

Notes on Multivariable Calculus II

Izabella Laba

Caleb Marshall

NOTES ON MULTIVARIABLE CALCULUS II

Izabella LABA

Caleb MARSHALL

Copyright information

- Copyright © 2026 Izabella Laba and Caleb Marshall
- Cover page photo by Izabella Laba
- Cover page design by Krysten Casumpang
- This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. You can view a copy of the license at <https://creativecommons.org/licenses/by-nc-sa/4.0/>.



About Math 227

Math 226 and 227 are multivariable calculus classes recommended for Honours students in Mathematics at the University of British Columbia. Math 227 is the second course in the sequence, with Math 226 as a required prerequisite. It covers curves in two and three dimensions, surfaces in three dimensions, line and surface integrals, vector fields, and the main theorems of vector calculus: Green's Theorem, Divergence (Gauss's) Theorem, Stokes's Theorem. As in Math 226, there is more emphasis on rigorous proofs and critical mathematical thinking, and less on computations and applications.

A major difference between Math 227 and its more conventional counterparts is that the last few weeks are devoted to an introduction to differential forms. This is a general theory that unites many of the geometric concepts introduced earlier in class, extends them to arbitrary dimensions, and provides a glimpse into the beauty of more abstract mathematics. The class ends with the Generalized Stokes Theorem, a common framework for the main theorems of vector calculus.

We spend less time on epsilon-delta proofs and local regularity issues, having covered that type of questions quite extensively in Math 226. Instead, we usually assume enough regularity to proceed with the main constructions, and then focus on the geometry involved. The key new skills include visualizing 3-dimensional geometric structures (such as surfaces or moving frames of vectors), writing up geometric arguments involving descriptions and manipulations of such structures, and, eventually, extrapolating arguments of this type to higher dimensions.

These notes do not aim to be a complete treatment of the subject. I recommend using them together with a textbook for additional examples, applications, and further details of some of the proofs that are only sketched here. In Math 226 and 227, we have been using *Calculus: Several Variables* (also available as the second half of *Calculus: A Complete Course*), by Robert A. Adams and Christopher Essex, published by Pearson. It is one of the few calculus books that cover differential forms, including in higher dimensions. For a fully rigorous treatment, I would refer the interested student to Folland's *Advanced Calculus*.

Most of the problems here are taken from my exams and homework assignments. The Worked Problems come with solutions; the solutions to the Practice Problems are provided in a separate Solutions Manual volume. Some are routine computations: even as we develop a more general theory, it still needs to be grounded in specifics. I encourage students to use their textbooks and online systems such as WebWork for more practice of this type. The proof questions may require more time and some creative thinking. I ask my students to write up the solutions clearly, in full sentences, with correct and complete explanations provided.

About these notes

The making of this book follows the same story that was described in the preface to the Math 226 volume, so I will be brief here. This volume is, again, based on the class notes that I developed while teaching the course repeatedly over the last two decades. I first made the notes available to students for the online and partially-online editions of Math 227 between 2020 and 2022. They were mostly handwritten then, except that I typed up parts of the differential forms material. All remaining notes (including most of this book) were initially typeset by Caleb Marshall. Caleb has also prepared the professional-quality illustrations based on my hand-drawn drafts, and worked with me on some of the problems and examples (for instance, he contributed Example 2.2.3 in Section 2.2). Between 2023 and 2025, we both edited various parts of the initial text. Mihai Marian compiled the Math 227 files into two volumes, this one and the separate Solutions Manual. I made the final revisions, and added the more recent exam and homework problems, in the Spring and Summer of 2026.

Acknowledgements

This book was written on the Point Grey campus of the University of British Columbia, located on the traditional, ancestral, and unceded territory of the Musqueam people.

I am grateful to the Department of Mathematics at the University of British Columbia, and especially to Daniel Coombs (Head) and Fok-Shuen Leung (Undergraduate Chair), for approving and supporting the project. I owe many thanks to Caleb Marshall for the two years of working on it together. Krysten Casumpang, a staff member in the department, did a great job designing the cover pages.

I also would like to acknowledge all other Teaching Assistants who helped me run these courses over the years. In Math 227, this includes Sebastian Gant, Elliot Cheung, Manuel Ruivo de Oliveira, Angel Cruz, Chenjian Wang, Yuveshen Moorooogen, Hao Ouyang, and Noureldein Hanafi. Unfortunately I have not kept TA records from 2016 and earlier, but nonetheless I would like to thank any of my former TAs who might be reading this.

