $F(x_1, x_2) = \omega(-x_2, x_1)$ and

$$\frac{\partial F_1}{\partial x_2} = -1 \neq 1 = \frac{\partial F_2}{\partial x_1}.$$

Note that (3.1.9) is necessary, but not sufficient, for a field to be a potential field.

3.1.20. Example. The field

$$F: \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}^2, \quad F(x_1, x_2) = \frac{1}{x_1^2 + x_2^2} \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}.$$

satisfies

$$\frac{\partial F_1}{\partial x_2} = \frac{\partial F_2}{\partial x_1}$$

but $\oint_{S^1} F d\vec{s} \neq 0$, so F is not a potential field in $\mathbb{R}^2 \setminus \{0\}$. Details are left to the assignments.

Criteria for Potential Fields

On certain “nice” sets, however, we do have a converse theorem:

3.1.21. Theorem. Let $\Omega \subset \mathbb{R}^n$ be a **simply connected** open set and suppose that $F: \Omega \rightarrow \mathbb{R}^n$ is continuously differentiable. If for all $i, j = 1, \dots, n$

$$\frac{\partial F_i}{\partial x_j} = \frac{\partial F_j}{\partial x_i},$$

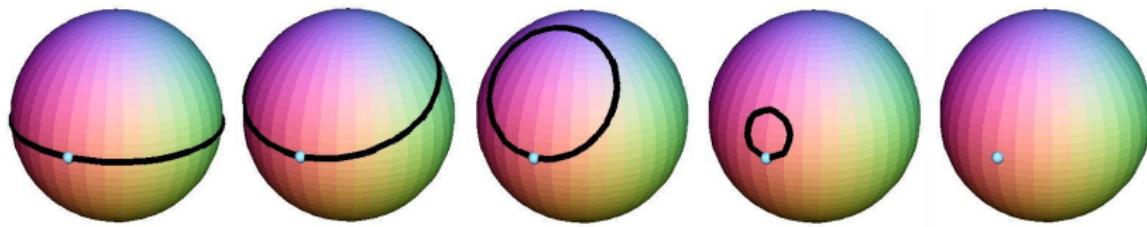
then F is a potential field.

We will not have time prove this result here. However, we do need to explain what a “simply connected” set is.

Loosely speaking, a set $\Omega \subset \mathbb{R}^n$ is said to be simply connected if

- (i) Ω is pathwise connected and
- (ii) every closed curve in Ω can be contracted to a single point within Ω .

Simply Connected Sets



Salix alba. *A homotopy of a circle around a sphere can be reduced to single point.* 2006. Wikipedia. Wikimedia Foundation. Web. 12 July 2012

For example, the unit sphere $S^2 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 = 1\}$ is simply connected, because any closed curve can be “continuously contracted”, staying on the sphere the entire time, until it becomes a single point.

Intuitively, a closed curve can be imagined as a stretched rubber band. If any rubber band can be contracted to a single point within a set, then the set is simply connected.

Simply Connected Sets

3.1.22. Examples.

- (i) $\mathbb{R}^2 \setminus \{(x_1, x_2) : x_1^2 + x_2^2 \leq 1\}$ is not simply connected.
- (ii) $\mathbb{R}^2 \setminus \{0\}$ is not simply connected.
- (iii) $\mathbb{R}^3 \setminus \{0\}$ is simply connected.
- (iv) A torus is not simply connected.

A closed curve \mathcal{C} in a set Ω can be thought of as the image of a continuous function $g: S^1 \rightarrow \mathcal{C}$, where $S^1 = \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1\}$. Let us write $D = \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1\}$.

3.1.23. Definition. Let $\Omega \subset \mathbb{R}^n$ be an open set.

- (i) A closed curve $\mathcal{C} \subset \Omega$ given as the image of a map $g: S^1 \rightarrow \mathcal{C}$ is said to be **contractible to a point** if there exists a continuous function $G: D \rightarrow \Omega$ such that $G|_{S^1} = g$.
- (ii) The set Ω is said to be **simply connected** if it is connected and every closed curve in Ω is contractible to a point.

Determining Potentials

We will develop a practical way of obtaining a potential function for a vector field. The idea is simply to integrate the components of the field and compare the results, then try to find a compatible potential. This is best demonstrated by an example.

3.1.24. Example. Consider the field $F(x_1, x_2) = (x_1^2 + x_2^2, 2x_2x_1 + x_2^2)$. Since F is defined on the simply connected set \mathbb{R}^2 and

$$\frac{\partial F_1}{\partial x_2} = 2x_2 = \frac{\partial F_2}{\partial x_1}$$

the field is a potential field, i.e., $F_1 = \partial_1 \varphi$, $F_2 = \partial_2 \varphi$ for some $U: \mathbb{R}^2 \rightarrow \mathbb{R}$. We integrate the components to find U :

$$U(x_1, x_2) = \int F_1(x_1, x_2) dx_1 = \frac{1}{3}x_1^3 + x_2^2 x_1 + C_1(x_2) \quad (3.1.10)$$

where the integration constant C_1 is now allowed to be a function depending on x_2 .

Determining Potentials

We repeat this for the second component:

$$U(x_1, x_2) = \int F_2(x_1, x_2) dx_2 = \frac{1}{3}x_2^3 + x_2^2 x_1 + C_2(x_1) \quad (3.1.11)$$

where the integration constant C_2 is allowed to depend on x_1 . Comparing (3.1.10) with (3.1.11), we see that

$$U(x_1, x_2) = \frac{1}{3}(x_1^3 + x_2^3) + x_2^2 x_1$$

is a potential function for F (of course, we can add any constant to U if we like).

This procedure works analogously for vector fields in \mathbb{R}^n .

Differential Forms

The transpose of a vector field is called a **differential form**:

$$\begin{aligned} F(x)^T &= (F_1(x), \dots, F_n(x)) \\ &= F_1(x) dx_1 + \cdots + F_n(x) dx_n, \end{aligned}$$

where the differentials dx_j , $j = 1, \dots, n$, are simply the standard basis row vectors, as defined in (2.4.1).

3.1.25. Definition. Let $F_1, \dots, F_n: \mathbb{R}^n \rightarrow \mathbb{R}$ be scalar functions. Then

$$\alpha = F_1 dx_1 + \cdots + F_n dx_n$$

is said to be a **differential one-form**.

Integrals of Differential Forms

Note that

$$\int_{\mathcal{C}^*} \alpha = \int_{\mathcal{C}^*} F d\vec{s}$$

where $F = (F_1, \dots, F_n)^T$ is the transpose of the differential form α . We thus see that the integral of a differential form can be expressed through that of a vector field, which we have shown to be independent of parametrization. Thus the integral of a form is likewise independent of parametrization.

3.1.26. Example. We integrate the form $4y \, dx + 2x^2y \, dy$ in counter-clockwise direction along the unit circle $S^1 \subset \mathbb{R}^2$. We parametrize the circle by $\gamma(\theta) = (\cos \theta, \sin \theta)$, $0 \leq \theta < 2\pi$,

$$\begin{aligned} \oint_{S^1} 4x \, dx + 2x^2y \, dy &= \int_0^{2\pi} 4 \sin \theta (-\sin \theta) \, d\theta + \int_0^{2\pi} 2 \cos^2 \theta \sin \theta (\cos \theta) \, d\theta \\ &= 2 \int_0^{2\pi} \sin \theta \cos \theta - 2 \, d\theta = -8\pi. \end{aligned}$$

Vector Fields and Line Integrals

Circulation and Flux

The Riemann Integral and Measurable Sets

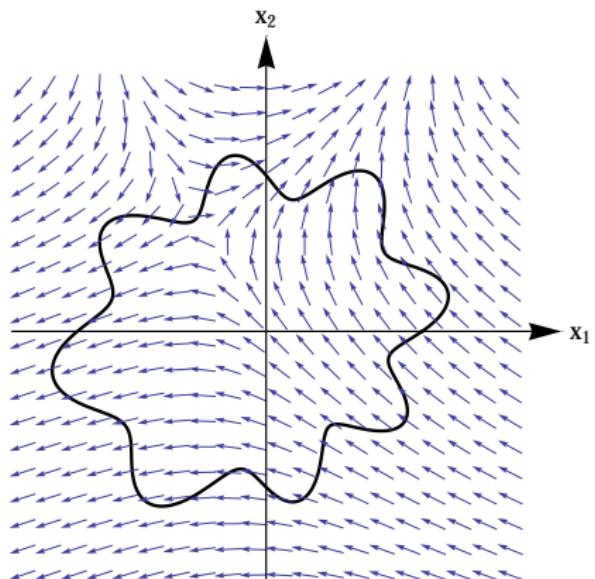
Integration in Practice

Surfaces and Surface Integrals

The Theorems of Gauß and Stokes

Vector Fields of Fluids

In the previous section we have primarily motivated line integrals of vector fields through the concept of work in a force field. Another physical approach is to motivate vector fields through velocity fields of fluids. This turns out to yield further useful concepts in field theory.



We will consider fluid flows in \mathbb{R}^2 to introduce general concepts. Observe the vector field illustrated at left, interpreted as the direction field of a fluid flow, and the closed curve, interpreted as the boundary of a region.

We can decompose the vector field at the boundary into a tangential component and a normal component.

Circulation and Flux

We interpret the normal component of the vector field as the part that flows through the boundary of the region, i.e., into or out of the region. This is called the **flux** of the vector field through the boundary.

The tangential component is the part of the vector field that flows around the boundary, called the **circulation** of the field.

3.2.1. Example. Let $S^1 = \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1\}$ be the unit circle, bounding the unit disc. Consider the two vector fields

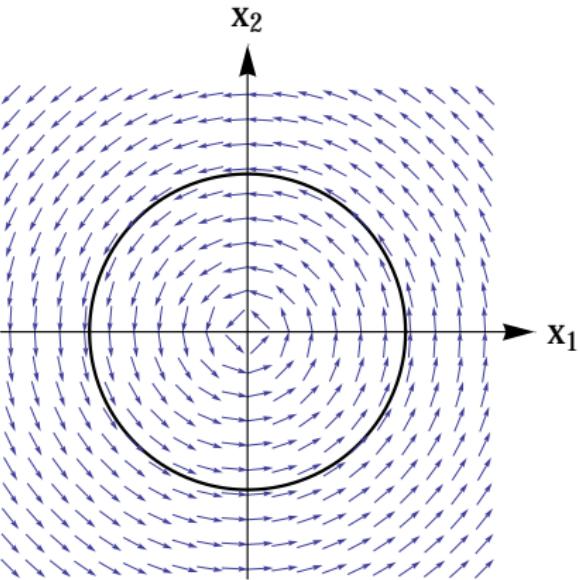
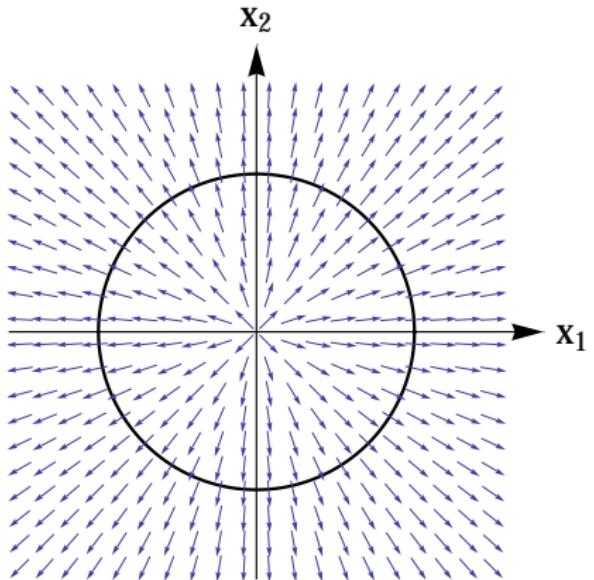
$$F, G: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad F(x_1, x_2) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad G(x_1, x_2) = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}.$$

The unit tangent vector to S^1 at (x_1, x_2) is given by $T = (-x_2, x_1)$, so

$$\langle T, F \rangle|_{S^1} = -x_2 x_1 + x_1 x_2 = 0,$$

$$\langle T, G \rangle|_{S^1} = -x_2(-x_2) + x_1 x_1 = 1.$$

Circulation and Flux



The unit normal vector at $x \in S^1$ is given by $N = -x$, so

$$\langle N, F \rangle|_{S^1} = -x_1 x_1 + -x_2 x_2 = -1,$$

$$\langle N, G \rangle|_{S^1} = x_1(-x_2) + x_2 x_1 = 0.$$

Circulation and Flux

3.2.2. Definition. Let $\Omega \subset \mathbb{R}^2$ be an open set, $F: \Omega \rightarrow \mathbb{R}^2$ a continuously differentiable vector field and \mathcal{C}^* a positively oriented closed curve in \mathbb{R}^2 . Then

$$\int_{\mathcal{C}^*} \langle F, T \rangle \, ds \quad (3.2.1)$$

is called the (total) **circulation** of F along \mathcal{C} and

$$\int_{\mathcal{C}^*} \langle F, N \rangle \, ds \quad (3.2.2)$$

is called the (total) **flux** of F through \mathcal{C} .

Here T denotes the usual unit tangent vector to \mathcal{C} . However, N is taken to the “unit normal vector” in the sense that

1. $\|N\| = 1$,
2. $\langle N, T \rangle = 0$,
3. N points **outwards** from the region bounded by \mathcal{C} .

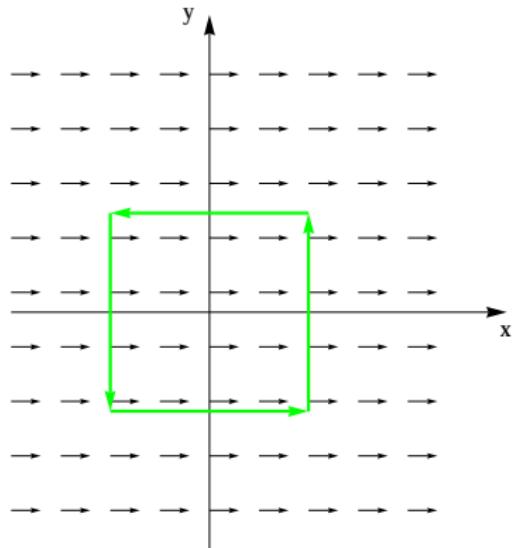
Circulation and Flux

3.2.3. Remarks.

- ▶ The integral (3.2.1) coincides with the line integral we defined in (3.1.5) and hence also gives the amount of work needed to move a particle along the closed curve \mathcal{C} . In a non-rotating fluid, this work should be zero.
- ▶ The integral (3.2.2) is a special case of a **surface integral**, which we will study later. In \mathbb{R}^2 , the “surfaces” of open sets are simply boundary curves, while in \mathbb{R}^3 , surfaces of open sets will be (essentially) two-dimensional objects.

Circulation and Flux

3.2.4. Example. For the vector field $F(x_1, x_2) = (1, 0)$ and the square pictured below, both the circulation and the flux are zero.



Total and Infinitesimal Flux

The previous example shows clearly that the total flux of a vector field through a boundary is the difference between “influx” and “efflux” of the field. In the context of fluid flow, zero (total) flux through a boundary means

“what flows in also flows out”

or

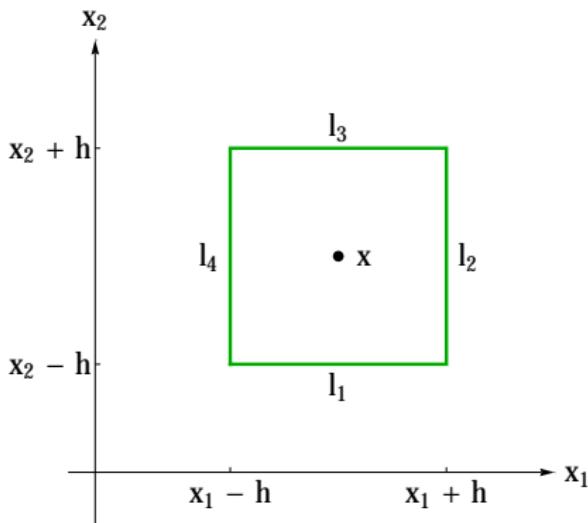
“there are no fluid sources or sinks within the boundary”

We now want to characterize vector fields where the flux through **any** boundary is zero. In fluid flow, these correspond to fluid fields where the fluid volume is preserved (incompressible fluids with no external influx or efflux).

This approach will lead to **infinitesimal flux**, i.e., the flux of the field through a given point (instead of along a boundary curve).

The Flux Through a Square

Let $\Omega \subset \mathbb{R}^2$ be open and consider the flux of a vector field $F: \Omega \rightarrow \mathbb{R}^2$ through a square of sidelength $2h$, $h > 0$, centered at a point $x \in \Omega$.



In particular, the square is given by

$$S_h = [x_1 - h, x_1 + h] \times [x_2 - h, x_2 + h]$$

and the boundary consists of four lines,

$$\partial S_h = l_1 \cup l_2 \cup l_3 \cup l_4.$$

We find the flux through the boundary by integrating

$$\int_{\partial S_h} \langle F, N \rangle \, ds = \sum_{k=1}^4 \int_{l_k} \langle F, N_k \rangle \, ds$$

The Flux Through a Square

We have the following parametrizations and normal vectors:

$$l_1: N_1 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \gamma_1(t) = \begin{pmatrix} x_1 \\ x_2 - h \end{pmatrix} + t \begin{pmatrix} 1 \\ 0 \end{pmatrix}, -h \leq t \leq h,$$

$$l_2: N_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \gamma_2(t) = \begin{pmatrix} x_1 + h \\ x_2 \end{pmatrix} + t \begin{pmatrix} 0 \\ 1 \end{pmatrix}, -h \leq t \leq h,$$

$$l_3: N_3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \gamma_3(t) = \begin{pmatrix} x_1 \\ x_2 + h \end{pmatrix} - t \begin{pmatrix} 1 \\ 0 \end{pmatrix}, -h \leq t \leq h,$$

$$l_4: N_4 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \gamma_4(t) = \begin{pmatrix} x_1 - h \\ x_2 \end{pmatrix} - t \begin{pmatrix} 0 \\ 1 \end{pmatrix}, -h \leq t \leq h.$$

Hence,

$$\begin{aligned} \int_{\partial S_h} \langle F, N \rangle ds &= - \int_{-h}^h F_2(x_1 + t, x_2 - h) dt + \int_{-h}^h F_1(x_1 + h, x_2 + t) dt \\ &\quad + \int_{-h}^h F_2(x_1 - t, x_2 + h) dt - \int_{-h}^h F_1(x_1 - h, x_2 - t) dt \end{aligned}$$

In the last two integrals, we substitute $\tau = -t$ and then rename the variable t .

The Flux Through a Square

This gives

$$\begin{aligned}\int_{\partial S_h} \langle F, N \rangle \, ds &= \int_{-h}^h (F_2(x_1 + t, x_2 + h) - F_2(x_1 + t, x_2 - h)) \, dt \\ &\quad + \int_{-h}^h (F_1(x_1 + h, x_2 + t) - F_1(x_1 - h, x_2 + t)) \, dt\end{aligned}$$

We note that (as $h \rightarrow 0$, uniformly for $t \in [-h, h]$)

$$F_2(x_1 + t, x_2 \pm h) = F_2(x_1 + t, x_2) \pm \frac{\partial F_2(x_1 + t, x_2)}{\partial x_2} \cdot h + o(h),$$

$$F_1(x_1 \pm h, x_2 + t) = F_1(x_1, x_2 + t) \pm \frac{\partial F_1(x_1, x_2 + t)}{\partial x_1} \cdot h + o(h).$$

The Flux Through a Square

Inserting this expansion and substituting in the integrals

$$\begin{aligned}\int_{\partial S_h} \langle F, N \rangle \, ds &= 2h \int_{-h}^h \left(\frac{\partial F_1(x_1, x_2 + t)}{\partial x_1} + o(1) \right) \, dt \\ &\quad + 2h \int_{-h}^h \left(\frac{\partial F_2(x_1 + t, x_2)}{\partial x_2} + o(1) \right) \, dt \\ &= (2h)^2 \int_{-1/2}^{1/2} \left(\frac{\partial F_1(x_1, x_2 + 2ht)}{\partial x_1} + o(1) \right) \, dt \\ &\quad + (2h)^2 \int_{-1/2}^{1/2} \left(\frac{\partial F_2(x_1 + 2ht, x_2)}{\partial x_2} + o(1) \right) \, dt\end{aligned}$$

The Flux Through a Square

By the Mean Value Theorem of integral calculus (see assignments of Vv186), for any value of h there exists a $t_0 \in [-1/2, 1/2]$ such that

$$\int_{-1/2}^{1/2} \frac{\partial F_1(x_1, x_2 + 2ht)}{\partial x_1} dt = \frac{\partial F_1(x_1, x_2 + 2ht_0)}{\partial x_1}$$
$$\xrightarrow{h \rightarrow 0} \frac{\partial F_1(x_1, x_2)}{\partial x_1}.$$

This implies that

$$\lim_{h \rightarrow 0} \frac{1}{(2h)^2} \int_{\partial S_h} \langle F, N \rangle ds = \left. \frac{\partial F_1}{\partial x_1} \right|_x + \left. \frac{\partial F_2}{\partial x_2} \right|_x.$$

Flux Density and the Divergence

We have shown that if S_h is a square of side length $2h$ centered at x ,

$$\frac{\text{Flux through } S_h}{\text{Area}(S_h)} \xrightarrow{h \rightarrow 0} \left. \frac{\partial F_1}{\partial x_1} \right|_x + \left. \frac{\partial F_2}{\partial x_2} \right|_x$$

Hence, the limit on the right corresponds to a **flux density**. There is a special term for this flux density:

3.2.5. Definition. Let $\Omega \subset \mathbb{R}^n$ and $F: \Omega \rightarrow \mathbb{R}^n$ be a continuously differentiable vector field. Then

$$\operatorname{div} F := \frac{\partial F_1}{\partial x_1} + \cdots + \frac{\partial F_n}{\partial x_n}$$

is called the **divergence** of F .