Last but not least, I'm grateful to all undergraduate students who took Math 227 with me. They have provided valuable and actionable feedback, asked challenging questions, and otherwise pushed me to improve this course. They have also pointed out many typos and errors in earlier versions of these notes. It was a pleasure to work with them.

Izabella Łaba

Vancouver, July 2026

Unit tangent and curvature

(11)

Question: How quickly does the curve change direction?

Direction is represented by the tangent vector, so a change of direction should be represented by the derivative of the tangent vector. We will consider the unit tangent vector, since we are only interested in its direction, not magnitude.

We want this to have a geometric meaning independent of parametrization (depending only on the shape of the curve). Therefore we will use arclength parametrization (i.e. motion along the curve with constant speed 1.)

Note on arclength parametrization: we can use it to define concepts even when it would be difficult or impossible to find the arclength parametrization explicitly. Eventually, we will have formulas for the same concepts that do not require us to know an explicit formula for arclength.

Assume $\vec{r}(t): [a, b] \rightarrow \mathbb{R}^3$ is differentiable, $\vec{r}'(t)$ is continuous, $\vec{r}'(t) \neq 0$ on (a, b) .

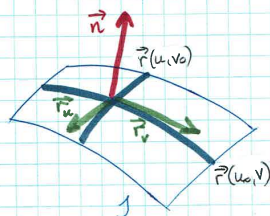
We want formulas for:

(3/2)

(1) Normal vector / tangent plane, if the surface is given in terms of parametrization $\vec{r}(u, v)$

(2) Surface area element, for the purpose of integration.

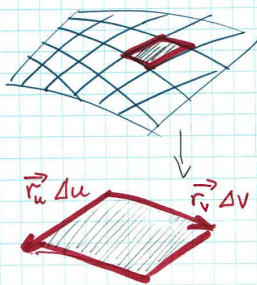
(1) Normal vector at $P = \vec{r}(u_0, v_0)$: the curves $\vec{r}(u, v_0)$ and $\vec{r}(u_0, v)$ pass through P and have tangent vectors $\vec{r}_u(u_0, v_0)$, $\vec{r}_v(u_0, v_0)$.



Assuming that $\vec{r}_u$ and $\vec{r}_v$ are not parallel, we can take

$$\vec{n} = \vec{r}_u \times \vec{r}_v$$

(2) Area element: for the purpose of integration, subdivide S using the (not necessarily rectangular) grid consisting of curves $u = \text{const}$, $v = \text{const}$. This divides S into approximate parallelograms spanned by vectors $\vec{r}_u \Delta u$ and $\vec{r}_v \Delta v$.



The area of such a parallelogram is

$$\begin{aligned} dS &= |(\vec{r}_u \Delta u) \times (\vec{r}_v \Delta v)| \\ &= |\vec{r}_u \times \vec{r}_v| \Delta u \Delta v \\ &= |\vec{n}| \Delta u \Delta v \end{aligned}$$

with $\vec{n}$ as above

Theorem (Divergence as flux density)

(4/4)

If J_E is the sphere of radius E centered at P , $\vec{F}$ smooth, $\vec{N}$ points outward, then

$$\text{div } \vec{F}(P) = \lim_{E \rightarrow 0} \frac{1}{\frac{4}{3}\pi E^3} \int_{J_E} \vec{F} \cdot \vec{N} dS$$

volume of the ball enclosed by J_E .

Approximately,

$$\int_{J_E} \vec{F} \cdot \vec{N} dS \approx \left(\frac{4}{3}\pi E^3\right) \cdot \text{div } \vec{F}(P)$$