The flux density at a point x is given by the divergence of the field at x . Although we have only proven this in the case of fields in \mathbb{R}^2 , this holds in any dimension $n \geq 2$ (we will prove this using surface integrals in \mathbb{R}^3 later).

Circulation Around a Parallelogram

We now turn to the **circulation** of a vector field. Again, our goal is to find an expression for the infinitesimal circulation around a point $x \in \mathbb{R}^n$. In contrast to the flux, where we had to restrict ourselves to \mathbb{R}^2 , we now consider a line integral in \mathbb{R}^n .

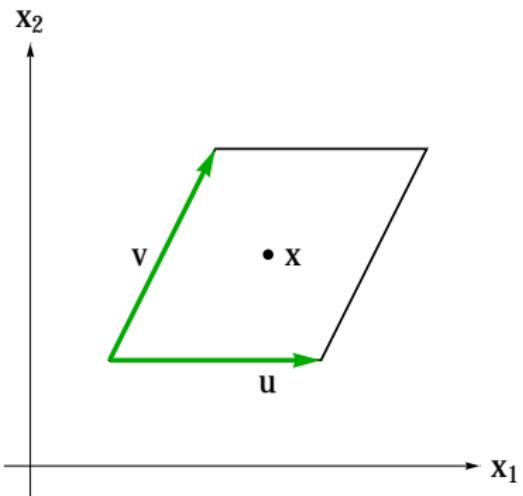
Let $\Omega \subset \mathbb{R}^n$ and $F: \Omega \rightarrow \mathbb{R}^n$ a continuously differentiable vector field. Let $x \in \Omega$ and denote by $P(u, v)$ a parallelogram spanned by vectors u and v and centered at x . We want to calculate the circulation of F around the boundary of the parallelogram, given by

$$\int_{\partial P(u,v)} F \, d\vec{s}.$$

Our goal is to analyze the integral when $\|u\| + \|v\| \rightarrow 0$.

Circulation Around a Parallelogram

The parallelogram is the union of four straight line segments parametrized by



$$\gamma_1(t) = x - \frac{u + v}{2} + tu,$$

$$\gamma_2(t) = x + \frac{u - v}{2} + tv,$$

$$\gamma_3(t) = x + \frac{u + v}{2} - tu,$$

$$\gamma_4(t) = x + \frac{v - u}{2} - tv$$

for $t \in [0, 1]$.

We will use that for every point $\gamma_j(t)$ on the parallelogram, we can write

$$F(\gamma_j(t)) = F(x) + DF|_x(\gamma_j(t) - x) + o(\|u\| + \|v\|)$$

as $\|u\| + \|v\| \rightarrow 0$.

Circulation Around a Parallelogram

The line integral is then

$$\begin{aligned}\int_{\partial P(u,v)} F \, d\vec{s} &= \sum_{j=1}^4 \int_0^1 \langle F \circ \gamma_j(t), \gamma'_j(t) \rangle \, dt \\&= \int_0^1 \langle F \circ \gamma_1(t), u \rangle \, dt + \int_0^1 \langle F \circ \gamma_2(t), v \rangle \, dt \\&\quad - \int_0^1 \langle F \circ \gamma_3(t), u \rangle \, dt - \int_0^1 \langle F \circ \gamma_4(t), v \rangle \, dt \\&= \int_0^1 \langle DF|_x \gamma_1(t), u \rangle \, dt + \int_0^1 \langle DF|_x \gamma_2(t), v \rangle \, dt \\&\quad - \int_0^1 \langle DF|_x \gamma_3(t), u \rangle \, dt - \int_0^1 \langle DF|_x \gamma_4(t), v \rangle \, dt \\&\quad + o((\|u\| + \|v\|)^2)\end{aligned}$$

Circulation Around a Parallelogram

From the linearity of the integral, the inner product and $DF|_x$ we can write

$$\begin{aligned} \int_{\partial P(u,v)} F d\vec{s} &= \int_0^1 \langle DF|_x(\gamma_1(t) - \gamma_3(t)), u \rangle dt \\ &\quad + \int_0^1 \langle DF|_x(\gamma_2(t) - \gamma_4(t)), v \rangle dt + o((\|u\| + \|v\|)^2) \\ &= \int_0^1 -\langle DF|_x(u + v - 2tu), u \rangle dt \\ &\quad + \int_0^1 \langle DF|_x(u - v + 2tv), v \rangle dt + o((\|u\| + \|v\|)^2) \\ &= \langle DF|_x u, v \rangle - \langle DF|_x v, u \rangle + o((\|u\| + \|v\|)^2) \end{aligned}$$

In leading order, the circulation is thus $\langle DF|_x u, v \rangle - \langle DF|_x v, u \rangle$.

The Circulation Density - Rotation / Curl

The expression $\langle DF|_x u, v \rangle - \langle DF|_x v, u \rangle$ is clearly anti-symmetric (it changes sign when u and v are interchanged) and bilinear. As we will see, it is the main term describing the **circulation density** in the plane spanned by u and v . It therefore deserves a special mention.

3.2.6. Definition. Let $\Omega \subset \mathbb{R}^n$ be open and $F: \Omega \rightarrow \mathbb{R}^n$ a continuously differentiable vector field. Then the anti-symmetric, bilinear form

$$\text{rot } F|_x: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}, \quad \text{rot } F|_x(u, v) := \langle DF|_x u, v \rangle - \langle DF|_x v, u \rangle \quad (3.2.3)$$

is called the **rotation** (in mainland Europe) or **curl** (in anglo-saxon countries) of the vector field F at $x \in \mathbb{R}^n$.

We will study this bilinear form in more detail for the case of fields in \mathbb{R}^2 and \mathbb{R}^3 .

The Rotation in \mathbb{R}^2

In \mathbb{R}^2 , the area of the parallelogram is given by $|\det(u, v)|$. The circulation density (circulation per unit area) is then

$$\frac{1}{|\det(u, v)|} \int_{\partial P(u, v)} F d\vec{s} = \frac{\langle DF|_x u, v \rangle - \langle DF|_x v, u \rangle}{\det(u, v)} + o(\|u\| + \|v\|)$$

For instance, if we set $u = h \cdot u_0$ and $v = h \cdot v_0$ where $h > 0$ and $\|u_0\|, \|v_0\| > 0$ are fixed, we have

$$\frac{1}{|\det(u, v)|} \int_{\partial P(u, v)} F d\vec{s} \xrightarrow{h \rightarrow 0} \frac{\langle DF|_x u_0, v_0 \rangle - \langle DF|_x v_0, u_0 \rangle}{|\det(u_0, v_0)|}.$$

Therefore,

$$\frac{\langle DF|_x u, v \rangle - \langle DF|_x v, u \rangle}{|\det(u, v)|} = \frac{\text{rot } F|_x(u, v)}{\text{Area}(P(u, v))} \quad (3.2.4)$$

represents the infinitesimal circulation around a parallelogram centered at a point $x \in \mathbb{R}^2$.

The Rotation in \mathbb{R}^2

3.2.7. Theorem. Let $\Omega \subset \mathbb{R}^2$ be open and $F: \Omega \rightarrow \mathbb{R}^2$ a continuously differentiable vector field. Then there exists a uniquely defined continuous potential function $\text{rot } F: \Omega \rightarrow \mathbb{R}$ such that

$$\text{rot } F|_x(u, v) = \text{rot } F(x) \cdot \det(u, v). \quad (3.2.5)$$

Proof.

The determinant is the unique alternating, normed, bilinear form in \mathbb{R}^2 . Since $\text{rot } F|_x$ is alternating and bilinear (but not normed) it must be a multiple of the determinant. In fact, we have

$$\begin{aligned} \text{rot } F(x) &= \text{rot } F(x) \cdot \det \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) = \text{rot } F|_x \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) \\ &= \langle DF|_x e_1, e_2 \rangle - \langle DF|_x e_2, e_1 \rangle \\ &= \frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2}. \end{aligned}$$



The Rotation in \mathbb{R}^2

3.2.8. Remark. Comparing (3.2.5) with the circulation density (3.2.4), it is clear that the function $\operatorname{rot} F$ gives this circulation density. We note:

The circulation density of a vector field F in \mathbb{R}^2 is represented by a scalar function, $\operatorname{rot} F$.

This scalar function is given by

$$\operatorname{rot} F = \frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2}. \quad (3.2.6)$$

The Rotation in \mathbb{R}^3

Let us now consider a field in three dimensions. The area of the parallelogram spanned by u and v is equal to the volume of the parallelepiped formed by u , v , and an orthogonal unit vector. Hence, by (1.7.5),

$$\text{Area}(P(u, v)) = \det \left(u, v, \frac{u \times v}{\|u \times v\|} \right) = \|u \times v\|.$$

Similarly to the situation in two dimensions, if we set $u = h \cdot u_0$ and $v = h \cdot v_0$ we have the circulation density

$$\begin{aligned} \frac{1}{\|u \times v\|} \int_{\partial P(u, v)} F \, d\vec{s} &\xrightarrow{h \rightarrow 0} \frac{\langle DF|_x u_0, v_0 \rangle - \langle DF|_x v_0, u_0 \rangle}{\|u_0 \times v_0\|} \\ &= \frac{\text{rot } F|_x(u, v)}{\text{Area}(P(u, v))} \end{aligned}$$

The Rotation in \mathbb{R}^3

3.2.9. Theorem. Let $\Omega \subset \mathbb{R}^3$ be open and $F: \Omega \rightarrow \mathbb{R}^3$ a continuously differentiable vector field. Then there exists a uniquely defined continuous vector field $\text{rot } F: \Omega \rightarrow \mathbb{R}^3$ such that

$$\text{rot } F|_x(u, v) = \det(\text{rot } F(x), u, v) = \langle \text{rot } F(x), u \times v \rangle. \quad (3.2.7)$$

Proof.

Suppose that the vector field $\text{rot } F$ exists for $F = (F_1, F_2, F_3)$; then we can use (3.2.7) to calculate the first component of $\text{rot } F$:

$$\begin{aligned} (\text{rot } F(x))_1 &= \langle \text{rot } F(x), e_1 \rangle = \langle \text{rot } F(x), e_2 \times e_3 \rangle = \text{rot } F|_x(e_2, e_3) \\ &= \langle DF|_x e_2, e_3 \rangle - \langle DF|_x e_3, e_2 \rangle \\ &= \frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3}. \end{aligned}$$

The Rotation in \mathbb{R}^3

Proof (continued).

Similarly, we calculate the other components of $\text{rot } F$ to obtain

$$\text{rot } F(x) = \begin{pmatrix} \frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3} \\ \frac{\partial F_1}{\partial x_3} - \frac{\partial F_3}{\partial x_1} \\ \frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \end{pmatrix}.$$

Hence, the vector $\text{rot } F$ is uniquely determined from F by (3.2.7).

The existence of a vector $y \in \mathbb{R}^3$ such that $\text{rot } F|_x(u, v) = \det(y, u, v)$ will be proven in the assignments. Since F is continuously differentiable as a function of x , so is the vector $y = y(x)$ and the vector field $\text{rot } F(x) = y(x)$ exists. □

The Rotation in \mathbb{R}^3

3.2.10. Remark. Comparing (3.2.7) with the circulation density (3.2.4), the circulation density in the plane spanned by u and v at x is given by

$$\left\langle \operatorname{rot} F|_x, \frac{u \times v}{\|u \times v\|} \right\rangle.$$

We note:

The circulation density of a vector field F in \mathbb{R}^3 is represented by a vector field, $\operatorname{rot} F$.

This vector field is given by

$$\operatorname{rot} F = \begin{pmatrix} \frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3} \\ \frac{\partial F_1}{\partial x_3} - \frac{\partial F_3}{\partial x_1} \\ \frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \end{pmatrix}. \quad (3.2.8)$$

The Rotation in \mathbb{R}^3

3.2.11. Remark. We can consider \mathbb{R}^2 as being a subspace of \mathbb{R}^3 by identifying points $(x_1, x_2) \in \mathbb{R}^2$ with $(x_1, x_2, 0) \in \mathbb{R}^3$. Similarly, a vector field in \mathbb{R}^2 can be considered as a field in \mathbb{R}^3 of the form

$$F(x_1, x_2, x_3) = \begin{pmatrix} F_1(x_1, x_2) \\ F_2(x_1, x_2) \\ 0 \end{pmatrix}.$$

We then obtain

$$\text{rot } F(x) = \begin{pmatrix} 0 \\ 0 \\ \frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \end{pmatrix},$$

effectively regaining (3.2.6) from (3.2.8).

The Rotation in Higher Dimensions

In general, the space of bilinear maps in \mathbb{R}^n is isomorphic to the space of $n \times n$ matrices (see (2.5.6) and the discussion there). In particular, there must exist a matrix $A(x)$ such that

$$\text{rot } F|_x(u, v) = \langle u, A(x)v \rangle \quad (3.2.9)$$

Since the rotation is anti-symmetric, $\text{rot } F|_x(v, u) = -\text{rot } F|_x(u, v)$, i.e.,

$$\langle u, A(x)v \rangle = -\langle v, A(x)u \rangle = \langle -A(x)^T v, u \rangle = \langle u, -A(x)^T v \rangle.$$

Hence, A must satisfy

$$A(x)^T = -A(x).$$

for all x . In fact, we can see directly from the definition (3.2.3) that

$$A(x) = (DF|_x)^T - DF|_x.$$

The Rotation in Higher Dimensions

The subspace

$$\{A \in \text{Mat}(n \times n; \mathbb{R}) : A^T = -A\}$$

has dimension 1 if $n = 2$, so $\text{rot } F|_x$ can be represented by a scalar. For $n = 3$, the space of anti-symmetric matrices has dimension 3, so $\text{rot } F|_x$ can be represented as a vector in \mathbb{R}^3 .

In \mathbb{R}^4 , the space of antisymmetric matrices has dimension 6, and the rotation can not be easily represented as a vector. This is the case for all higher dimensions, where the rotation can only be expressed through an anti-symmetric matrix.

Irrotational Fields

A continuously differentiable field $F: \Omega \rightarrow \mathbb{R}^n$ such that $\text{rot } F|_x = 0$ for all $x \in \Omega$ is said to be **irrotational**. Writing an irrotational field in the form (3.2.9), we see that $A(x) = 0$ for all x . This implies

$$(DF|_x)^T = DF|_x$$

which means that

$$\frac{\partial F_i}{\partial x_j} = \frac{\partial F_j}{\partial x_i}, \quad i, j = 1, \dots, n.$$

Hence, a potential field is irrotational. Conversely, if $F: \Omega \rightarrow \mathbb{R}^n$ is defined on a simply connected domain, we may apply Theorem 3.1.21 to deduce that F is a potential field.

Fluid Statics

Fluid statics is the study of time-independent flows. In particular, the streamlines are assumed to be given by a direction field F in \mathbb{R}^n (most often, $n = 2$ or 3) that does not depend on time. Irrotational fluid flow is often modeled by a potential field, i.e., one assumes

$$F = \nabla U$$

for some potential U . (The resulting flow is known as **potential flow**.) If there are no sources or sinks and the fluid is incompressible, one additionally has

$$\operatorname{div} F = 0.$$

Combining these two equations, one obtains

$$\operatorname{div}(\nabla U) = \operatorname{div} \begin{pmatrix} \frac{\partial U}{\partial x_1} \\ \vdots \\ \frac{\partial U}{\partial x_n} \end{pmatrix} = \frac{\partial^2 U}{\partial x_1^2} + \cdots + \frac{\partial^2 U}{\partial x_n^2} = \Delta U = 0.$$

The Laplace Equation

The equation

$$\Delta U = 0$$

is a partial differential equation and known as the **Laplace equation**.

Together with boundary conditions for the flow it can (in principle) be solved to yield the streamlines of an irrotational, incompressible fluid in any physical situation.

However, practical solutions must often rely on numerical or approximate methods, as solving the equation explicitly is possible only in the simplest situations (such as fluid flow around a sphere). Finding solutions to this (and other) partial differential equations is one of the main research problems in applied mathematics and engineering.

Solutions of the Laplace equation play a minor role in Vv286 (Honors Math II), a major role in Vv454 (Partial Differential Equations and Boundary Value Problems) and are a principal topic of Vv557 (Methods of Applied Mathematics II).

Triangle Calculus

It is convenient to introduce the formal “vector”

$$\nabla := \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix}$$

Then the gradient of a potential function f is just

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

The divergence of a vector field F can be expressed as

$$\operatorname{div} F = \langle \nabla, F \rangle = \left\langle \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix}, \begin{pmatrix} F_1 \\ \vdots \\ F_n \end{pmatrix} \right\rangle = \frac{\partial F_1}{\partial x_1} + \cdots + \frac{\partial F_n}{\partial x_n}.$$

Triangle Calculus

The rotation of a vector field F can be formally written as

$$\text{rot } F = \nabla \times F(x) = \det \begin{pmatrix} e_1 & e_2 & e_3 \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ F_1 & F_2 & F_3 \end{pmatrix},$$

where e_1, e_2, e_3 are the standard basis vectors in \mathbb{R}^3 . Finally, we can formally write

$$\begin{aligned} \langle \nabla, \nabla \rangle &= \left(\frac{\partial}{\partial x_1} \right)^2 + \cdots + \left(\frac{\partial}{\partial x_n} \right)^2 \\ &= \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2} \\ &= \Delta \end{aligned}$$

so it is common for physicists to write ∇^2 instead of Δ .

Vector Fields and Line Integrals

Circulation and Flux

The Riemann Integral and Measurable Sets

Integration in Practice

Surfaces and Surface Integrals

The Theorems of Gauß and Stokes

Integration in \mathbb{R}^n

Contrary to the general approach we took in differential calculus, we will now focus on functions defined on \mathbb{R}^n , even restricting ourselves to \mathbb{R}^2 and \mathbb{R}^3 in some cases. The reason for this is that the geometry of objects in general vector spaces (even \mathbb{R}^n) is quite complex, and in order to understand even the finite-dimensional case, we would need to introduce a variety of abstract concepts (manifolds, tangent and cotangent spaces etc.). This is generally done in courses on **vector analysis**; regrettably, we do not have time to pursue these things here.

When integrating a function defined on a subset $\Omega \subset \mathbb{R}^2$, there are some difficulties that do not occur for functions of a single variable. In particular, while integrating across an interval $[a, b]$ is straightforward, the shape of Ω now plays a significant role.

We will discuss the concept of volume of sets and introduce integrals of step functions on cuboids before attempting to extend integration to continuous functions defined on more general sets.

Cuboids

We wish to assign a volume to general sets in \mathbb{R}^n .

3.3.1. Definition. Let $a_k, b_k, k = 1, \dots, n$ be pairs of numbers with $a_k < b_k$. Then the set $Q \subset \mathbb{R}^n$ given by

$$\begin{aligned} Q &= [a_1, b_1] \times \cdots \times [a_n, b_n] \\ &= \{x \in \mathbb{R}^n : x_k \in [a_k, b_k], k = 1, \dots, n\} \end{aligned}$$

is called an ***n*-cuboid**. We define the volume of Q to be

$$|Q| := \prod_{k=1}^n (b_k - a_k).$$

We will denote the set of all *n*-cuboids by \mathcal{Q}_n .

3.3.2. Remark. Clearly, an *n*-cuboid is a compact set in \mathbb{R}^n .

Upper and Lower Volumes of Sets

The idea for assigning volume to a subset $\Omega \subset \mathbb{R}^n$ is similar to that for the Riemann integral: consider volumes of enclosing and enclosed n -cuboids; if there infimum and supremum (respectively) are equal, assign this number as the volume of Ω .

3.3.3. Definition. Let $\Omega \subset \mathbb{R}^n$ be a bounded non-empty set. We define the **outer** and **inner volume** of Ω by

$$\overline{V}(\Omega) := \inf \left\{ \sum_{k=0}^r |Q_k| : r \in \mathbb{N}, Q_0, \dots, Q_r \in \mathcal{Q}_n, \Omega \subset \bigcup_{k=1}^r Q_k \right\},$$

$$\underline{V}(\Omega) := \sup \left\{ \sum_{k=0}^r |Q_k| : r \in \mathbb{N}, Q_0, \dots, Q_r \in \mathcal{Q}_n, \Omega \supset \bigcup_{k=1}^r Q_k, \bigcap_{k=1}^r Q_k = \emptyset \right\}.$$

It is easy to see that $0 \leq \underline{V}(\Omega) \leq \overline{V}(\Omega)$.

Measurable Sets

Sets for which we can define a volume are called **measureable**. The volume is referred to as the **measure** of a set.

3.3.4. Definition. Let $\Omega \subset \mathbb{R}^n$ be a bounded set. Then Ω is said to be **(Jordan) measurable** if either

- (i) $\overline{V}(\Omega) = 0$ or
- (ii) $\overline{V}(\Omega) = \underline{V}(\Omega)$.

In the first case, we say that Ω has (Jordan) **measure zero**, in the second case we say that

$$|\Omega| := \overline{V}(\Omega) = \underline{V}(\Omega)$$

is the Jordan measure of Ω .

Sets of Measure Zero

For a set $\Omega \subset \mathbb{R}^n$ to have measure zero, $\underline{V}(\Omega)$ does not need to exist (possibly because there is no n -cuboid that can be a subset of Ω).

3.3.5. Examples.

- (i) A set $\{x\}$ consisting of a single point $x \in \mathbb{R}^n$ is a set of measure zero.
- (ii) A subset of \mathbb{R}^n consisting of a finite number of single points is a set of measure zero.
- (iii) A curve of finite length $\mathcal{C} \subset \mathbb{R}^n$, $n \geq 2$, is a set of measure zero.
- (iv) A bounded section of a plane in \mathbb{R}^3 is a set of measure zero.
- (v) The set of rational numbers in the interval $[0, 1]$ has measure zero.
- (vi) The set of irrational numbers in the interval $[0, 1]$ is not (Jordan) measurable.

“Almost Everywhere” Properties

3.3.6. Definition. A function f on \mathbb{R}^n that has a property for all $x \in \mathbb{R}^n \setminus \Omega$, where Ω is a set of measure zero, is said to have this property **almost everywhere** (often abbreviated by “a.e.”).

3.3.7. Example.

- (i) The function $f: \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = |x|$ is differentiable almost everywhere.
- (ii) The function $f: [0, 1] \times [0, 1] \rightarrow \mathbb{R}$,

$$f(x, y) = \begin{cases} \frac{1}{x-y} & x \neq y, \\ 0 & \text{otherwise} \end{cases},$$

is continuous almost everywhere.

Partitions of Cuboids

Let $Q \subset \mathbb{R}^n$ be an n -cuboid. We can then define step functions on Q just as we did for intervals. First, recall from 186 Definition 4.1.1 that a partition P of an interval $[a, b]$ is a finite sequence of numbers $P = (a_0, \dots, a_n)$ with

$$a = a_0 < a_1 < \dots < a_m = b.$$

3.3.8. Definition. A **partition** P of an n -cuboid $Q = [a_1, b_1] \times \cdots \times [a_n, b_n]$ is a tuple $P = (P_1, \dots, P_n)$ such that $P_k = (a_{k0}, \dots, a_{km_k})$ is a partition of the interval $[a_k, b_k]$.

The partition P of Q induces cuboids of the form

$$Q_{j_1 j_2 \dots j_n} := [a_{1(j_1-1)}, a_{1j_1}] \times \cdots \times [a_{n(j_n-1)}, a_{nj_n}] \subset Q.$$

Step Functions on Cuboids

The intersection of the cuboids $Q_{j_1 j_2 \dots j_n}$ is a subset of

$$\Omega = \{x \in Q : x_k = a_{k i_k} \text{ for some } k = 1, \dots, n\},$$

which is a set of measure zero. We say that the union of sets whose intersection is a set of measure zero is **almost disjoint**. Thus,

$$Q = \bigcup_{\substack{1 \leq j_1 \leq m_1 \\ \vdots \\ 1 \leq j_n \leq m_n}} Q_{j_1 j_2 \dots j_n}$$

is the union of almost disjoint cuboids induced by a partition P of Q .

Step Functions on Cuboids

3.3.9. Definition. Let $Q \subset \mathbb{R}^n$ be an n -cuboid. A function $f: Q \rightarrow \mathbb{R}$ is called a **step function with respect to a partition P** if there exist numbers $y_{j_1 j_2 \dots j_n} \in \mathbb{R}$ such that $f(x) = y_{j_1 j_2 \dots j_n}$ whenever $x \in \text{int } Q_{j_1 j_2 \dots j_n}$, $j_k = 1, \dots, m_k$, $k = 1, \dots, n$.

3.3.10. Remarks.

- (i) It doesn't matter how the step function is defined on the set

$$\Omega = \{x \in Q : x_k = a_{k i_k} \text{ for some } k = 1, \dots, n\},$$

which is a set of measure zero.

- (ii) We call f simply a step function on Q if there exists some partition P of Q with respect to which it is a step function.
- (iii) The set of step functions on Q is a vector space (the sum of two step functions is a step function w.r.t. a common subpartition of the partitions). This vector space is a subspace of the space of bounded functions on Q .

Integration of Step Functions

3.3.11. Theorem. Let $Q \subset \mathbb{R}^n$ be a cuboid and $f: Q \rightarrow \mathbb{R}$ a step function with respect to some partition P of Q . Then

$$I_P(f) := \sum_{\substack{j_1=1, \dots, m_1 \\ \vdots \\ j_n=1, \dots, m_n}} |Q_{j_1 \dots j_n}| \cdot y_{j_1 \dots j_n}$$

is independent of the choice of the partition P and is called the **integral** of f .

We thus define

$$\int_Q f := I_P(f)$$

for any partition P to be the integral of a step function over a cuboid Q .

The Regulated Integral

Recall that the regulated integral was defined through the following procedure:

- ▶ We defined the set of step functions on an interval $[a, b] \subset \mathbb{R}$.
- ▶ We defined the integral of a step function.
- ▶ Those functions that were uniform limits of sequences of step functions were termed ***regulated functions***. Their integral was defined as the limit of the integrals of a corresponding sequence of step functions.
- ▶ We showed that the continuous functions are regulated. The same is true for piecewise-continuous functions.

We can not extend this strategy to sets in \mathbb{R}^n ; the reason it breaks down is that functions $f: \Omega \rightarrow \mathbb{R}$ on general domains $\Omega \subset \mathbb{R}^n$ **can not be approximated uniformly** by step functions.

The Riemann / Darboux Integral

However, we have an alternative approach at hand in the form of the Riemann integral. This was defined for a function $f: [a, b] \rightarrow \mathbb{R}$ as follows:

- ▶ We defined the set of step functions on an interval $[a, b] \subset \mathbb{R}$.
- ▶ We defined the set of lower step functions w.r.t. f , i.e., those whose values are less than the values of f .
- ▶ We defined the set of upper step functions w.r.t. f , i.e., those whose values are greater than the values of f .
- ▶ If the greatest lower bound for the values of the integrals of upper step functions coincides with the least upper bound for the integrals of lower step functions, this must be the integral of f . This integral is called the **Darboux integral** and is equivalent to the Riemann integral.
- ▶ The Riemann integral coincides with the integral for regulated functions, but can be applied even to functions that are not regulated.

Integration over Cuboids

We will now formulate the definition of the Riemann integral for functions of several variables with real values.

3.3.12. Definition. Let $Q \subset \mathbb{R}^n$ be an n -cuboid and f a bounded real function on Q . let \mathcal{U}_f denote the set of all step functions u on Q such that $u \geq f$ and \mathcal{L}_f the set of all step functions v on Q such that $v \leq f$. The function f is then said to be **(Darboux)-integrable** if

$$\sup_{v \in \mathcal{L}_f} \int_Q v = \inf_{u \in \mathcal{U}_f} \int_Q u.$$

In this case, the **(Darboux) integral of f over Q** , $\int_Q f$, is defined to be this common value.

3.3.13. Theorem. A bounded function $f: Q \rightarrow \mathbb{R}$ is Riemann-integrable if and only if for every $\varepsilon > 0$ there exist step functions u_ε and v_ε such that $v_\varepsilon \leq f \leq u_\varepsilon$ and

$$\int_Q u_\varepsilon - \int_Q v_\varepsilon \leq \varepsilon.$$

Integration over Cuboids

3.3.14. Proposition. Let $Q \subset \mathbb{R}^n$ be an n -cuboid and $f: Q \rightarrow \mathbb{R}$ be bounded and continuous almost everywhere. Then f is (Riemann) integrable.

Proof.

Since f is continuous almost everywhere, we can find a partition of Q such that the set of points where f is discontinuous is contained in cuboids of arbitrarily small measure. Furthermore, since f is bounded, we can find some $C > 0$ such that

$$-C/2 < f(x) < C/2.$$

For any partition P of Q we denote by Q' the union of the induced cuboids on which f is discontinuous. Fix $\varepsilon > 0$ and choose a partition of Q such that

$$|Q'| < \frac{\varepsilon}{2C}.$$

Integration over Cuboids

Proof (continued).

Let $Q'' := Q \setminus Q'$ be the union of the cuboids where f is continuous.
 Choose the partition P in such a way that

$$\sup_{x,y \in Q_{j_1 \dots j_n}} |f(x) - f(y)| \leq \frac{\varepsilon}{2|Q|}$$

for any $Q_{j_1 \dots j_n} \subset Q''$.

Define a step function u as follows:

$$u(x) = \begin{cases} C/2 & x \in Q_{j_1 \dots j_n}, \quad Q_{j_1 \dots j_n} \subset Q', \\ \sup_{x \in Q_{j_1 \dots j_n}} f(x) & x \in Q_{j_1 \dots j_n}, \quad Q_{j_1 \dots j_n} \subset Q'', \\ f(x) & x \in \partial Q_{j_1 \dots j_n}. \end{cases}$$

Clearly, $f \leq u$ on Q .

Integration over Cuboids

Proof (continued).

Similarly, set

$$v(x) = \begin{cases} -C/2 & x \in Q_{j_1 \dots j_n}, \quad Q_{j_1 \dots j_n} \subset Q', \\ \inf_{x \in Q_{j_1 \dots j_n}} f(x) & x \in Q_{j_1 \dots j_n}, \quad Q_{j_1 \dots j_n} \subset Q'', \\ f(x) & x \in \partial Q_{j_1 \dots j_n}, \end{cases}$$

so $v \leq f$ on Q .

Integration over Cuboids

Proof (continued).

Then

$$\begin{aligned} \left| \int_Q u - \int_Q v \right| &\leq \left| \int_{Q'} (u - v) \right| + \left| \int_{Q''} (u - v) \right| \\ &\leq |Q'| \cdot \sup_{x \in Q'} |u(x) - v(x)| + |Q''| \cdot \sup_{x \in Q''} |u(x) - v(x)| \leq |Q'| \cdot \\ &\leq \frac{\varepsilon}{2C} \cdot C + |Q''| \cdot \frac{\varepsilon}{2|Q|} < \varepsilon. \end{aligned}$$

By Theorem 3.3.13, f is integrable.



Integration over Jordan-Measurable Sets

When integrating a function $f: U \rightarrow \mathbb{R}$ defined on some set $U \subset \mathbb{R}^n$, we automatically consider the domain of f to be extended to all of \mathbb{R}^n by setting $f(x) = 0$ for $x \in \mathbb{R}^n \setminus U$. At the same time, we define the **indicator function** for a set $\Omega \subset \mathbb{R}^n$:

$$\mathbb{1}_\Omega(x) = \begin{cases} 1, & x \in \Omega, \\ 0, & \text{otherwise.} \end{cases}$$

Then

$$\mathbb{1}_\Omega(x)f(x) = \begin{cases} f(x) & x \in \Omega, \\ 0 & x \notin \Omega. \end{cases}$$

Integration over Jordan-Measurable Sets

3.3.15. Definition. Let $U \subset \mathbb{R}^n$, $f: U \rightarrow \mathbb{R}$ and $\Omega \subset \mathbb{R}^n$ be a bounded Jordan-measurable set. Then f is said to be integrable on Ω if for every n -cuboid $Q \subset \mathbb{R}^n$ such that $\Omega \subset Q$ the function $\mathbb{1}_\Omega f: Q \rightarrow \mathbb{R}$ is integrable. We then write

$$\int_{\Omega} f := \int_Q f \cdot \mathbb{1}_{\Omega}.$$

for any n -cuboid $Q \supset \Omega$.

We omit the proof of the following result:

3.3.16. Lemma. Let $\Omega \in \mathbb{R}^n$ be a bounded set. Then Ω is Jordan-measurable if and only if its boundary $\partial\Omega$ has Jordan measure zero.

Proposition 3.3.14 and Lemma 3.3.16 immediately yield:

3.3.17. Corollary. Let $\Omega \subset \mathbb{R}^n$ be a bounded Jordan-measurable set and let $f: \Omega \rightarrow \mathbb{R}$ be continuous a.e. Then f is integrable on Ω .

Basic Properties of the Integral

From the definition of the integral and measurability of sets, we have the following result:

3.3.18. Lemma.

- (i) Let $\Omega \subset \mathbb{R}^n$ be a measurable set. Then

$$|\Omega| = \int_{\Omega} 1.$$

- (ii) Let $\Omega \subset \mathbb{R}^n$ be a set of measure zero and and $f: \Omega \rightarrow \mathbb{R}$ some function that is integrable on Ω . Then $\int_{\Omega} f = 0$.
- (iii) Let $\Omega \subset \mathbb{R}^n$ and $\Omega' \subset \Omega$ be measurable sets and $f: \mathbb{R}^n \rightarrow \mathbb{R}$ integrable on Ω . Then f is also integrable on Ω' .
- (iv) Let $\Omega, \Omega' \subset \mathbb{R}^n$ measurable sets and $f: \mathbb{R}^n \rightarrow \mathbb{R}$ integrable on both of them. Then f is integrable on $\Omega \cup \Omega'$ and

$$\int_{\Omega \cup \Omega'} f = \int_{\Omega} f + \int_{\Omega'} f - \int_{\Omega \cap \Omega'} f.$$

Vector Fields and Line Integrals

Circulation and Flux

The Riemann Integral and Measurable Sets

Integration in Practice

Surfaces and Surface Integrals

The Theorems of Gauß and Stokes

Practical Integration over Cuboids

The following result lets us reduce integrals over n -cuboids to separate integrals over n_1 - and n_2 -cuboids, where $n_1 + n_2 = n$. Since we know how to integrate over 1-cuboids (intervals!), this is a powerful tool for evaluating general integrals.

3.4.1. Fubini's Theorem. Let Q_1 be an n_1 -cuboid and Q_2 an n_2 -cuboid so that $Q := Q_1 \times Q_2 \subset \mathbb{R}^{n_1+n_2}$ is an $(n_1 + n_2)$ -cuboid. Assume that $f: Q \rightarrow \mathbb{R}$ is integrable on Q and that for every $x \in Q_1$ the integral

$$g(x) = \int_{Q_2} f(x, \cdot)$$

exists. Then

$$\int_Q f = \int_{Q_1 \times Q_2} f = \int_{Q_1} g = \int_{Q_1} \left(\int_{Q_2} f \right).$$

We omit the proof of this theorem; the statement can be shown as in the previous results by considering step functions.

Practical Integration over Cuboids

3.4.2. Example. Consider the 2-cuboid $Q = [0, 1] \times [0, 2]$ and the function $f: Q \rightarrow \mathbb{R}$, $f(x, y) = yx^2 + 2y^2$. The integral

$$g(x) = \int_{[0,2]} f(x, \cdot) = \int_0^2 f(x, y) dy = \left[\frac{x^2}{2} y^2 + \frac{2}{3} y^3 \right]_{y=0}^2 = 2x^2 + \frac{16}{3}$$

exists for every $x \in [0, 1]$, so we can apply Fubini's Theorem to yield

$$\begin{aligned} \int_Q f &= \int_{[0,1]} g = \int_0^1 g(x) dx = \int_0^1 \left(\int_0^2 f(x, y) dy \right) dx \\ &= \frac{2}{3} [x^3 + 8x]_0^1 = \frac{18}{3}. \end{aligned}$$

We can thus use Fubini's theorem to iteratively reduce integrals over n -cuboids to integrals over intervals, which we know well how to calculate.

Practical Integration over Cuboids

We will often omit the parentheses when evaluating multiple integrals, e.g., we will write

$$\int_0^1 \int_0^2 f(x, y) dy dx = \int_0^1 \left(\int_0^2 f(x, y) dy \right) dx$$

If the hypotheses of Fubini's Theorem are satisfied, we can therefore write

$$\int_Q f = \int_{a_n}^{b_n} \cdots \int_{a_2}^{b_2} \int_{a_1}^{b_1} f(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n.$$

We will often abbreviate $dx := dx_1 dx_2 \dots dx_n$, writing

$$\int_Q f = \int_Q f(x) dx.$$

Ordinate and Simple Regions in \mathbb{R}^2

Of course, we do not just want to integrate over cuboids but also over more complicated domains. For most purposes it is sufficient to consider regions whose boundaries can be expressed by the graphs of functions.

3.4.3. Definition. A set $D \subset \mathbb{R}^2$ is called an **ordinate region with respect to x_2** , if there exists an interval $I \subset \mathbb{R}$ and continuous, almost everywhere differentiable functions $\varphi_1, \varphi_2: I \rightarrow \mathbb{R}$ such that

$$D = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \in I, \varphi_1(x_1) \leq x_2 \leq \varphi_2(x_1)\}.$$

If the role of x_1 and x_2 above is interchanged, we say that D is **an ordinate region with respect to x_1** .

If $D \subset \mathbb{R}^2$ is an ordinate region both with respect to x_1 and x_2 , we say that D is a **simple region**.

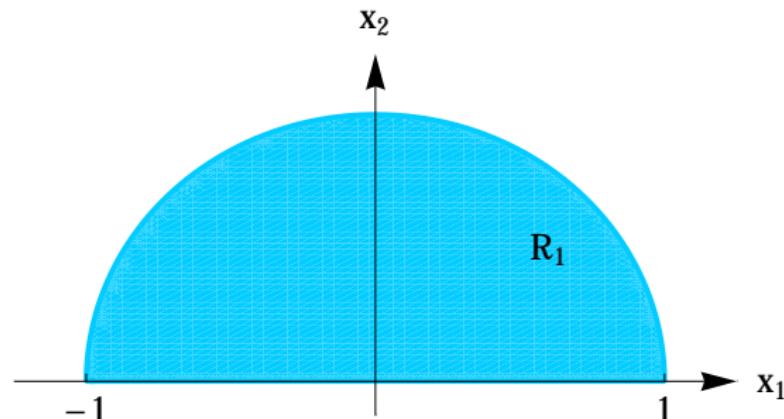
Ordinate and Simple Regions in \mathbb{R}^2

3.4.4. Example. The half-disk region

$$R_1 = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 \geq 0, x_1^2 + x_2^2 \leq 1\}$$

is a simple region, because we can write

$$\begin{aligned} R_1 &= \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \in [-1, 1], 0 \leq x_2 \leq \sqrt{1 - x_1^2}\} \\ &= \{(x_1, x_2) \in \mathbb{R}^2 : x_2 \in [0, 1], -\sqrt{1 - x_2^2} \leq x_1 \leq \sqrt{1 - x_2^2}\}. \end{aligned}$$



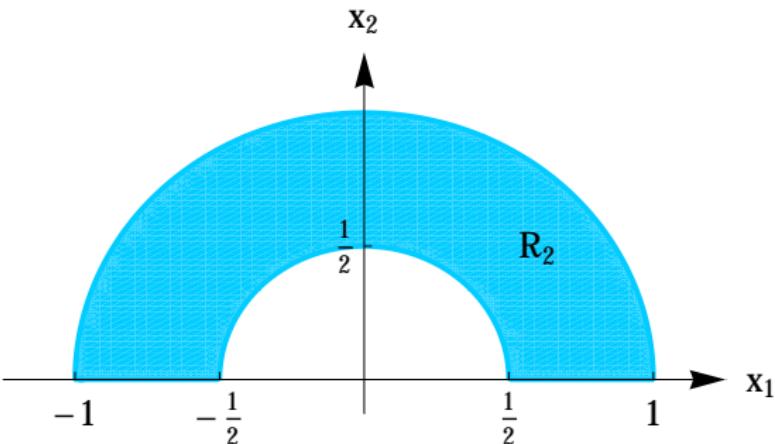
Ordinate and Simple Regions in \mathbb{R}^2

3.4.5. Example. The upper half-annulus

$$R_2 = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 \geq 0, 1/4 \leq x_1^2 + x_2^2 \leq 1\}$$

is an ordinate region with respect to x_2 but not with respect to x_1 . We can write

$$R_2 = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \in [-1, 1], f(x_1) \leq x_2 \leq \sqrt{1 - x_1^2}\}.$$



where

$$f(x) = \sqrt{1/4 - x^2}$$

if $|x| < 1/2$ and

$$f(x) = 0$$

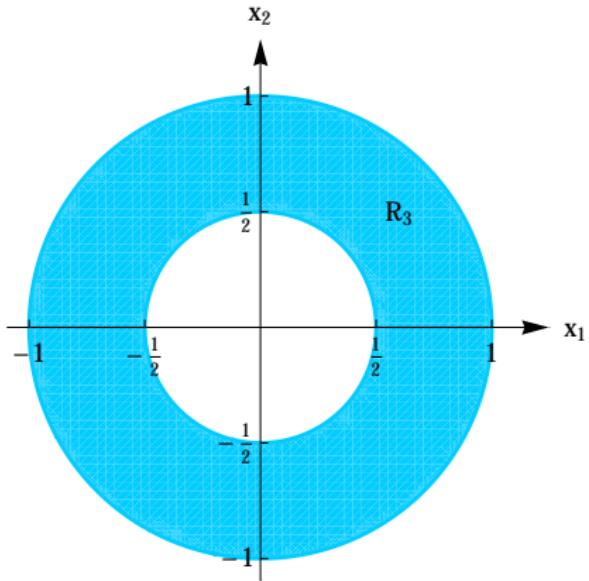
otherwise.

Ordinate and Simple Regions in \mathbb{R}^2

3.4.6. Example. The annulus

$$R_3 = \{(x_1, x_2) \in \mathbb{R}^2 : 1/4 \leq x_1^2 + x_2^2 \leq 1\}$$

is not an ordinate region (but can be expressed as the union of two ordinate regions).



Ordinate Regions in \mathbb{R}^n

We now generalize ordinate regions to \mathbb{R}^n . For $x \in \mathbb{R}^n$ we define

$$\hat{x}^{(k)} := (x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n) \in \mathbb{R}^{n-1}$$

as the vector x with the k th component omitted.

3.4.7. Definition. A subset $U \subset \mathbb{R}^n$ is said to be an **ordinate region (with respect to x_k)** if there exists a measurable set $\Omega \subset \mathbb{R}^{n-1}$ and continuous, almost everywhere differentiable functions $\varphi_1, \varphi_2: \Omega \rightarrow \mathbb{R}$, such that

$$U = \{(x \in \mathbb{R}^n: x \in \Omega, \varphi_1(\hat{x}^{(k)}) \leq x_k \leq \varphi_2(\hat{x}^{(k)})\}.$$

If U is an ordinate region with respect to each x_k , $k = 1, \dots, n$, it is said to be a **simple region**.

3.4.8. Lemma. Any ordinate region is measurable.

Ordinate Regions

3.4.9. Example. The unit ball in \mathbb{R}^3 ,

$$B^3 := \{x \in \mathbb{R}^3 : \|x\| \leq 1\}$$

is an ordinate region, since we can write

$$B^3 = \left\{ x \in \mathbb{R}^3 : (x_1, x_2) \in B^2, -\sqrt{1 - x_1^2 - x_2^2} \leq x_3 \leq \sqrt{1 - x_1^2 - x_2^2} \right\}$$

where $B^2 := \{x \in \mathbb{R}^2 : \|x\| \leq 1\}$. Of course, we still need to check that B^2 is measurable. However,

$$B^2 = \left\{ x \in \mathbb{R}^2 : x_1 \in [-1, 1], -\sqrt{1 - x_1^2} \leq x_2 \leq \sqrt{1 - x_1^2} \right\}$$

is itself an ordinate set, and since the interval $[-1, 1]$ is measurable, so is B^2 .

Integrals on Ordinate Regions

For an ordinate region $U \in \mathbb{R}^n$ with respect to x_k over a measurable set Ω the indicator function $\mathbb{1}_U$ takes the form

$$\mathbb{1}_U(x) = \mathbb{1}_{\Omega} \cdot \mathbb{1}_{[\varphi_1(\hat{x}^{(k)}), \varphi_2(\hat{x}^{(k)})]}(x_k)$$

It then follows that

$$\int_U f(x) dx_1 \dots dx_n = \int_{\Omega} \left(\int_{\varphi_1(\hat{x}^{(k)})}^{\varphi_2(\hat{x}^{(k)})} f(x) dx_k \right) d\hat{x}^{(k)}$$

if $\int_{\varphi_1(\hat{x}^{(k)})}^{\varphi_2(\hat{x}^{(k)})} f(x) dx_k$ exists for every $\hat{x}^{(k)} \in \Omega$.

Integrals on Ordinate Regions

3.4.10. Example. The volume of a Jordan-measurable measurable set $\Omega \subset \mathbb{R}^n$ is given by

$$|\Omega| = \int_{\Omega} 1.$$

As an example, we calculate the volume of the three-dimensional unit ball B^3 . Writing B^3 as ordinate regions, we have

$$\begin{aligned} |B^3| &= \int_{B^3} 1 = \int_{B^2} \int_{-\sqrt{1-x_1^2-x_2^2}}^{\sqrt{1-x_1^2-x_2^2}} 1 \, dx_3 \, dx_2 \, dx_1 \\ &= 2 \int_{B^2} \sqrt{1 - x_1^2 - x_2^2} \, dx_2 \, dx_1 \\ &= 2 \int_{-1}^1 \left(\int_{-\sqrt{1-x_1^2}}^{\sqrt{1-x_1^2}} \sqrt{1 - x_1^2 - x_2^2} \, dx_2 \right) dx_1 \end{aligned}$$

We now substitute $y_2 = x_2 / \sqrt{1 - x_1^2}$ in the inner integral.

Integrals on Ordinate Regions

$$\begin{aligned}|B^3| &= 2 \int_{-1}^1 \left((1 - x_1^2) \int_{-1}^1 \sqrt{1 - y_2^2} dy_2 \right) dx_1 \\&= 2 \int_{-1}^1 (1 - x_1^2) dx_1 \cdot \int_{-1}^1 \sqrt{1 - y_2^2} dy_2 \\&= 8 \int_0^1 (1 - x_1^2) dx_1 \cdot \int_0^1 \sqrt{1 - y_2^2} dy_2 \\&= \frac{16}{3} \int_0^1 \sqrt{1 - y_2^2} dy_2\end{aligned}$$

Substituting $y_2 = \sin \theta$, we obtain

$$|B^3| = \frac{16}{3} \int_0^{\pi/2} \cos^2 \theta d\theta = \frac{4}{3}\pi,$$

as expected.

Bodies, Moments and Center of Mass

A **rigid body** is a set $B \subset \mathbb{R}^n$ (in physics, $n = 2, 3$) with a **mass distribution** $\varrho: B \rightarrow \mathbb{R}$. The **mass** of the body is given by

$$M(B) = \int_B \varrho.$$

We define the (first) moments of B by

$$m_k(B) = \int_B x_k \varrho, \quad k = 1, \dots, n.$$

Then the center of mass is given by

$$x_c(B) = \frac{1}{M(B)} \begin{pmatrix} m_1(B) \\ \vdots \\ m_n(B) \end{pmatrix}.$$

If $\varrho = 1$ on B , then $x_c(B)$ represents the geometric center and $M(B) = |B|$ the volume of B .

Bodies, Moments and Center of Mass

3.4.11. Example. Let $B \subset \mathbb{R}^2$ be given by

$$B = \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, 0 \leq y \leq x^2\}.$$

with $\varrho(x, y) = x + y$. Then

$$\begin{aligned} M(B) &= \int_B \varrho = \int_0^1 \int_0^{x^2} (x + y) dy dx = \int_0^1 [xy + y^2/2]_0^{x^2} dx \\ &= \int_0^1 (x^3 + x^4/2) dx = 1/4 + 1/10 = 7/20. \end{aligned}$$

The moments of B are

$$\begin{aligned} m_1(B) &= \int_B x\varrho = \int_0^1 \int_0^{x^2} x(x + y) dy dx \\ &= \int_0^1 (x^4 + x^5/2) dx = 1/5 + 1/12 = 17/60, \end{aligned}$$

Bodies, Moments and Center of Mass

and

$$\begin{aligned}m_2(B) &= \int_B y\varrho = \int_0^1 \int_0^{x^2} y(x+y) dy dx \\&= \int_0^1 (x^5/2 + x^6/3) dx = 1/12 + 1/21 = 11/84.\end{aligned}$$

Hence the center of mass is given by

$$x_c(B) = \frac{1}{M(B)} \begin{pmatrix} m_1(B) \\ m_2(B) \end{pmatrix} = \frac{20}{7} \begin{pmatrix} 17/60 \\ 11/84 \end{pmatrix} = \begin{pmatrix} 17/21 \\ 55/147 \end{pmatrix}$$

The Substitution Rule

A powerful tool in evaluating integrals is the substitution rule, which takes on an analogous form to that for functions of one variable. We will merely state, and not prove, this result.

3.4.12. Substitution Rule. Let $\Omega \subset \mathbb{R}^n$ be open and $g: \Omega \rightarrow \mathbb{R}^n$ injective and continuously differentiable. Suppose that $\det J_g(y) \neq 0$ for all $y \in \Omega$. Let K be a compact measurable subset of Ω . Then $g(K)$ is compact and measurable and if $f: g(K) \rightarrow \mathbb{R}$ is integrable, then

$$\int_{g(K)} f(x) dx = \int_K f(g(y)) \cdot |\det J_g(y)| dy.$$

Polar Coordinates

3.4.13. Examples. The most important substitutions are transformations to cylindrical or spherical/polar coordinates.

- (i) Polar coordinates in \mathbb{R}^2 are defined by a map

$$\phi: (0, \infty) \times [0, 2\pi) \rightarrow \mathbb{R}^2 \setminus \{0\}, \quad (r, \phi) \mapsto (x, y)$$

where

$$x = r \cos \phi, \quad y = r \sin \phi.$$

Note that this map is bijective and even C^∞ in the interior of its domain. An alternative (but rarely used) version of polar coordinates would map $x = r \sin \phi$, $y = r \cos \phi$. This simply corresponds to a different geometrical interpretation of the angle ϕ . In any case,

$$|\det J_\phi(r, \phi)| = \left| \det \begin{pmatrix} \cos \phi & -r \sin \phi \\ \sin \phi & r \cos \phi \end{pmatrix} \right| = r$$

Cylindrical Coordinates

(ii) Cylindrical coordinates in \mathbb{R}^3 are given through a map

$$\phi: (0, \infty) \times [0, 2\pi) \times \mathbb{R} \rightarrow \mathbb{R}^3 \setminus \{0\}, \quad (r, \phi, \zeta) \mapsto (x, y, z)$$

defined by

$$x = r \cos \phi, \quad y = r \sin \phi, \quad z = \zeta$$

In this case,

$$|\det J_\phi(r, \phi, \zeta)| = \left| \det \begin{pmatrix} \cos \phi & -r \sin \phi & 0 \\ \sin \phi & r \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \right| = r$$

Spherical Coordinates in \mathbb{R}^3

(iii) Spherical coordinates in \mathbb{R}^3 are often defined through a map

$$\begin{aligned}\phi: (0, \infty) \times [0, 2\pi) \times (0, \pi) &\rightarrow \mathbb{R}^3 \setminus \{0\}, \quad (r, \phi, \theta) \mapsto (x, y, z), \\ x &= r \cos \phi \sin \theta, \quad y = r \sin \phi \sin \theta, \quad z = r \cos \theta.\end{aligned}$$

Of course, there is a certain freedom in defining θ and ϕ , so there are alternative formulations. The modulus of the determinant of the Jacobian is given by

$$\begin{aligned}|\det J_\phi(r, \phi, \theta)| &= \left| \det \begin{pmatrix} \cos \phi \sin \theta & -r \sin \phi \sin \theta & r \cos \phi \cos \theta \\ \sin \phi \sin \theta & r \cos \phi \sin \theta & r \sin \phi \cos \theta \\ \cos \theta & 0 & -r \sin \theta \end{pmatrix} \right| \\ &= r^2 \sin \theta\end{aligned}$$

Spherical Coordinates in \mathbb{R}^n

(iv) In \mathbb{R}^n , we can define spherical coordinates by

$$x_1 = r \cos \theta_1$$

$$x_2 = r \sin \theta_1 \cos \theta_2$$

$$x_3 = r \sin \theta_1 \sin \theta_2 \cos \theta_3$$

$$\vdots$$

$$x_{n-1} = r \sin \theta_1 \sin \theta_2 \dots \sin \theta_{n-2} \cos \theta_{n-1}$$

$$x_n = r \sin \theta_1 \sin \theta_2 \dots \sin \theta_{n-2} \sin \theta_{n-1}$$

with $r > 0$ and $0 < \theta_k < \pi$, $k = 1, \dots, n-2$, and $0 < \theta_{n-1} < 2\pi$.

Here, the determinant of the Jacobian can be shown to be

$$|\det J_\phi(r, \theta_1, \dots, \theta_{n-1})| = r^{n-1} \sin^{n-2} \theta_1 \sin^{n-3} \theta_2 \dots \sin \theta_{n-1}.$$

The Substitution Rule in Practice

Using spherical coordinates in \mathbb{R}^3 as an example, we write

$$\int_{\Omega} f = \int_{\Omega} f(x) dx = \int_{\phi^{-1}(\Omega)} f \circ \Phi(r, \theta, \phi) \cdot |\det J_{\phi}(r, \theta, \phi)| dr d\theta d\phi$$

The terms dx and $|\det J_{\phi}(r, \theta, \phi)| dr d\theta d\phi$ are often referred to as **volume elements**, and one sometimes writes

$$dx = |\det J_{\phi}(r, \theta, \phi)| dr d\theta d\phi.$$

Physicists like to interpret dx as an “infinitesimally small volume” whose volume is changed when transforming by ϕ^{-1} to $dr d\theta d\phi$. Thus $|\det J_{\phi}(r, \theta, \phi)|$ (which can be interpreted as the size of the parallelepiped spanned by the tangent vectors $\frac{\partial x}{\partial r}, \frac{\partial x}{\partial \theta}, \frac{\partial x}{\partial \phi}$ at x) corrects this change in volume. These ideas can be made rigorous, but we will not pursue them further.

The Substitution Rule in Practice

3.4.14. Example. We can again calculate the volume of the unit ball $B^3 \subset \mathbb{R}^3$. Using spherical coordinates,

$$\begin{aligned}|B^3| &= \int_{B^3} 1 = \int_0^{2\pi} \int_0^\pi \int_0^1 r^2 \sin \theta \, dr \, d\theta \, d\phi \\&= 2\pi \int_0^\pi \sin \theta \, d\theta \cdot \int_0^1 r^2 \, dr \\&= \frac{4\pi}{3} \int_0^{\pi/2} \sin \theta \, d\theta = \frac{4\pi}{3}.\end{aligned}$$

Note that B^3 is **not** given by

$$\{(x, y, z) \in \mathbb{R}^3 : x = r \cos \phi \sin \theta, y = r \sin \phi \sin \theta, z = r \cos \theta, 0 \leq \phi < 2\pi, 0 < \theta < \pi, 0 < r \leq 1\}$$

because this set does not include the set $\{(0, 0, x) : x \in [-1, 1]\}$. Since this set is of measure zero (as is the boundary S^2 of B^3) our calculation remains correct.

Gravitational Potential

3.4.15. Example. We want to calculate the gravitational potential of a homogenous solid ball in \mathbb{R}^3 of mass M and radius R at a point $p \in \mathbb{R}^3$ with distance $r > R$ from the center of the sphere. This potential is given by

$$U(p) = -G \int_{B^3} \frac{\varrho(\cdot)}{\text{dist}(p, \cdot)}$$

where ϱ is the mass density of the sphere. In our case,

$$\varrho = \frac{M}{|B^3|} = \frac{3M}{4\pi R^3} \cdot \mathbb{1}_{B^3}.$$

Due to the symmetry of the problem, we may choose coordinates such that $p = (0, 0, r)$ and introduce polar coordinates

$$x_1 = \rho \cos \phi \sin \theta, \quad x_2 = \rho \sin \phi \sin \theta, \quad x_3 = \rho \cos \theta$$

with $0 \leq \phi < 2\pi$, $0 \leq \theta < \pi$, $0 < \rho \leq R$.

Gravitational Potential

Then

$$\begin{aligned}\text{dist}(p, x(r, \phi, \theta)) &= \sqrt{(\rho \cos \phi \sin \theta)^2 + (\rho \sin \phi \sin \theta)^2 + (\rho \cos \theta - r)^2} \\ &= \sqrt{\rho^2 + r^2 - 2r\rho \cos \theta}\end{aligned}$$

and

$$\begin{aligned}U(p) &= -\frac{3MG}{4\pi R^3} \int_{B^3} \frac{dx}{\text{dist}(p, x)} \\ &= -\frac{3MG}{4\pi R^3} \int_0^R \int_0^{2\pi} \int_0^\pi \frac{\rho^2 \sin \theta \, d\theta \, d\phi \, d\rho}{\sqrt{\rho^2 + r^2 - 2r\rho \cos \theta}} \\ &= -\frac{3MG}{4\pi R^3} \int_0^R \int_0^{2\pi} \frac{\rho^2}{r\rho} \sqrt{\rho^2 + r^2 - 2r\rho \cos \theta} \Big|_0^\pi \, d\phi \, d\rho \\ &= -\frac{3MG}{2R^3 r} \int_0^R \rho \left(\sqrt{\rho^2 + r^2 + 2r\rho} - \sqrt{\rho^2 + r^2 - 2r\rho} \right) \, d\rho\end{aligned}$$

Gravitational Potential

Continuing,

$$\begin{aligned} U(p) &= -\frac{3MG}{2R^3r} \int_0^R \rho \left(\sqrt{(\rho+r)^2} - \sqrt{(\rho-r)^2} \right) d\rho \\ &= -\frac{3MG}{2R^3r} \int_0^R \rho (\rho+r - |\rho-r|) d\rho \end{aligned}$$

Since $r > R > \rho$, we have

$$\begin{aligned} U(p) &= -\frac{3MG}{2R^3r} \int_0^R \rho (\rho+r - (r-\rho)) d\rho \\ &= -\frac{3MG}{R^3r} \int_0^R \rho^2 d\rho = -\frac{MG}{r} \end{aligned}$$

Thus the potential induced by a sphere of mass M and radius R at a point with distance $r > R$ from the center of the sphere is the same as that induced by a point mass with mass M situated at the center of the sphere.

Gravitational Potential

3.4.16. Remarks. In the physical literature, this is part of what is called the ***shell theorem***. You will study the other parts of this theorem in the assignments.

An analogous formula holds for the electrostatic potential induced by a body with charge density ϱ .

If the mass/charge distribution of the sphere is not uniform, then the integral becomes much more difficult to solve. One then expands the integrand using the ***generating function of the Legendre polynomials*** P_l , $l \in \mathbb{N}$,

$$\frac{1}{\sqrt{1 + 2xt + x^2}} = \sum_{l=0}^{\infty} P_l(x) t^l$$

which for every $x \in [-1, 1]$ has radius of convergence 1. The same expansion can be used when summing over several discrete point charges/masses, where it is then called a ***multi-pole expansion***.

Improper Integrals

Just as for integrals of a single variable, we can treat improper Riemann integrals of functions $f : \mathbb{R}^n \rightarrow \mathbb{R}$ over measurable sets $\Omega \subset \mathbb{R}^n$. These occur if either

1. f is unbounded or
2. Ω is unbounded.

In either case, one considers the improper integral as a suitable limit of “proper” integrals; if the limit exists, so does the improper integral.

3.4.17. Example. Our aim is to prove that the Gauß integral

$$\int_{-\infty}^{\infty} e^{-x^2/2} dx$$

exists and equals $\sqrt{2\pi}$. First, consider the integral

$$I(a) = \int_{-a}^a e^{-x^2/2} dx.$$

The Gauß Integral

Since the integrand is positive and continuous, $I(a)$ exists and is increasing. For $a > 1$,

$$\begin{aligned} I(a) &< \int_{-a}^{-1} -xe^{-x^2/2} dx + \int_{-1}^1 e^{-x^2/2} dx + \int_1^a xe^{-x^2/2} dx \\ &= 2e^{-1/2} - 2e^{-a^2/2} + \int_{-1}^1 e^{-x^2/2} dx \\ &\xrightarrow{a \rightarrow \infty} 2e^{-1/2} + \int_{-1}^1 e^{-x^2/2} dx < \infty, \end{aligned}$$

so $I(a)$ is bounded. It follows that $\lim_{a \rightarrow \infty} I(a) =: \int_{-\infty}^{\infty} e^{-x^2/2} dx$ exists.

We now consider

$$I(a)^2 = \left(\int_{-a}^a e^{-x^2/2} dx \right) \left(\int_{-a}^a e^{-y^2/2} dy \right).$$

The Gauß Integral

By Fubini's theorem, we can write

$$I(a)^2 = \int_{Q_a} e^{-(x^2+y^2)/2} dx dy,$$

where $Q_a = [-a, a] \times [-a, a]$.

Now $B_a(0) \subset Q_a \subset B_{2a}(0)$, where $B_r(0) = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < r^2\}$, so

$$\int_{B_a(0)} e^{-(x^2+y^2)/2} dx dy \leq I(a)^2 \leq \int_{B_{2a}(0)} e^{-(x^2+y^2)/2} dx dy.$$

Using polar coordinates, we calculate

$$\begin{aligned} \int_{B_R(0)} e^{-(x^2+y^2)/2} dx dy &= \int_0^{2\pi} \int_0^R e^{-r^2/2} r dr d\phi \\ &= 2\pi(1 - e^{-R^2/2}) \xrightarrow{R \rightarrow \infty} 2\pi \end{aligned}$$

This implies $\lim_{a \rightarrow \infty} I(a)^2 = 2\pi$ and hence $\int_{-\infty}^{\infty} e^{-x^2/2} dx = \sqrt{2\pi}$.

Fundamental Theorems of Integration

In our previous integrals, we have often been able to formulate a “fundamental theorem” of the type

“integral of f over domain” = “values of primitive of f on the boundary”

For example, for a real integrable function defined on an interval,

$$\int_a^b f(x) dx = F(b) - F(a)$$

A similar formula was found to hold for a potential field integrated along a curve, where the potential of the vector field played the role of the primitive.

It turns out that suitable generalizations of this principle hold in higher dimensions. These **fundamental theorems of vector analysis** are among the most important theorems for engineering applications that we will study this term.

Green's Theorem

3.4.18. **Green's Theorem.** Let $R \subset \mathbb{R}^2$ be a bounded, simple region and $\Omega \supset R$ an open set containing R . Let $F: \Omega \rightarrow \mathbb{R}^2$ be a continuously differentiable vector field. Then

$$\int_{\partial R^*} F \cdot d\vec{s} = \int_R \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) dx \quad (3.4.1)$$

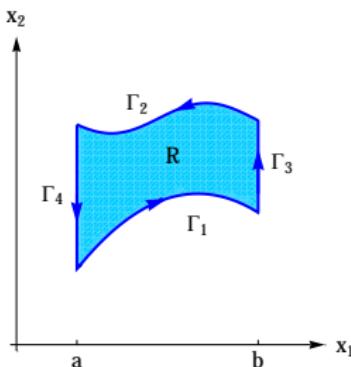
where ∂R^* denotes the boundary curve of R with positive (counter-clockwise) orientation.

3.4.19. **Remark.** Actually, Green's Theorem is valid not just for simple regions, but more generally for bounded domains whose boundary is a rectifiable curve (i.e., a curve of finite length). However, the proof becomes much more complicated in this general case and so we restrict ourselves to simple domains here.

Green's Theorem

Proof.

Let R be an ordinate region with respect to x_2 , Ω an open set containing R and $F_1: \Omega \rightarrow \mathbb{R}$ continuously differentiable.



The boundary ∂R of R is the union of the curves

$$\Gamma_1 = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \in [a, b], x_2 = \varphi_1(x_1)\},$$

$$\Gamma_2 = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \in [a, b], x_2 = \varphi_2(x_1)\},$$

$$\Gamma_3 = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 = b, x_2 \in [\varphi_1(b), \varphi_2(b)]\},$$

$$\Gamma_4 = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 = a, x_2 \in [\varphi_1(a), \varphi_2(a)]\}.$$

We will later imbue the boundary with positive orientation, so we discuss integrals along the curves $\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4$ in the directions indicated by the arrows.

Green's Theorem

Proof (continued).

Consider the integral

$$\begin{aligned} \int_R \frac{\partial F_1}{\partial x_2} dx &= \int_a^b \int_{\varphi_1(x_1)}^{\varphi_2(x_1)} \frac{\partial F_1}{\partial x_2} dx_2 dx_1 \\ &= \int_a^b F_1(x_1, \varphi_2(x_1)) dx_1 - \int_a^b F_1(x_1, \varphi_1(x_1)) dx_1. \quad (3.4.2) \end{aligned}$$

We now integrate the vector field

$$\tilde{F}: \Omega \rightarrow \mathbb{R}^2, \quad \tilde{F}(x) = \begin{pmatrix} F_1(x) \\ 0 \end{pmatrix},$$

along Γ_1 and $-\Gamma_2$, using the respective parametrizations

$$\gamma^{(1)}(t) = \begin{pmatrix} t \\ \varphi_1(t) \end{pmatrix}, \quad \gamma^{(2)}(t) = \begin{pmatrix} t \\ \varphi_2(t) \end{pmatrix}, \quad t \in [a, b].$$

Green's Theorem

Proof (continued).

A quick calculation yields

$$\int_{\Gamma_1} \begin{pmatrix} F_1 \\ 0 \end{pmatrix} d\vec{s} = \int_a^b F_1 \circ \gamma^{(1)}(t) dt = \int_a^b F_1(t, \varphi_1(t)) dt,$$

$$\int_{-\Gamma_2} \begin{pmatrix} F_1 \\ 0 \end{pmatrix} d\vec{s} = - \int_a^b F_1 \circ \gamma^{(2)}(t) dt = - \int_a^b F_1(t, \varphi_2(t)) dt.$$

It is also easy to see that

$$\int_{\Gamma_3} \begin{pmatrix} F_1 \\ 0 \end{pmatrix} d\vec{s} = \int_{\Gamma_4} \begin{pmatrix} F_1 \\ 0 \end{pmatrix} d\vec{s} = 0$$

so we find

$$\oint_{\partial R} \begin{pmatrix} F_1 \\ 0 \end{pmatrix} d\vec{s} = \int_a^b F_1(t, \varphi_1(t)) dt - \int_a^b F_1(t, \varphi_2(t)) dt \quad (3.4.3)$$

Green's Theorem

Proof (continued).

Putting (3.4.2) and (3.4.2) together,

$$\int_{\partial R} \begin{pmatrix} F_1 \\ 0 \end{pmatrix} d\vec{s} = - \int_R \frac{\partial F_1}{\partial x_2} dx. \quad (3.4.4)$$

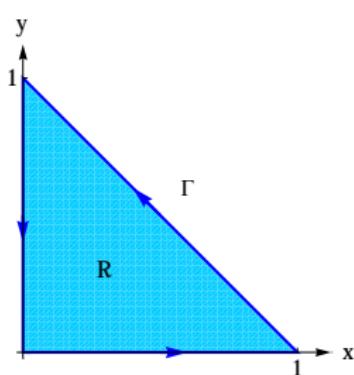
Repeating a similar argument with a scalar function $F_2: \Omega \rightarrow \mathbb{R}$ and representing R as an ordinate region with respect to x_1 yields

$$\int_{\partial R} \begin{pmatrix} 0 \\ F_2 \end{pmatrix} d\vec{s} = \int_R \frac{\partial F_2}{\partial x_1} dx. \quad (3.4.5)$$

Adding (3.4.4) and (3.4.5) then gives (3.4.1). □

Green's Theorem

3.4.20. Example. We wish to calculate the integral $\int_{\Gamma} x^4 dx + xy dy$, where $\Gamma \subset \mathbb{R}^2$ is the triangle with vertices $(0, 0)$, $(0, 1)$ and $(1, 0)$, oriented positively.



We note that the vector field $F(x, y) = (x^4, xy)$ is defined on all of \mathbb{R}^2 and the region R bounded by the triangle is a simple region. Instead of evaluating three separate line integrals (one for each edge of the triangle) we can apply Green's Theorem. We note that

$$R = \{(x, y) \in \mathbb{R}^2 : x \in [0, 1], 0 \leq y \leq 1 - x\}.$$

Then

$$\int_{\Gamma} x^4 dx + xy dy = \int_R (y - 0) d(x, y) = \int_0^1 \int_0^{1-x} y dy dx = \frac{1}{6}$$

Measurement of Area – the Planimeter

For $F(x_1, x_2) = (-x_2, x_1)$ we obtain

$$|R| = \int_R 1 \, dx = \frac{1}{2} \int_{\partial R} \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix} \cdot d\vec{s}.$$

Hence the integral of a vector field around the boundary of a region can be used to determine the area of that region.

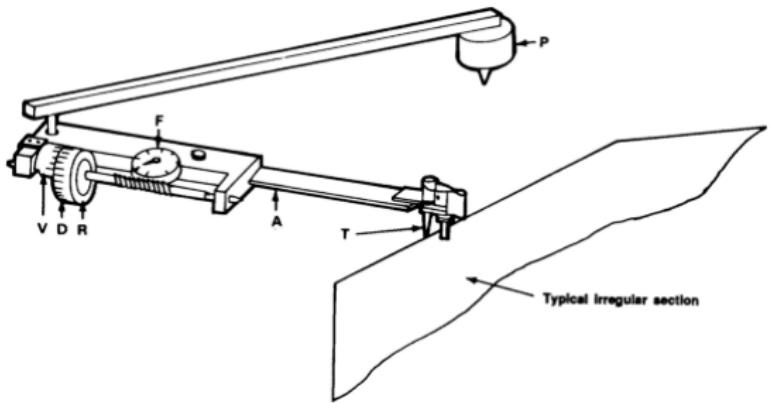
Several measurement instruments, known as **planimeters**, have been developed to implement this. The most successful version is the **polar planimeter**, invented by the Swiss mathematician Jakob Amsler in 1854. Previous planimeters (the first was constructed in 1814) were not as accurate as the polar planimeter, and this remains the most common form today.

A description of the basic functioning of the polar planimeter is quoted on the next slide. The source is an article on military surveys and earthwork constructions by GlobalSecurity.org.

Functioning of a Polar Planimeter

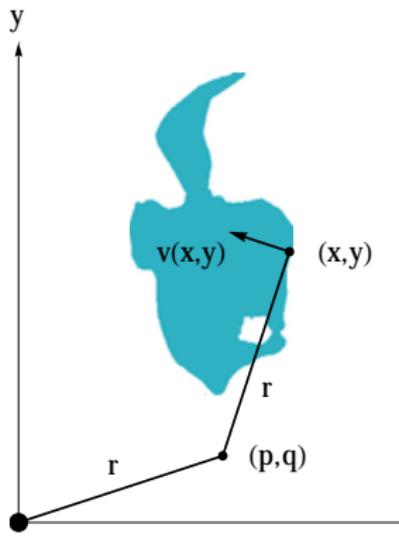
"The planimeter [...] touches the paper at three points: the anchor point, P; the tracing point, T; and the roller, R. The adjustable arm, A, is graduated to permit adjustment to the scale of the plot. This adjustment provides a direct ratio between the area traced by the tracing point and the revolutions of the roller. As the tracing point is moved over the paper, the drum, D, and the disk, F, revolve. The disk records the revolutions of the roller in units of tenths; the drum, in hundredths; and the vernier, V, in thousandths."

<http://www.globalsecurity.org/military/library/policy/army/fm/5-430-00-1/CH3.htm>



Polar Planimeter in Use. GlobalSecurity.org. Web. 22 July 2012

Principle of a Polar Planimeter



To understand the principle of the polar planimeter, consider the sketch at right:

Assume for simplicity that both arms of the planimeter have equal length r . When tracing the boundary curve of the pictured shape, the position of the pivot (p, q) changes as a function of (x, y) (the point (p, q) is unique if we require that the angle between the two arms is less than π).

The planimeter vector field v is given by

$$v(x, y) = \frac{1}{r} \begin{pmatrix} -(y - q(x, y)) \\ x - p(x, y) \end{pmatrix}$$

where $r^2 = (y - q(x, y))^2 + (x - p(x, y))^2$. The disk, roller and vernier of the planimeter record the integral of $v(x, y)$ along the boundary.

Principle of a Polar Planimeter

The point $(p(x, y), q(x, y))$ is given by the intersection of two circles of radius r about the origin and about (x, y) . From the two equations

$$p^2 + q^2 = r^2, \quad (p - x)^2 + (q - y)^2 = r^2,$$

we obtain

$$x^2 + y^2 - 2px - 2qy = 0.$$

The implicit equations for p and q ,

$$x^2 + y^2 - 2px - 2\sqrt{r^2 - p^2}y = 0,$$

$$x^2 + y^2 - 2qy - 2\sqrt{r^2 - q^2}x = 0,$$

yield (see assignments)

$$\frac{\partial p}{\partial x} + \frac{\partial q}{\partial y} = 1.$$

Principle of a Polar Planimeter

We then have

$$\begin{aligned}\text{planimeter reading} &= \int_{\partial R^*} v \, d\vec{s} \\ &= \int_R \left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y} \right) dx \, dy \\ &= \frac{1}{r} \int_R dx \, dy = \frac{1}{r} |R|.\end{aligned}$$

Hence, with r known, the area of the enclosed domain R has been found.

Finding Areas by Green's Theorem

3.4.21. Example. Let $E = \{(x, y) \in \mathbb{R}^2 : (x/a)^2 + (y/b)^2 = 1\}$ be the ellipse centered at the origin with half-axes of length $a > 0$ and $b > 0$. We can find the area of the ellipse from

$$|E| = \frac{1}{2} \int_{\partial E} \begin{pmatrix} -y \\ x \end{pmatrix} d\vec{s}.$$

First, ∂E is parametrized by

$$\gamma(t) = \begin{pmatrix} a \cos t \\ b \sin t \end{pmatrix}, \quad t \in [0, 2\pi].$$

We obtain

$$\begin{aligned} |E| &= \frac{1}{2} \int_0^{2\pi} \left\langle \begin{pmatrix} -b \sin t \\ a \cos t \end{pmatrix}, \begin{pmatrix} -a \sin t \\ b \cos t \end{pmatrix} \right\rangle dt \\ &= \frac{1}{2} \int_0^{2\pi} ab(\cos^2 t + \sin^2 t) dt = \pi ab. \end{aligned}$$

Physical Interpretation of Green's Theorem

Green's theorem can be interpreted both in terms of the rotation and the divergence of a vector field.

Let $\Omega \subset \mathbb{R}^2$ be open, $F: \Omega \rightarrow \mathbb{R}^2$ and $R \subset \Omega$ a simple region with boundary ∂R . Then the circulation (see (3.2.1) and the following discussion) around R is given by

$$\int_{\partial R^*} F \, d\vec{s}$$

where ∂R^* is oriented positively. Furthermore, the rotation of F is given by

$$\text{rot } F = \frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2}$$

(see (3.2.6)).

Physical Interpretation of Green's Theorem

Green's Theorem then states that

$$\begin{aligned}\text{circulation along } \partial R &= \int_{\partial R^*} \mathbf{F} \cdot d\vec{s} \\ &= \int_R \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) dx \\ &= \int_R \operatorname{rot} \mathbf{F} dx \\ &= \text{integral of circulation density over } R\end{aligned}$$

Physical Interpretation of Green's Theorem

In a similar manner, we can show that Green's Theorem relates the total flux through ∂R with the divergence of F . Define

$$\tilde{F}(x) := \begin{pmatrix} -F_2(x) \\ F_1(x) \end{pmatrix}.$$

Then $\langle F, N \rangle = \langle \tilde{F}, T \rangle$ for a tangent vector of ∂R (positively oriented) and the outward-pointing normal vector. This yields

$$\begin{aligned} \text{flux through } \partial R &= \int_{\partial R^*} \langle F, N \rangle \, ds = \int_{\partial R^*} \tilde{F} \, d\vec{s} \\ &= \int_R \left(\frac{\partial \tilde{F}_2}{\partial x_1} - \frac{\partial \tilde{F}_1}{\partial x_2} \right) \, dx \\ &= \int_R \operatorname{div} F \, dx \\ &= \text{integral of flux density over } R \end{aligned}$$

Vector Fields and Line Integrals

Circulation and Flux

The Riemann Integral and Measurable Sets

Integration in Practice

Surfaces and Surface Integrals

The Theorems of Gauß and Stokes

Parametrized Surfaces

We will now introduce surfaces in \mathbb{R}^n . While it is possible to discuss surfaces without references to a parametrization and then consider different parametrizations and reparametrizations, this requires more mathematical background than we have developed at this point. (The implicit function theorem plays a major role.) Therefore, we restrict ourselves to parametrized surfaces, i.e., surfaces that are accompanied by a fixed parametrization.

3.5.1. Definition. A *smooth parametrized m-surface in \mathbb{R}^n* is a subset $\mathcal{S} \subset \mathbb{R}^n$ together with a locally bijective, continuously differentiable map (parametrization)

$$\varphi: \Omega \rightarrow \mathcal{S}, \quad \Omega \subset \mathbb{R}^m,$$

such that

$$\text{rank } D\varphi|_x = m$$

for almost every $x \in \Omega$. If $m = n - 1$, \mathcal{S} is said to be a *parametrized hypersurface*.

Parametrized Surfaces

3.5.2. Example. The unit sphere in \mathbb{R}^3 ,

$$S^2 := \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 = 1\}$$

is a two-surface with parametrization

$$\varphi: [0, 2\pi] \times [0, \pi] \rightarrow S^2, \quad \varphi(\phi, \theta) = \begin{pmatrix} \cos \phi \sin \theta \\ \sin \phi \sin \theta \\ \cos \theta \end{pmatrix}.$$

We note that

$$\begin{aligned} \text{rank } D\varphi|_{(\phi, \theta)} &= \text{rank} \begin{pmatrix} -\sin \phi \sin \theta & \cos \phi \cos \theta \\ \cos \phi \sin \theta & \sin \phi \cos \theta \\ 0 & -\sin \theta \end{pmatrix} \\ &= \text{rank} \begin{pmatrix} -\sin \phi & \cos \phi \cos \theta \\ \cos \phi & \sin \phi \cos \theta \\ 0 & 1 \end{pmatrix} = 2 \end{aligned}$$

for $\sin \theta \neq 0$. Hence $\text{rank } D\varphi = 2$ almost everywhere on $[0, 2\pi] \times [0, \pi]$.

Parametrized Surfaces

3.5.3. Example. Consider the graph of a potential function $f: \Omega \rightarrow \mathbb{R}$, $\Omega \subset \mathbb{R}^n$,

$$\Gamma(f) = \{(x_1, \dots, x_n, x_{n+1}) \in \mathbb{R}^{n+1} : x = (x_1, \dots, x_n) \in \Omega, x_{n+1} = f(x)\}.$$

This is a hypersurface in \mathbb{R}^{n+1} with parametrization

$$\varphi: \Omega \rightarrow \Gamma(f), \quad \varphi(x) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ f(x_1, \dots, x_n) \end{pmatrix}.$$

The rank of the Jacobian is

$$\text{rank } D\varphi|_x = \text{rank} \begin{pmatrix} \mathbb{1} \\ Df|_x \end{pmatrix} = n,$$

written in block matrix form, where $\mathbb{1} \in \text{Mat}(n \times n; \mathbb{R})$ is the $n \times n$ unit matrix.

Tangent Spaces of Surfaces

We now want to define the tangent space of a parametrized m -surface $\mathcal{S} \subset \mathbb{R}^n$. The parametrization $\varphi: \Omega \rightarrow \mathcal{S}$ satisfies

$$\varphi(x_0 + h) = \varphi(x_0) + D\varphi|_{x_0} h + o(h) \quad \text{as } h \rightarrow 0.$$

Hence we consider the map

$$h \rightarrow D\varphi|_x h$$

as a linear approximation to φ near x_0 . The range of this map is given by

$$\{x \in \mathbb{R}^n : x = D\varphi|_{x_0} h, h \in \mathbb{R}^m\}.$$

and is equal to the span of the column vectors of $D\varphi$. The elements of the range give good approximations to \mathcal{S} at points near x_0 .

Tangent Spaces of Surfaces

Hence, it is natural to make the following definition.

3.5.4. Definition. Let $\mathcal{S} \subset \mathbb{R}^n$ be a parametrized m -surface with parametrization $\varphi: \Omega \rightarrow \mathcal{S}$. Then

$$t_k(p) = \frac{\partial}{\partial x_k} \begin{pmatrix} \varphi_1(x) \\ \vdots \\ \varphi_n(x) \end{pmatrix} \Bigg|_{x=\varphi^{-1}(p)}, \quad k = 1, \dots, m.$$

is called the ***kth tangent vector of \mathcal{S} at $p \in \mathcal{S}$*** and

$$T_p \mathcal{S} := \text{ran } D\varphi|_x = \text{span}\{t_1(p), \dots, t_m(p)\}$$

is called the ***tangent space*** to \mathcal{S} at p . The vector field

$$t_k: \mathcal{S} \rightarrow \mathbb{R}^n, \quad p \mapsto t_k(p)$$

is called the ***kth tangent vector field*** on \mathcal{S} .

Tangent Vectors to the Unit Sphere

3.5.5. Example. For the unit sphere $S^2 \subset \mathbb{R}^3$ parametrized with

$$\varphi(\phi, \theta) = \begin{pmatrix} \cos \phi \sin \theta \\ \sin \phi \sin \theta \\ \cos \theta \end{pmatrix}.$$

we have the tangent vectors at $p \in S^2$ given by

$$t_\phi(p) = \frac{\partial}{\partial \phi} \begin{pmatrix} \cos \phi \sin \theta \\ \sin \phi \sin \theta \\ \cos \theta \end{pmatrix} = \begin{pmatrix} -\sin \phi \sin \theta \\ \cos \phi \sin \theta \\ 0 \end{pmatrix}$$

$$t_\theta(p) = \frac{\partial}{\partial \theta} \begin{pmatrix} \cos \phi \sin \theta \\ \sin \phi \sin \theta \\ \cos \theta \end{pmatrix} = \begin{pmatrix} \cos \phi \cos \theta \\ \sin \phi \cos \theta \\ -\sin \theta \end{pmatrix}$$

taken at $(\phi, \theta) = \varphi^{-1}(p)$.

Tangent Vectors to the Unit Sphere

At $p = (1/\sqrt{2}, 0, 1/\sqrt{2}) = \varphi(0, \pi/4)$ the tangent vectors are

$$t_\phi(p) = \begin{pmatrix} 0 \\ \sqrt{2}/2 \\ 0 \end{pmatrix}, \quad t_\theta(p) = \begin{pmatrix} \sqrt{2}/2 \\ 0 \\ -\sqrt{2}/2 \end{pmatrix}$$

and the tangent space is

$$\begin{aligned} T_p S^2 &= \text{span}\{t_\phi(p), t_\theta(p)\} \\ &= \left\{ x \in \mathbb{R}^3 : x = \alpha \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \alpha, \beta \in \mathbb{R} \right\}. \end{aligned}$$

3.5.6. Example. The graph $\Gamma(f)$ of a potential function is a surface (see Example 3.5.3). The tangent space to $\Gamma(f)$ coincides with the tangent plane of Definition 2.4.6. The verification is left to you.

The Normal Vector to Hypersurfaces

The tangent space of an m -surface in \mathbb{R}^n is an m -dimensional subspace of \mathbb{R}^n . If \mathcal{S} is a hypersurface, i.e., an $(n - 1)$ -surface in \mathbb{R}^n , then $(T_p \mathcal{S})^\perp$ is a 1-dimensional subspace of \mathbb{R}^n and there exists a unit basis vector of this space. This basis vector is uniquely defined up its sign.

3.5.7. Definition. Let $\mathcal{S} \subset \mathbb{R}^n$ be a hypersurface. Then a unit vector that is orthogonal to all tangent vectors to \mathcal{S} at p is called a **unit normal vector to \mathcal{S} at p** and denoted by $N(p)$. The vector field

$$N: \mathcal{S} \rightarrow \mathbb{R}^n, \quad p \mapsto N(p)$$

is called the **normal vector field** on \mathcal{S} .

Orientation of Hypersurfaces

3.5.8. Example. Returning to Example 3.5.5, the unit normal vector at $p = (1/\sqrt{2}, 0, 1/\sqrt{2})$ is orthogonal to both $t_\phi(p)$ and $t_\theta(p)$ and given by

$$N(p) = \pm \begin{pmatrix} \sqrt{2}/2 \\ 0 \\ \sqrt{2}/2 \end{pmatrix} = \pm p$$

where we are free to choose a sign arbitrarily.

Since the unit normal vector is uniquely determined up to its sign, there are two possible choices for a normal vector at each $p \in \mathcal{S}$. Usually, one chooses the direction of the normal vector at a single point of \mathcal{S} and attempts to choose the normal vector at all other points of \mathcal{S} in such a way that the normal vector field is continuous on \mathcal{S} .

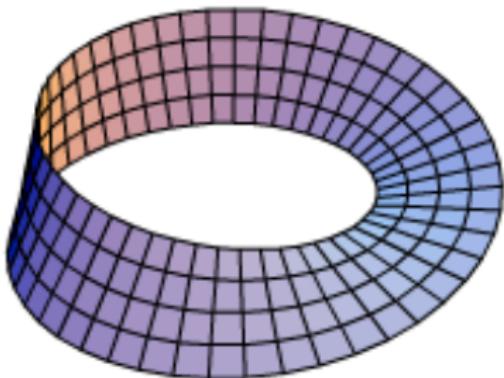
Orientation of Hypersurfaces

3.5.9. Definition.

- (i) A hypersurface $\mathcal{S} \subset \mathbb{R}^n$ such that it admits a continuous normal vector field is said to be **orientable**.
- (ii) A choice of direction for the normal vector field is called an **orientation of \mathcal{S}** .
- (iii) A hypersurface that is the boundary of a measurable set $\Omega \subset \mathbb{R}^n$ with non-zero measure is said to be a **closed surface**.
- (iv) A closed hypersurface is said to have **positive orientation** if the normal vector field is chosen so that the normal vectors point **outwards** from Ω .

Orientation of Surfaces

3.5.10. Example. The classic example of a 2-surface that is not orientable is the **Möbius strip** in \mathbb{R}^3 :



A parametrization is given by

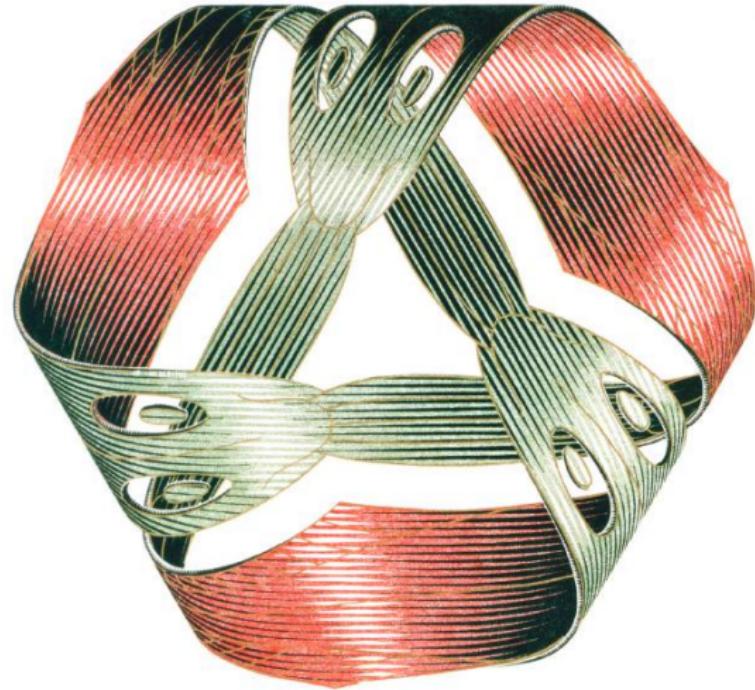
$$\varphi: [-w, w] \times [0, 2\pi) \rightarrow \mathbb{R}^3,$$

$$\varphi(s, t) = \begin{pmatrix} (R + s \cos(t/2)) \cos t \\ (R + s \cos(t/2)) \sin t \\ s \sin(t/2) \end{pmatrix}$$

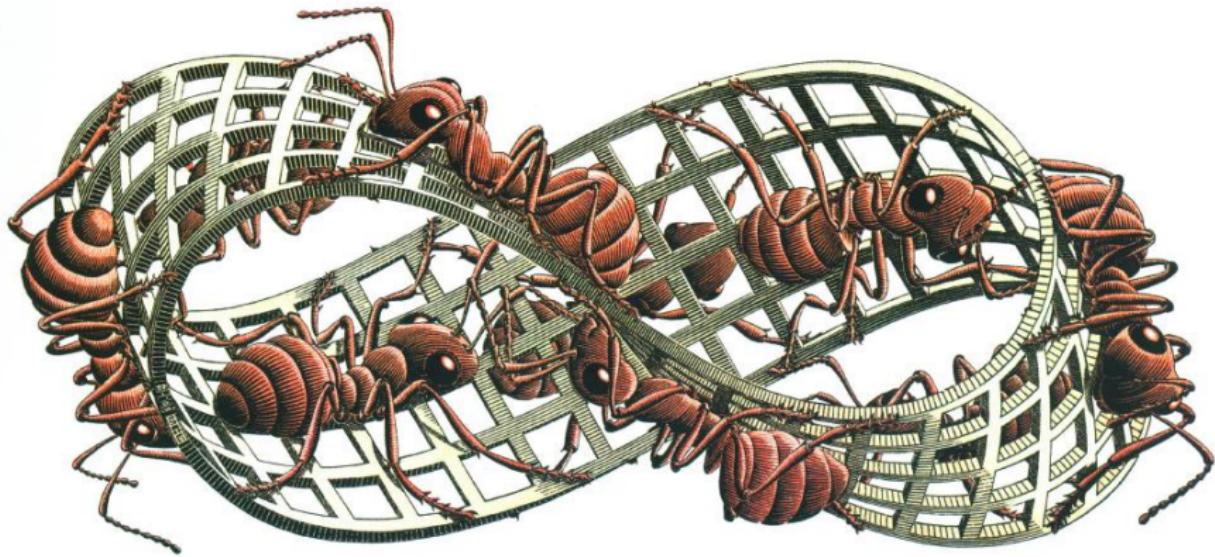
The above parametrization gives a Möbius strip lying in the x_1 - x_2 plane of width $2w > 0$.

Suppose a normal vector is chosen at some point p . Moving the normal vector around the strip back to its initial position p , it then points in the other direction. Hence, the normal vector field is not continuous.

Möbius 1 by M. C. Escher



Möbius 2 by M. C. Escher



Infinitesimal Surface Elements of Hypersurfaces

Our goal now is to define the area of surfaces. Consider a parametrized 2-surface \mathcal{S} in \mathbb{R}^3 . At any point $p \in \mathcal{S}$ there exist two tangent vectors $t_1(p)$ and $t_2(p)$. Suppose $\varphi = \varphi(x_1, x_2)$ is a parametrization of \mathcal{S} and $p = \varphi(x)$. We would like to define an “infinitesimal surface element”

$$dA = \text{area of the parallelogram spanned by } t_1 \text{ and } t_2 \text{ at } \varphi(x_1, x_2) \cdot dx_1 dx_2$$

We do not have a simple expression for the area of this parallelogram, but fortunately we have a unit normal vector N at $p = \varphi(x)$ (because \mathcal{S} is a hypersurface). We can hence replace the area of the parallelogram spanned by t_1 and t_2 by the volume of the parallelepiped spanned by t_1 , t_2 and N .

Volume (Area) of Hypersurfaces

We define the **scalar surface element of a hypersurface in \mathbb{R}^3** by

$$dA = |\det(t_1, t_2, N) \circ \varphi| dx_1 dx_2.$$

Of course, we can generalize this to hypersurfaces in \mathbb{R}^n , setting

$$dA = |\det(t_1, t_2, \dots, t_{n-1}, N) \circ \varphi| dx_1 dx_2 \dots dx_{n-1}.$$

3.5.11. Definition. Let $\mathcal{S} \subset \mathbb{R}^n$ be a hypersurface with parametrization $\varphi \in C^1(\Omega, \mathbb{R}^n)$, $\Omega \subset \mathbb{R}^{n-1}$. Let $t_j = D\varphi e_j$, $j = 1, \dots, n-1$, be the tangent vector fields on \mathcal{S} . Let N be a chosen normal vector field on \mathcal{S} (so that \mathcal{S} is oriented). Then the **volume** or **area** of \mathcal{S} is defined as

$$|\mathcal{S}| := \int_{\Omega} |\det(t_1, \dots, t_{n-1}, N) \circ \varphi(x)| dx_1 dx_2 \dots dx_{n-1}.$$

Area of the Unit Sphere in \mathbb{R}^3

3.5.12. Example. In Example 3.5.5 we have seen that S^2 with parametrization

$$\varphi(\phi, \theta) = \begin{pmatrix} \cos \phi \sin \theta \\ \sin \phi \sin \theta \\ \cos \theta \end{pmatrix}.$$

has tangent vectors

$$t_\phi \circ \varphi(\phi, \theta) = \begin{pmatrix} -\sin \phi \sin \theta \\ \cos \phi \sin \theta \\ 0 \end{pmatrix}, \quad t_\theta \circ \varphi(\phi, \theta) = \begin{pmatrix} \cos \phi \cos \theta \\ \sin \phi \cos \theta \\ -\sin \theta \end{pmatrix}$$

To calculate the normal vector, we can simply take

$$t_\phi \times t_\theta = \begin{pmatrix} -\cos \phi \sin^2 \theta \\ -\sin \phi \sin^2 \theta \\ -\cos \theta \sin \theta \end{pmatrix}$$

Area of the Unit Sphere in \mathbb{R}^3

Taking account of $|t_\phi \times t_\theta| = \sin \theta$, we have

$$N \circ \varphi(\phi, \theta) = - \begin{pmatrix} \cos \phi \sin \theta \\ \sin \phi \sin \theta \\ \cos \theta \end{pmatrix}.$$

Then the area of the unit sphere is given by

$$\begin{aligned} |S^2| &= \int_0^{2\pi} \int_0^\pi \left| \det \begin{pmatrix} -\sin \phi \sin \theta & \cos \phi \cos \theta & -\cos \phi \sin \theta \\ \cos \phi \sin \theta & \sin \phi \cos \theta & -\sin \phi \sin \theta \\ 0 & -\sin \theta & -\cos \theta \end{pmatrix} \right| d\theta d\phi \\ &= \int_0^{2\pi} \int_0^\pi \sin \theta \left| \det \begin{pmatrix} -\sin \phi & \cos \phi \cos \theta & \cos \phi \sin \theta \\ \cos \phi & \sin \phi \cos \theta & \sin \phi \sin \theta \\ 0 & -\sin \theta & \cos \theta \end{pmatrix} \right| d\theta d\phi \\ &= 2\pi \int_0^\pi \sin \theta d\theta = 4\pi \end{aligned}$$

Infinitesimal Surface Elements of Arbitrary Surfaces

We would like to generalize the concept of area and infinitesimal surface elements from hypersurfaces to arbitrary surfaces in \mathbb{R}^n . From the beginning, the introduction of the normal vector to calculate the surface area by means of the volume was undertaken by necessity rather than through any other considerations.

We note that, in block matrix notation.

$$\begin{aligned}\det(t_1, \dots, t_{n-1}, N)^2 &= \det((t_1, \dots, t_{n-1}, N)^T) \cdot \det(t_1, \dots, t_{n-1}, N) \\ &= \det((t_1, \dots, t_{n-1}, N)^T \cdot (t_1, \dots, t_{n-1}, N)) \\ &= \det\left(\begin{pmatrix} t_1^T \\ \vdots \\ t_{n-1}^T \\ N^T \end{pmatrix} \cdot (t_1, \dots, t_{n-1}, N)\right)\end{aligned}$$

Infinitesimal Surface Elements of Arbitrary Surfaces

Performing the row-by-column matrix multiplication, we see that

$$\begin{aligned} \det(t_1, \dots, t_{n-1}, N)^2 &= \det \begin{pmatrix} \langle t_1, t_1 \rangle & \cdots & \langle t_1, t_{n-1} \rangle & \langle t_1, N \rangle \\ \vdots & \ddots & \vdots & \vdots \\ \langle t_{n-1}, t_1 \rangle & \cdots & \langle t_{n-1}, t_{n-1} \rangle & \langle t_{n-1}, N \rangle \\ \langle N, t_1 \rangle & \cdots & \langle N, t_{n-1} \rangle & \langle N, N \rangle \end{pmatrix} \\ &= \det \begin{pmatrix} \langle t_1, t_1 \rangle & \cdots & \langle t_1, t_{n-1} \rangle & 0 \\ \vdots & \ddots & \vdots & \vdots \\ \langle t_{n-1}, t_1 \rangle & \cdots & \langle t_{n-1}, t_{n-1} \rangle & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} \\ &= \det \begin{pmatrix} \langle t_1, t_1 \rangle & \cdots & \langle t_1, t_{n-1} \rangle \\ \vdots & \ddots & \vdots \\ \langle t_{n-1}, t_1 \rangle & \cdots & \langle t_{n-1}, t_{n-1} \rangle \end{pmatrix} \end{aligned}$$

where we have used that the normal vector is orthogonal to all tangent vectors and has unit length.

The Metric Tensor

3.5.13. Definition. Let $\mathcal{S} \subset \mathbb{R}^n$ be an m -surface with parametrization φ and tangent vector fields t_1, \dots, t_m . Then $G \in \text{Mat}(m \times m; \mathbb{R})$ given by

$$G := \begin{pmatrix} \langle t_1, t_1 \rangle & \cdots & \langle t_1, t_m \rangle \\ \vdots & \ddots & \vdots \\ \langle t_m, t_1 \rangle & \cdots & \langle t_m, t_m \rangle \end{pmatrix}$$

is said to be the **metric tensor** on \mathcal{S} with respect to φ . The coefficients

$$g_{ij} := \langle t_i, t_j \rangle, \quad i, j = 1, \dots, m,$$

are called the **metric coefficients** of G . We often write

$$g(x) = \det G(\varphi(x))$$

for short.

The Metric Tensor

3.5.14. Remarks.

- (i) We have proved that if \mathcal{S} is a hypersurface in \mathbb{R}^n , then

$$|\det(t_1, \dots, t_{n-1}, N)| = \sqrt{\det G}.$$

This will allow us to extend the definition of area/volume to general surfaces.

- (ii) In the case $n = 3$, $m = 2$, we have

$$\begin{aligned} g &= \det \begin{pmatrix} \|t_1\|^2 & \langle t_1, t_2 \rangle \\ \langle t_2, t_1 \rangle & \|t_2\|^2 \end{pmatrix} = \|t_1\|^2 \|t_2\|^2 - \langle t_1, t_2 \rangle^2 \\ &= \|t_1\|^2 \|t_2\|^2 (1 - \cos^2 \angle(t_1, t_2)) \\ &= \|t_1\|^2 \|t_2\|^2 \sin^2 \angle(t_1, t_2) \\ &= \|t_1 \times t_2\|^2 \end{aligned}$$

so that

$$dA = \|t_1 \times t_2\| \circ \varphi(x) dx_1 dx_2.$$

Scalar Surface Integrals

3.5.15. Definition. Let \mathcal{S} be a parametrized m -surface with parametrization $\varphi: \Omega \rightarrow \mathcal{S}$, $\Omega \subset \mathbb{R}^m$. Then

$$|\mathcal{S}| := \int_{\Omega} \sqrt{g(x)} dx$$

defines the **volume** or **area** of \mathcal{S} .

Let $f: \mathcal{S} \rightarrow \mathbb{R}$ be a potential function. Then the **(scalar) surface integral of f over \mathcal{S}** is defined as

$$\int_{\mathcal{S}} f dA := \int_{\Omega} f \circ \varphi(x) \sqrt{g(x)} dx$$

3.5.16. Remark. As usual,

$$dA := \sqrt{g(x)} dx$$

is called the **scalar surface element** of \mathcal{S} .

Electrostatic Potential of a Surface Charge

3.5.17. Example. The electrostatic potential $V(p)$ at a point $p \in \mathbb{R}^3$ induced by a charged surface \mathcal{S} is given by

$$V(p) = \frac{1}{4\pi\varepsilon_0} \int_{\mathcal{S}} \frac{\varrho(\cdot)}{\text{dist}(p, \cdot)} dA$$

where ϱ is the charge density of the surface. Let

$$\mathcal{S} = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = z^2, 0 \leq z \leq 1\}$$

and assume that ϱ is constant on \mathcal{S} . We calculate the potential at the point $p = (0, 0, 1)$. Introducing polar coordinates, we have

$$\begin{aligned} \mathcal{S} = \{(x, y, z) \in \mathbb{R}^3 : x &= r \cos \theta, y = r \sin \theta, z = r, \\ &0 \leq \theta \leq 2\pi, 0 \leq r \leq 1\}. \end{aligned}$$

Electrostatic Potential of a Surface Charge

We can read off that a parametrization of \mathcal{S} is given by

$$\varphi: [0, 2\pi] \times [0, 1] \rightarrow \mathcal{S}, \quad \varphi(\theta, r) = \begin{pmatrix} r \cos \theta \\ r \sin \theta \\ r \end{pmatrix}.$$

The tangent vectors are

$$t_\theta \circ \varphi(\theta, r) = \begin{pmatrix} -r \sin \theta \\ r \cos \theta \\ 0 \end{pmatrix}, \quad t_r \circ \varphi(\theta, r) = \begin{pmatrix} \cos \theta \\ \sin \theta \\ 1 \end{pmatrix}.$$

Hence,

$$\begin{aligned} g(\theta, r) &= \det \begin{pmatrix} \langle t_\theta, t_\theta \rangle & \langle t_\theta, t_r \rangle \\ \langle t_r, t_\theta \rangle & \langle t_r, t_r \rangle \end{pmatrix} \Big|_{\varphi(\theta, r)} \\ &= \det \begin{pmatrix} r^2 & 0 \\ 0 & 2 \end{pmatrix} \\ &= 2r^2. \end{aligned}$$

Electrostatic Potential of a Surface Charge

It follows that the volume element is given by

$$dA = \sqrt{2r} dr d\theta.$$

We then have

$$\begin{aligned} V(p) &= \frac{\varrho}{4\pi\varepsilon_0} \int_{\mathcal{S}} \frac{dA}{\|p - (\cdot)\|} \\ &= \frac{\varrho}{4\pi\varepsilon_0} \int_0^{2\pi} \int_0^1 \frac{\sqrt{2r} dr d\theta}{\sqrt{r^2 \cos^2 \theta + r^2 \sin^2 \theta + (1-r)^2}} \\ &= \frac{\varrho}{\sqrt{2}\varepsilon_0} \int_0^1 \frac{r dr}{\sqrt{2r^2 - 2r + 1}} \\ &= \frac{\varrho}{4\varepsilon_0} \ln(3 + 2\sqrt{2}) \end{aligned}$$

Flux Through Hypersurfaces

If \mathcal{S} is a hypersurface, we can also define the flux of a vector field through \mathcal{S} :

3.5.18. Definition. Let $F: \mathbb{R}^{n+1} \rightarrow \mathbb{R}^{n+1}$ be a vector field defined in a neighborhood of a hypersurface \mathcal{S} with parametrization $\varphi: \Omega \rightarrow \mathbb{R}^{n+1}$, $\Omega \subset \mathbb{R}^n$. Then we define the **flux of F through \mathcal{S}** by

$$\begin{aligned}\int_{\mathcal{S}} F d\vec{A} &:= \int_{\mathcal{S}} \langle F, N \rangle dA \\ &= \int_{\Omega} \langle F \circ \varphi(x), N \circ \varphi(x) \rangle \sqrt{g(x)} dx_1 \dots dx_n\end{aligned}$$

Flux Through Hypersurfaces

3.5.19. Remark. We also sometimes write

$$\int_{\mathcal{S}} \langle F, d\vec{A} \rangle$$

for the flux integral. The term

$$d\vec{A} := N(\varphi(x)) \cdot \sqrt{g(x)} \, dx$$

is called the **vectorial surface element** of a hypersurface \mathcal{S} .

For a hypersurface in \mathbb{R}^3 we have

$$N = \frac{\mathbf{t}_1 \times \mathbf{t}_2}{\|\mathbf{t}_1 \times \mathbf{t}_2\|},$$

so that

$$d\vec{A} = \mathbf{t}_1(\varphi(x)) \times \mathbf{t}_2(\varphi(x)) \, dx_1 \, dx_2.$$

Flux Through Hypersurfaces

3.5.20. Example. A point charge Q at the origin induces a field

$$E(p) = \frac{1}{4\pi\varepsilon_0} \frac{Q}{\|p\|^3} p$$

at any point $p \in \mathbb{R}^3 \setminus \{0\}$. The flux of this field through the unit sphere S^2 is given by

$$\int_{S^2} \langle E, d\vec{A} \rangle.$$

As in Example 3.5.12, we can parametrize S^2 by

$$\varphi(\phi, \theta) = \begin{pmatrix} \cos \phi \sin \theta \\ \sin \phi \sin \theta \\ \cos \theta \end{pmatrix}, \quad t_\theta \times t_\phi = \begin{pmatrix} \cos \phi \sin^2 \theta \\ \sin \phi \sin^2 \theta \\ \cos \theta \sin \theta \end{pmatrix}$$

where $0 < \phi < 2\pi$ and $0 < \theta < \pi$. Here we have chosen the outward-pointing (positively oriented) normal vector.

Flux Through Hypersurfaces

It follows that

$$\begin{aligned}\int_{S^2} \langle E, d\vec{A} \rangle &= \frac{Q}{4\pi\varepsilon_0} \int_0^{2\pi} \int_0^\pi \underbrace{\frac{1}{\|\varphi(\phi, \theta)\|^3}}_{=1} \langle \varphi(\phi, \theta), t_\phi \times t_\theta(\phi, \theta) \rangle d\theta d\phi \\ &= \frac{Q}{4\pi\varepsilon_0} \int_0^{2\pi} \int_0^\pi \left\langle \begin{pmatrix} \cos \phi \sin \theta \\ \sin \phi \sin \theta \\ \cos \theta \end{pmatrix}, \begin{pmatrix} \cos \phi \sin^2 \theta \\ \sin \phi \sin^2 \theta \\ \cos \theta \sin \theta \end{pmatrix} \right\rangle d\theta d\phi \\ &= \frac{Q}{4\pi\varepsilon_0} \int_0^{2\pi} \int_0^\pi \sin \theta d\theta d\phi \\ &= \frac{Q}{\varepsilon_0}.\end{aligned}$$

The fact that this result is actually true for any closed surface (not just S^2) that contains the charge at the origin is known as **Gauß's law** in electrostatics.

Vector Fields and Line Integrals

Circulation and Flux

The Riemann Integral and Measurable Sets

Integration in Practice

Surfaces and Surface Integrals

The Theorems of Gauß and Stokes

Ordinate Regions in \mathbb{R}^n

The goal of this final section is to develop generalizations of Green's Theorem 3.4.18 to higher dimensions. In \mathbb{R}^2 , we have seen that Green's Theorem provides macroscopic equations for both the divergence and the rotation, summarized as

flux through $\partial R = \text{integral of flux density over } R,$

circulation along $\partial R = \text{integral of circulation density over } R.$

for a suitable region $R \subset \mathbb{R}^2$. It turns out that in \mathbb{R}^3 these physical statements lead to two separate theorems.

Admissible Regions

We have proven Green's theorem for simple regions in \mathbb{R}^2 , but we have not precisely specified for which regions it is valid precisely. Let us now do this:

3.6.1. Definition.

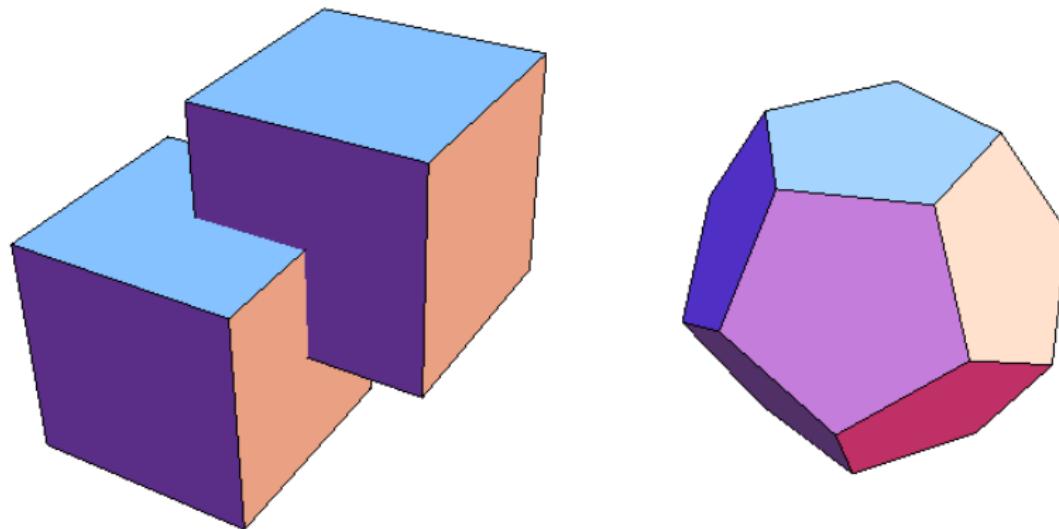
- (i) A subset $R \subset \mathbb{R}^n$ is called a **region** if it is open and (pathwise) connected.
- (ii) A region $R \subset \mathbb{R}^n$ is said to be **admissible** if it is bounded and its boundary is the union of a finite number of parametrized hypersurfaces whose normal vectors point outwards from R .

3.6.2. Theorem.

Green's Theorem is valid for any admissible region in \mathbb{R}^2 .

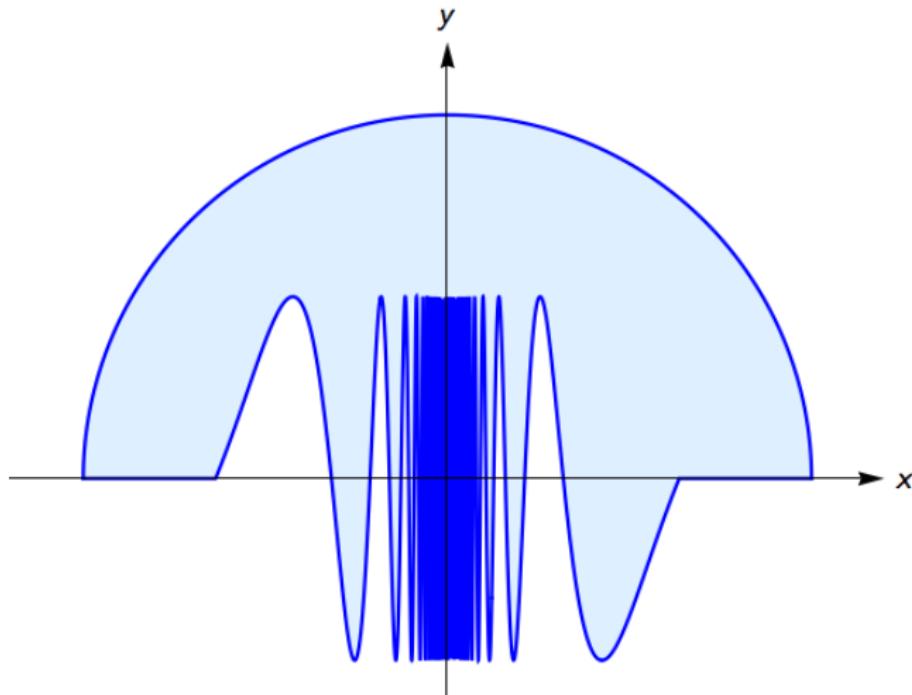
Admissible Regions

Admissible regions may have edges and corners:



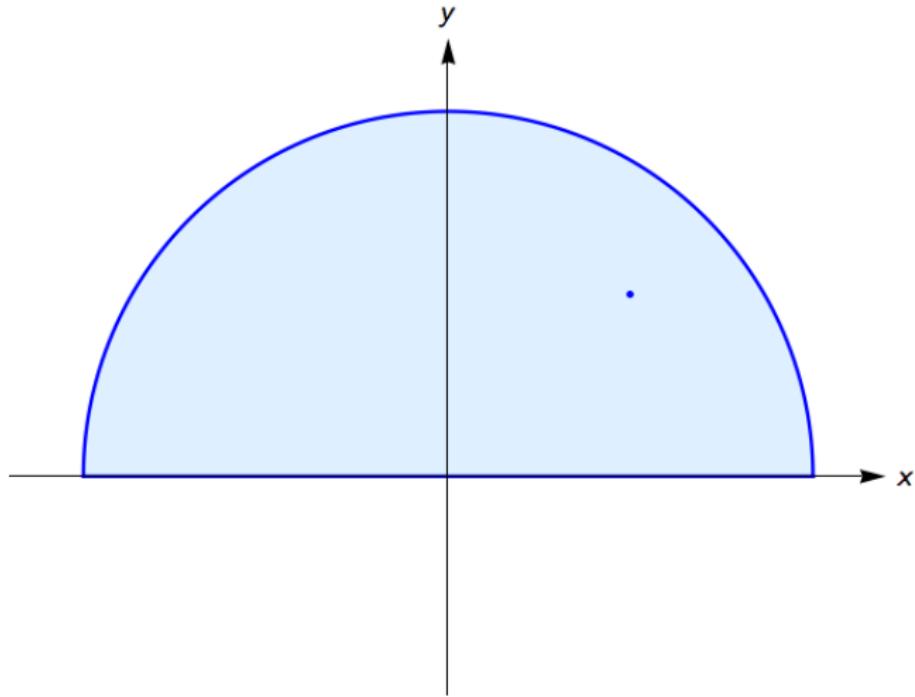
Admissible Regions

The boundary may not behave “too wildly”. This region is not admissible:



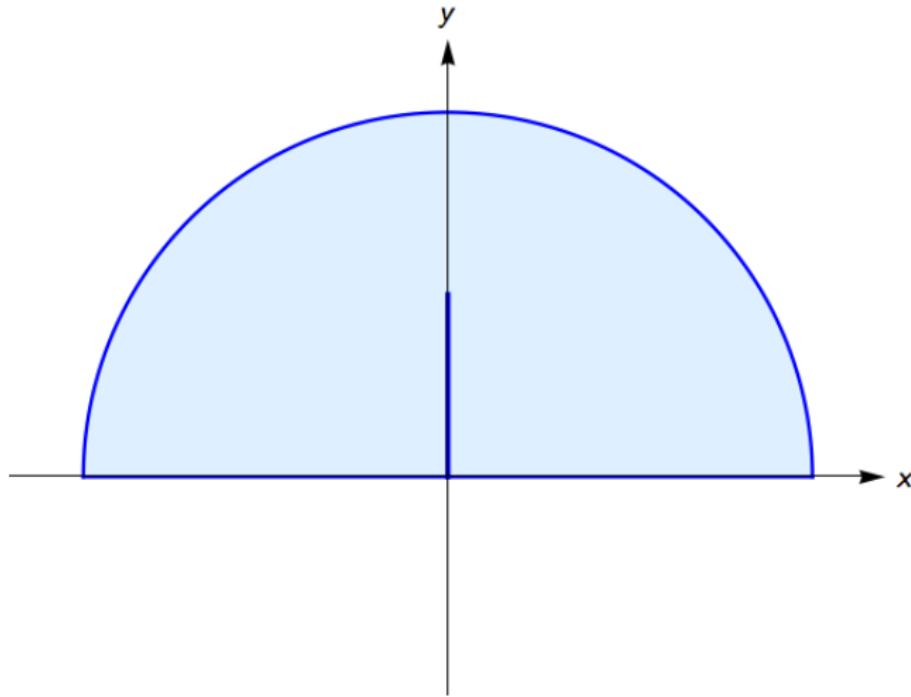
Admissible Regions

Removing an interior point means the boundary is not everywhere a hypersurface. This region is not admissible:



Admissible Regions

Removing part of the center line means that it is impossible to find outward-pointing normal vectors. This region is not admissible:



Admissible Hypersurfaces in \mathbb{R}^3

3.6.3. Definition. A hypersurface $\mathcal{S} \subset \mathbb{R}^3$ with parametrization $\varphi: R \rightarrow \mathcal{S}$ is said to be **admissible** if

- (i) the interior $\text{int } R$ is an admissible region in \mathbb{R}^2 with an oriented boundary curve ∂R^* and
- (ii) R is closed, i.e., $R = \overline{R}$.

In particular, for the boundary of the region R consists of a finite number of parametrized hypersurfaces in \mathbb{R}^2 , i.e., smooth curves. Let us write

$$\partial R = \mathcal{C}_1 \cup \mathcal{C}_2 \cup \cdots \cup \mathcal{C}_k.$$

This boundary of R is of course mapped by φ into a set of points of \mathcal{S} . We would like to formulate a criterion for determining whether $\varphi(\partial R)$ (or part of this set) constitutes an “actual boundary” of the surface \mathcal{S} or not.

Since ∂R^* is oriented, we define the **chain of curves** $\mathcal{C}_1^* \cup \mathcal{C}_2^* \cup \cdots \cup \mathcal{C}_k^*$ where the individual curves are traversed in the “correct” orientation determined by ∂R^* .

Closed Hypersurfaces in \mathbb{R}^3 and those with Boundary

3.6.4. Definition. Let $\mathcal{S} \subset \mathbb{R}^3$ be an admissible hypersurface with parametrization $\varphi: R \rightarrow \mathcal{S}$. Let $\partial R^* = \mathcal{C}_1^* \cup \mathcal{C}_2^* \cup \dots \cup \mathcal{C}_k^*$, where each \mathcal{C}_k^* is an oriented smooth curve in \mathbb{R}^2 and all \mathcal{C}_i are pairwise disjoint.

- (i) We say that φ **annihilates** a chain of curves $\mathcal{C}_{i_1} \cup \dots \cup \mathcal{C}_{i_j}$ if

$$\int_{\varphi(\mathcal{C}_{i_1} \cup \dots \cup \mathcal{C}_{i_j})} 1 \, ds = 0.$$

- (ii) If φ annihilates ∂R , \mathcal{S} is said to be a **closed surface**.
- (iii) Denote by $\mathcal{C}' \subset \partial R$ the largest chain of curves that is annihilated by φ . If $\mathcal{C}' \neq \partial R$ we say that \mathcal{S} is a **surface with boundary** and define

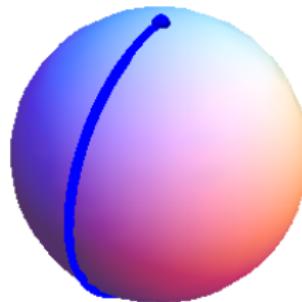
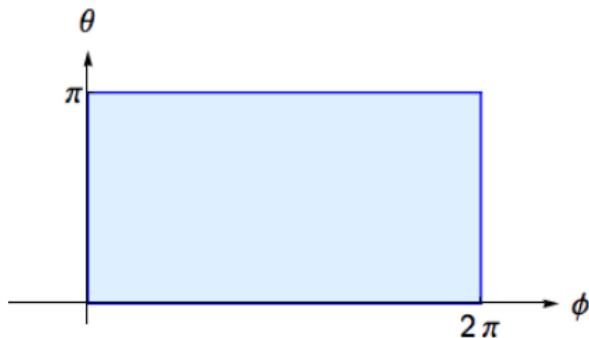
$$\partial \mathcal{S} := \varphi(\partial R \setminus \mathcal{C}'). \tag{3.6.1}$$

Admissible Hypersurfaces in \mathbb{R}^3

3.6.5. Example. The unit sphere $S^2 \subset \mathbb{R}^3$ is parametrized by

$$\varphi: [0, 2\pi] \times [0, \pi] \rightarrow S^2, \quad \varphi(\phi, \theta) = \begin{pmatrix} \cos \phi \sin \theta \\ \sin \phi \sin \theta \\ \cos \theta \end{pmatrix}.$$

The interior of $[0, 2\pi] \times [0, \pi]$ is clearly an admissible region in \mathbb{R}^2 , as it is closed and its boundary consists of four lines, which are hypersurfaces in \mathbb{R}^2 and the normal vectors can be taken to point outward. It is easily seen that φ annihilates the boundary of the rectangle, so S^2 is a closed, admissible surface.



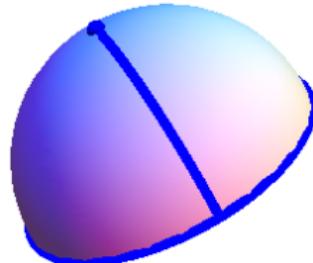
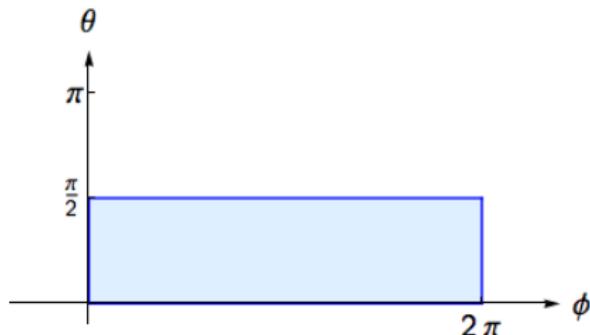
Admissible Hypersurfaces in \mathbb{R}^3

3.6.6. Example. The map

$$\varphi: [0, 2\pi] \times [0, \pi/2] \rightarrow \mathcal{S}, \quad \varphi(\phi, \theta) = \begin{pmatrix} \cos \phi \sin \theta \\ \sin \phi \sin \theta \\ \cos \theta \end{pmatrix}.$$

parametrizes a unit hemisphere \mathcal{S} . The boundary not annihilated, so the hemisphere is a surface with boundary, given by

$$\partial \mathcal{S} = \{x \in \mathbb{R}^3 : x_3 = 0, x_1^2 + x_2^2 = 1\}.$$



Stokes's Theorem in \mathbb{R}^3

There is a theorem that states that an oriented hypersurface in \mathbb{R}^3 is closed if and only if it divides \mathbb{R}^3 into two disjoint, connected components. Hence, an oriented hypersurface is closed if and only if it is the boundary of a region in \mathbb{R}^3 .

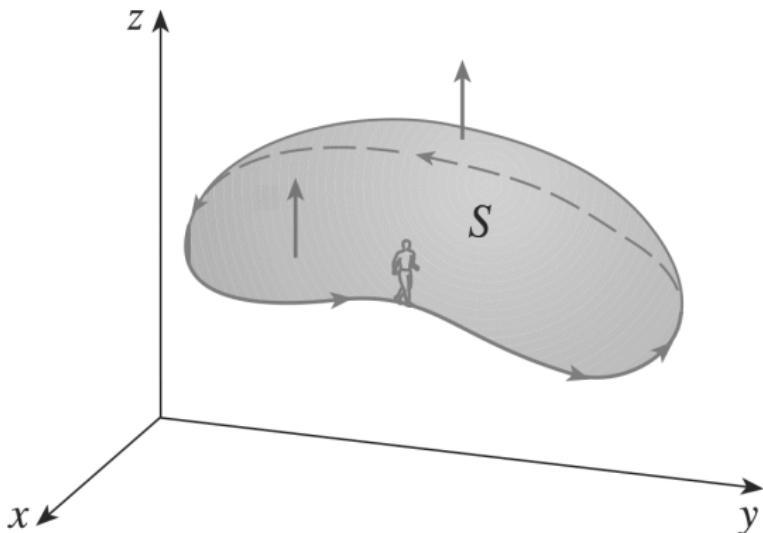
After these preparations we can finally formulate one generalization of Green's theorem to \mathbb{R}^3 :

3.6.7. Stokes's Theorem. Let $\Omega \subset \mathbb{R}^3$ be an open set, $\mathcal{S} \subset \Omega$ a parametrized, admissible surface in \mathbb{R}^3 with boundary $\partial\mathcal{S}$ and let $F: \Omega \rightarrow \mathbb{R}^3$ be a continuously differentiable vector field. Then

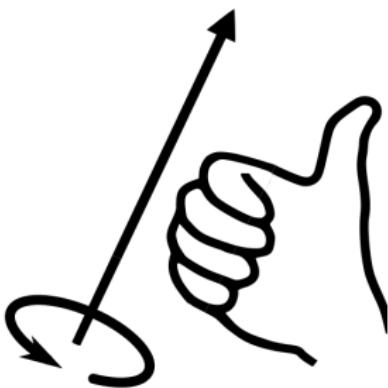
$$\int_{\partial\mathcal{S}^*} F \, d\vec{s} = \int_{\mathcal{S}^*} \operatorname{rot} F \, d\vec{A}$$

where the orientations of the boundary curve $\partial\mathcal{S}^*$ and the surface \mathcal{S}^* are chosen so that the normal vector to \mathcal{S}^* points in the direction of the thumb of the right hand if the four fingers point in the direction of the tangent vector to $\partial\mathcal{S}^*$.

Orientation for Stokes's Theorem in \mathbb{R}^3



Orientation for Stokes's Theorem. James Stewart, **Calculus**, 4th Ed., Brooks Cole



Right-hand Grip Rule Wikimedia Commons. Wikimedia Foundation. Web. 28 July 2012

Physical Interpretation of Stokes's Theorem

The physical interpretation of Stokes's theorem is the same as for Green's theorem: in the integral of infinitesimal circulations (the rotation) across a surface, the individual circulations cancel everywhere except at the boundary, so the total integral is just the circulation along the boundary:

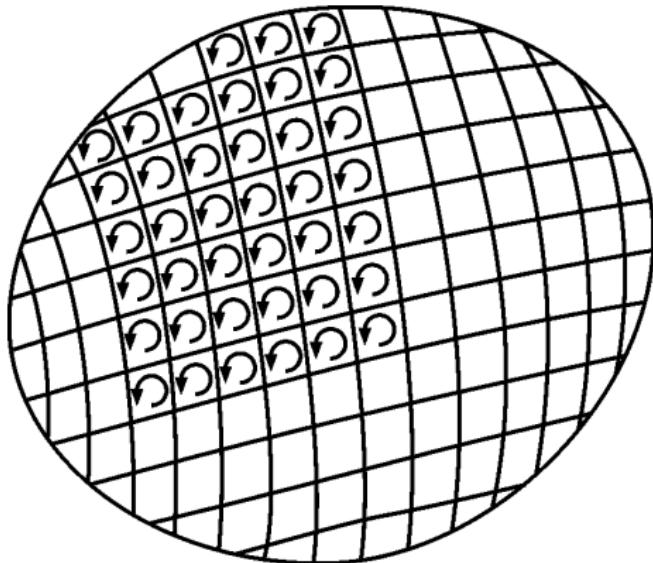
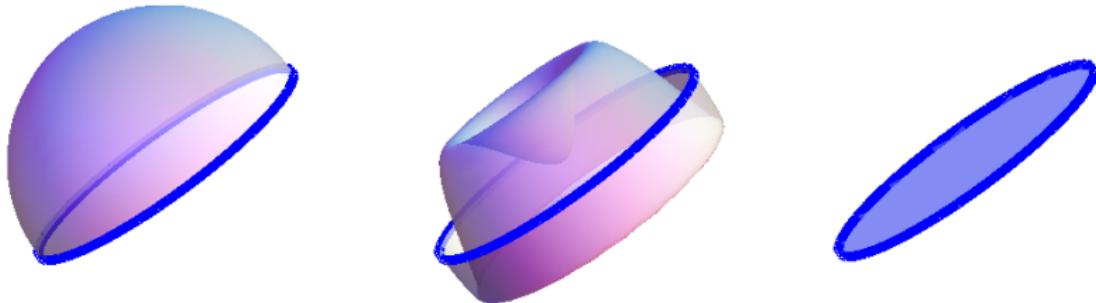


Illustration of Stokes's Theorem. Wikimedia Commons. Wikimedia Foundation. Web. 28 July 2012

Physical Interpretation of Stokes's Theorem

Furthermore, it does not matter how the surface is deformed if the boundary remains the same; the “infinitesimal circulations” continue to cancel each other. The integral of the circulation across the hemisphere below left is not affected by deformations (middle) and even equal to the circulation integrated over the disk with the same boundary (right).



Proof of Stokes's Theorem

Proof of Stokes's Theorem 3.6.7.

Stokes's Theorem in \mathbb{R}^3 can be reduced to Green's Theorem 3.4.18. By definition,

$$\begin{aligned}\int_{\partial^*} \operatorname{rot} F d\vec{A} &= \int_{\partial^*} \langle \operatorname{rot} F, N \rangle dA \\ &= \int_{\Omega} \langle \operatorname{rot} F(\varphi(x_1, x_2), t_1 \times t_2|_{\varphi(x_1, x_2)}) \rangle dx_1 dx_2.\end{aligned}$$

By (3.2.7) and (3.2.3) we have

$$\begin{aligned}\langle \operatorname{rot} F(\varphi(x), t_1 \times t_2|_{\varphi(x)}) \rangle &= \operatorname{rot} F|_{\varphi(x)}(t_1(\varphi(x)), t_2(\varphi(x))) \\ &= \operatorname{rot} F|_{\varphi(x)}\left(\frac{\partial \varphi}{\partial x_1}, \frac{\partial \varphi}{\partial x_2}\right) \\ &= \left\langle DF|_{\varphi(x)} \frac{\partial \varphi}{\partial x_1}, \frac{\partial \varphi}{\partial x_2} \right\rangle - \left\langle \frac{\partial \varphi}{\partial x_1}, DF|_{\varphi(x)} \frac{\partial \varphi}{\partial x_2} \right\rangle\end{aligned}$$

Proof of Stokes's Theorem

Proof of Stokes's Theorem 3.6.7 (continued).

By the chain rule,

$$\frac{\partial}{\partial x_1} \left\langle F|_{\varphi(x)}, \frac{\partial \varphi}{\partial x_2} \right\rangle = \left\langle DF|_{\varphi(x)} \frac{\partial \varphi}{\partial x_1}, \frac{\partial \varphi}{\partial x_2} \right\rangle + \left\langle F|_{\varphi(x)}, \frac{\partial^2 \varphi}{\partial x_1 \partial x_2} \right\rangle$$

and so

$$\begin{aligned} \langle \operatorname{rot} F(\varphi(x), t_1 \times t_2)|_{\varphi(x)} \rangle &= \frac{\partial}{\partial x_1} \left\langle F|_{\varphi(x)}, \frac{\partial \varphi}{\partial x_2} \right\rangle - \frac{\partial}{\partial x_2} \left\langle F|_{\varphi(x)}, \frac{\partial \varphi}{\partial x_1} \right\rangle \\ &= \operatorname{rot} \tilde{F}(x), \end{aligned}$$

where

$$\tilde{F}: \Omega \rightarrow \mathbb{R}^2, \quad \tilde{F}(x) = \begin{pmatrix} \left\langle F|_{\varphi(x)}, \frac{\partial \varphi}{\partial x_1} \right\rangle \\ \left\langle F|_{\varphi(x)}, \frac{\partial \varphi}{\partial x_2} \right\rangle \end{pmatrix}.$$

Proof of Stokes's Theorem

Proof of Stokes's Theorem 3.6.7 (continued).

We can now apply Green's theorem in the admissible region Ω . Then

$$\begin{aligned}\int_{\delta^*} \operatorname{rot} F d\vec{A} &= \int_{\Omega} \langle \operatorname{rot} F(\varphi(x_1, x_2), t_1 \times t_2|_{\varphi(x_1, x_2)}) \rangle dx_1 dx_2 \\ &= \int_{\Omega} \operatorname{rot} \tilde{F}(x) dx_1 dx_2 \\ &= \int_{\partial\Omega} \tilde{F} d\vec{s}.\end{aligned}$$

Let us suppose (for simplicity) that $\partial\Omega$ is given by a single parametrization $\gamma: I \rightarrow \partial\Omega$. Then

$$\int_{\delta^*} \operatorname{rot} F d\vec{A} = \int_{\partial\Omega} \tilde{F} d\vec{s} = \int_I \langle \tilde{F}(\gamma(t)), \gamma'(t) \rangle dt.$$

Proof of Stokes's Theorem

Proof of Stokes's Theorem 3.6.7 (continued).

Inserting the definition of \tilde{F} ,

$$\begin{aligned}\int_{\partial\mathcal{S}^*} \operatorname{rot} F d\vec{A} &= \int_I \left(\left\langle F|_{\varphi(x)}, \frac{\partial \varphi}{\partial x_1} \right\rangle \Big|_{x=\gamma(t)} \gamma'_1(t) \right. \\ &\quad \left. + \left\langle F|_{\varphi(x)}, \frac{\partial \varphi}{\partial x_2} \right\rangle \Big|_{x=\gamma(t)} \gamma'_2(t) \right) dt \\ &= \int_I \left\langle F|_{\varphi(\gamma(t))}, \frac{d}{dt} \varphi(\gamma(t)) \right\rangle dt \\ &= \int_{\partial\mathcal{S}^*} F d\vec{s}\end{aligned}$$

where we have used the chain rule and that $\varphi(\gamma(t))$ parametrizes $\partial\mathcal{S}$. We have assumed a single parametrization for the boundary; in general, this calculation can be applied to each boundary segment. □

Stokes's Theorem in \mathbb{R}^n

3.6.8. Remark. We have formulated and proved Stokes's theorem in \mathbb{R}^3 , since we have a good idea of the structure of the rotation (as a three-dimensional vector) in \mathbb{R}^3 .

To generalize Stokes's theorem to n dimensions would require working with the rotation as a bilinear form and would entail a fair amount of abstract algebra and geometry, including a closer study of differential forms. This is beyond the scope of our course; if you are interested in this, search for books on **vector analysis**.

For example, Michael Spivak's book ***Calculus on Manifolds*** is a good place to start.

Gauß's Theorem

The other aspect of Green's theorem is based on the physical idea of flux.
This has a straightforward generalization to n dimensions:

3.6.9. Gauß's Theorem. Let $R \subset \mathbb{R}^n$ be an admissible region and $F: \overline{R} \rightarrow \mathbb{R}^n$ a continuously differentiable vector field. Then

$$\int_R \operatorname{div} F \, dx = \int_{\partial R^*} F \, d\vec{A}.$$

The integrals make sense, as the boundary of an admissible region is a union of hypersurfaces. Recall that the surfaces are oriented in such a way that the normal vector points outward.

We will prove Gauß's theorem only for the case of ***simple regions***, whose definition we now recall.

Ordinate Regions in \mathbb{R}^n

We generalize ordinate regions to \mathbb{R}^n as follows. For $x \in \mathbb{R}^n$ we define

$$\hat{x}^{(k)} := (x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n) \in \mathbb{R}^{n-1}$$

as the vector x with the k th component omitted.

3.6.10. Definition. A subset $R \subset \mathbb{R}^n$ is said to be an **ordinate region (with respect to x_k)** if there exists a measurable set $\Omega \subset \mathbb{R}^{n-1}$ and continuous, almost everywhere differentiable functions $\varphi_1, \varphi_2: \Omega \rightarrow \mathbb{R}$, such that

$$R = \{x \in \mathbb{R}^n : \hat{x}^{(k)} \in \Omega, \varphi_1(\hat{x}^{(k)}) \leq x_k \leq \varphi_2(\hat{x}^{(k)})\}.$$

If R is an ordinate region with respect to each x_k , $k = 1, \dots, n$, it is said to be a **simple region**.

3.6.11. Theorem. A simple region is admissible.

Proof of Gauß's Theorem for Simple Regions

We will prove Gauß's theorem only for simple regions.

Proof of Gauß's Theorem 3.6.9.

Suppose that $F = (F_1, \dots, F_n)$. Then we have to prove

$$\int_R \operatorname{div} F \, dx = \int_{\partial R^*} F \, d\vec{A},$$

where ∂R^* is oriented by the normal vector pointing outwards. Since

$$\int_R \operatorname{div} F \, dx = \sum_{k=1}^n \int_R \frac{\partial F_k}{\partial x_k} \, dx, \quad \int_{\partial R^*} \langle F, N \rangle \, dA = \sum_{k=1}^n \int_{\partial R} F_k \langle e_k, N \rangle \, dA$$

it is sufficient to show that

$$\int_R \frac{\partial F_k}{\partial x_k} \, dx = \int_{\partial R^*} F_k \langle e_k, N \rangle \, dA \tag{3.6.2}$$

for $k = 1, \dots, n$.

Proof of Gauß's Theorem for Simple Regions

Proof of Gauß's Theorem 3.6.9 (continued).

We use that R is a simple region, so we can write

$$R = \{x \in \mathbb{R}^n : \hat{x}^{(k)} \in \Omega, \varphi_1(\hat{x}^{(k)}) \leq x_k \leq \varphi_2(\hat{x}^{(k)})\}.$$

for some $\Omega \subset \mathbb{R}^{n-1}$. The boundary of R is given by

$$\begin{aligned} \partial R^* &= \underbrace{\left\{x \in \mathbb{R}^n : \hat{x}^{(k)} \in \Omega, x_k = \varphi_1(\hat{x}^{(k)})\right\}}_{=: \partial_1} \\ &\cup \underbrace{\left\{x \in \mathbb{R}^n : \hat{x}^{(k)} \in \Omega, x_k = \varphi_2(\hat{x}^{(k)})\right\}}_{=: \partial_2} \\ &\cup \underbrace{\left\{x \in \mathbb{R}^n : \hat{x}^{(k)} \in \partial\Omega, \varphi_1(\hat{x}^{(k)}) \leq x_k \leq \varphi_2(\hat{x}^{(k)})\right\}}_{=: \partial_3} \end{aligned}$$

Proof of Gauß's Theorem for Simple Regions

Proof of Gauß's Theorem 3.6.9 (continued).

It is left as an exercise to show that the normal vector N to the “mantle” \mathcal{S}_3 is orthogonal to e_k (by writing down a parametrization and showing that e_3 is in fact a tangent vector). Then the surface integral in (3.6.2) becomes

$$\int_{\partial R^*} F_k \langle e_k, N \rangle dA = \int_{\partial \mathcal{S}_1^*} F_k \langle e_k, N_1 \rangle dA + \int_{\partial \mathcal{S}_2^*} F_k \langle e_k, N_2 \rangle dA. \quad (3.6.3)$$

To evaluate the surface integrals, we need to find the unit normal vectors N_1 and N_2 . The surface \mathcal{S}_1 is parametrized by

$$\Phi_1(\hat{x}^{(k)}) = \left(x_1, \dots, x_{x-1}, \varphi_1(\hat{x}^{(k)}), x_{k+1}, \dots, x_n \right)^T.$$

Proof of Gauß's Theorem for Simple Regions

Proof of Gauß's Theorem 3.6.9 (continued).

Then the j th tangent vector is given by

$$t_j = \frac{\partial \Phi(\hat{x}^{(k)})}{\partial x_j} = \left(0, \dots, 0, \underset{j}{\overset{\uparrow}{1}}, 0, \dots, 0, \frac{\partial \varphi_1}{\partial x_j}, 0, \dots, 0 \right)^T$$
$$\qquad\qquad\qquad \underset{k}{\overset{\uparrow}{}}$$

for $j \in \{1, \dots, n\} \setminus \{k\}$. The normal vector will be orthogonal to all $n - 1$ tangent vectors, so

$$N_1 = C \cdot \left(\frac{\partial \varphi_1}{\partial x_1}, \dots, \frac{\partial \varphi_1}{\partial x_{k-1}}, -1, \frac{\partial \varphi_1}{\partial x_{k+1}}, \dots, \frac{\partial \varphi_1}{\partial x_n} \right)^T$$

where $C \in \mathbb{R}$ is a normalization constant.

Proof of Gauß's Theorem for Simple Regions

Proof of Gauß's Theorem 3.6.9 (continued).

Since N_1 is to have unit length,

$$\frac{1}{C} = \left(1 + \sum_{\substack{j=1 \\ j \neq k}}^n \left(\frac{\partial \varphi_1}{\partial x_j} \right)^2 \right) = 1 + |D\varphi_1|^2.$$

We note that, after re-arranging rows,

$$\begin{aligned} |\det(t_1, \dots, t_{n-1}, N_1)| &= \left| (-1)^{n-k} \det \begin{pmatrix} \mathbb{1}_{n-1} & (D\varphi_1)^T \\ D\varphi_1 & -1 \end{pmatrix} \right| \\ &= \left| \det \begin{pmatrix} \mathbb{1}_{n-1} & 0 \\ D\varphi_1 & -1 - \sum_{j \neq k} \left(\frac{\partial \varphi_1}{\partial x_j} \right)^2 \end{pmatrix} \right| \\ &= \frac{1}{C}. \end{aligned}$$

Proof of Gauß's Theorem for Simple Regions

Proof of Gauß's Theorem 3.6.9 (continued).

We see that $\langle e_k, N_1 \rangle = -C$ and hence

$$\begin{aligned} \int_{\partial\delta_1^*} F_k \langle e_k, N_1 \rangle dA &= -C \int_{\Omega} F_k(\phi_1(\hat{x}^{(k)})) \underbrace{|\det(t_1, \dots, t_{n-1}, N_1)|}_{=1/C} d\hat{x}^{(k)} \\ &= - \int_{\Omega} F_k(x_1, \dots, x_{k-1}, \varphi_1(\hat{x}^{(k)}), x_{k+1}, \dots, x_n) d\hat{x}^{(k)} \end{aligned}$$

In the same way we can show that

$$\int_{\partial\delta_2^*} F_k \langle e_k, N_2 \rangle dA = \int_{\Omega} F_k(x_1, \dots, x_{k-1}, \varphi_2(\hat{x}^{(k)}), x_{k+1}, \dots, x_n) d\hat{x}^{(k)}$$

Proof of Gauß's Theorem for Simple Regions

Proof of Gauß's Theorem 3.6.9 (continued).

Finally, we obtain

$$\begin{aligned}\int_{\partial R^*} F_k \langle e_k, N \rangle dA &= \int_{\partial \mathcal{O}_2^*} F_k dA + \int_{\partial \mathcal{O}_1^*} F_k dA \\&= \int_{\Omega} F_k(x_1, \dots, x_{k-1}, \varphi_2(\hat{x}^{(k)}), x_{k+1}, \dots, x_n) d\hat{x}^{(k)} \\&\quad - \int_{\Omega} F_k(x_1, \dots, x_{k-1}, \varphi_1(\hat{x}^{(k)}), x_{k+1}, \dots, x_n) d\hat{x}^{(k)} \\&= \int_{\Omega} \int_{\varphi_1(\hat{x}^{(k)})}^{\varphi_2(\hat{x}^{(k)})} \frac{\partial F_k}{\partial x_k}(x) dx_k d\hat{x}^{(k)} \\&= \int_R \frac{\partial F_k}{\partial x_k} dx,\end{aligned}$$

which shows (3.6.2). □

Application to Electromagnetics

We have shown in Example 3.5.20 that the flux through the unit sphere S^2 of the electric field

$$E(p) = \frac{1}{4\pi\varepsilon_0} \frac{Q}{\|p\|^3} p$$

induced by a point charge at the origin is

$$\int_{S^2} \langle E, d\vec{A} \rangle = \frac{Q}{\varepsilon_0}.$$

This calculation can easily be modified so that it holds for any sphere $\partial B_r(0)$ of radius $r > 0$. Now let $R \subset \mathbb{R}^3$ be an admissible region containing the origin. Then the flux through R is given by

$$\begin{aligned} \int_R \langle E, d\vec{A} \rangle &= \int_{\partial B_r(0)} \langle E, d\vec{A} \rangle + \int_{\partial(R \setminus B_r(0))} \langle E, d\vec{A} \rangle \\ &= \frac{Q}{\varepsilon_0} + \int_{R \setminus B_r(0)} \operatorname{div} E \, dx. \end{aligned}$$

Application to Electromagnetics

Now

$$\begin{aligned}\frac{\partial E_i}{\partial x_i} &= \frac{Q}{4\pi\varepsilon_0} \frac{\partial}{\partial x_i} \frac{x_i}{(x_1^2 + x_2^2 + x_3^2)^{3/2}} \\ &= \frac{Q}{4\pi\varepsilon_0} \frac{1}{(x_1^2 + x_2^2 + x_3^2)^3} ((x_1^2 + x_2^2 + x_3^2)^{3/2} - 3(x_1^2 + x_2^2 + x_3^2)^{1/2} x_i^2).\end{aligned}$$

Since

$$\sum_{i=1}^3 ((x_1^2 + x_2^2 + x_3^2)^{3/2} - 3(x_1^2 + x_2^2 + x_3^2)^{1/2} x_i^2) = 0,$$

we see that $\operatorname{div} E(x) = 0$ for $x \neq 0$. This implies that

$$\int_R \langle E, d\vec{A} \rangle = \frac{Q}{\varepsilon_0}$$

for any admissible region $R \subset \mathbb{R}^3$.

The Normal Derivative

3.6.12. Definition. Let $R \subset \mathbb{R}^n$ be a region and its boundary $\mathcal{S} = \partial R$ a hypersurface parametrized by $\varphi: \Omega \rightarrow \mathcal{S}$. Let $N: \mathcal{S} \rightarrow \mathbb{R}^n$ be the normal vector field on \mathcal{S} and $u: \overline{R} \rightarrow \mathbb{R}$ be a continuously differentiable potential function. Then

$$\frac{\partial u}{\partial n}(x) := \langle \nabla u(\varphi(x)), N \circ \varphi(x) \rangle$$

is called the **normal derivative** of u on \mathcal{S} .

We have previously introduced the normal derivative in \mathbb{R}^2 ; see Definition 2.4.4.

Green's Identities

3.6.13. **Green's Identities.** Let $R \subset \mathbb{R}^n$ be an admissible region and $u, v: \overline{R} \rightarrow \mathbb{R}$ be twice continuously differentiable potential functions. Then

$$\int_R \langle \nabla u, \nabla v \rangle dx = - \int_R u \cdot \Delta v dx + \int_{\partial R^*} u \frac{\partial v}{\partial n} dA \quad (3.6.4)$$

and

$$\int_R (u \cdot \Delta v - v \cdot \Delta u) dx = \int_{\partial R^*} \left(u \frac{\partial v}{\partial n} - v \frac{\partial u}{\partial n} \right) dA. \quad (3.6.5)$$

(3.6.4) is commonly called ***Green's first identity*** and (3.6.5) ***Green's second identity***.

Green's identities can be regarded as “integration by parts for ∇ and Δ ”. They play an important role in problems relating to the Laplace partial differential equation, $\Delta u = 0$.

Green's Identities

Proof.

We have proven in the assignments that

$$\operatorname{div}(u \nabla v) = u \underbrace{\operatorname{div}(\nabla v)}_{=\Delta v} + \langle \nabla u, \nabla v \rangle.$$

Applying Gauß's theorem,

$$\begin{aligned}\int_R \langle \nabla u, \nabla v \rangle dx &= \int_R (\operatorname{div}(u \nabla v) - u \Delta v) dx \\&= \int_{\partial R^*} u \nabla v \cdot d\vec{A} - \int_R u \Delta v dx \\&= \int_{\partial R^*} u \langle \nabla v, N \rangle dA - \int_R u \Delta v dx \\&= \int_{\partial R^*} u \frac{\partial v}{\partial n} dA - \int_R u \Delta v dx.\end{aligned}$$

Green's Identities

Proof (continued).

This proves the first identity, (3.6.4). The second identity, (3.6.5), follows by subtracting the two equations

$$\int_R \langle \nabla u, \nabla v \rangle dx = - \int_R u \cdot \Delta v dx + \int_{\partial R^*} u \frac{\partial v}{\partial n} dA,$$
$$\int_R \langle \nabla v, \nabla u \rangle dx = - \int_R v \cdot \Delta u dx + \int_{\partial R^*} v \frac{\partial u}{\partial n} dA$$

from each other. □