Proof: Assume $P=0$, let $\vec{F} = F_1 \vec{i} + F_2 \vec{j} + F_3 \vec{k}$, expand F_1, F_2, F_3 in Taylor series, evaluate F_1 contribution:

$$F_1(\vec{r}) = F_1(\vec{0}) + F_{1,x}(\vec{0})x + F_{1,y}(\vec{0})y + F_{1,z}(\vec{0})z + O(E^2)$$

differential approximation error estimate

$$\vec{N} = \frac{1}{E}(x\vec{i} + y\vec{j} + z\vec{k}) \text{ on } J_E \text{ (using } |\vec{r}| = E)$$

$$\begin{aligned} F_1 \vec{i} \cdot \vec{N} &= F_1 \frac{x}{E} = \frac{1}{E} [F_1(\vec{0})x + F_{1,x}(\vec{0})x^2 + F_{1,y}(\vec{0})xy \\ &\quad + F_{1,z}(\vec{0})xz] + \frac{1}{E} O(E^2) \cdot x \end{aligned}$$

$O(E^2)$ since $|x| \leq E$

By symmetry and cancellation,

$$\int_{J_E} x dS = \int_{J_E} xy dS = \int_{J_E} xz dS = 0$$

Stokes's Theorem

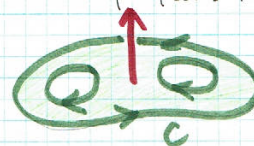
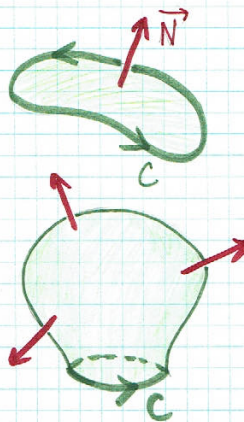
(4/3)

Theorem: Let S be a piecewise smooth, oriented surface in $\mathbb{R}^3$, with normal $\vec{N}$, and suppose that the boundary of S consists of one or more closed, piecewise smooth, simple curves with orientation inherited from S . Let $\vec{F}$ be a smooth vector field on an open set containing S .

$$\oint_C \vec{F} \cdot d\vec{r} = \iint_S \text{curl } \vec{F} \cdot \vec{N} dS$$

Inherited orientation: counterclockwise when viewed from the direction of $\vec{N}$

(reverse if the surface has "holes" - same as for Green's Theorem)



Green's Theorem is a special case of Stokes's Theorem, with $S = D$ (a planar region) and $\vec{N} = \vec{k}$.

Contents

1	Parametrized Curves	1
1.1	Curves and Vector-Valued Functions	1
1.2	Arc Length for Curves	8
1.3	Unit Tangent and Curvature	16
1.4	The Principal Normal and The Osculating Plane	22
1.5	Torsion	28
1.6	Velocity and Acceleration	33
2	Vector Fields	41
2.1	Introduction to Vector Fields	41
2.2	Conservative Vector Fields	51
2.3	Line Integrals	59
2.4	Path Independence	68
3	Parametric Surfaces and Surface Integration	77
3.1	Surfaces and Area Integrals	77
3.2	Flux Integrals	89
4	Main Theorems of Vector Calculus	97
4.1	Divergence and Curl	97
4.2	Green's Theorem in Two Dimensions	104
4.3	The Divergence Theorem	118
4.4	Stokes's Theorem	127
5	Differential Forms	139
5.1	An Introduction to Differential Forms	139
5.2	Exterior Derivative	148

5.3	Integrating Differential Forms	155
5.4	Orientation on Manifolds	160
5.5	Generalized Stokes's Theorem	169

Chapter 1

Parametrized Curves

1.1 Curves and Vector-Valued Functions

1.1.1 Continuity and Derivatives for Vector-Valued Functions

In this chapter, we will work with **vector-valued functions** of a single variable. These are functions $\mathbf{r} : I \rightarrow \mathbb{R}^n$ (for now, we will have $n = 2$ or $n = 3$), where $I \subset \mathbb{R}$ is an interval. We can always write the function $\mathbf{r} = \mathbf{r}(t)$ in terms of its component functions:

$$\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}$$

where $x, y, z : I \rightarrow \mathbb{R}$ are scalar-valued functions.

Definition 1.1.1. *The vector-valued function $\mathbf{r} : I \rightarrow \mathbb{R}^n$ is called **continuous** if and only if all component functions $x(t), y(t)$ and $z(t)$ are continuous on the interval I .*

A useful interpretation is to think of $\mathbf{r}$ as describing the motion of an object in space, so that the coordinates of that object at time t are $(x(t), y(t), z(t))$. The assumption that $\mathbf{r}$ is continuous guarantees that the object does not make any instantaneous jumps from one location to another.

Definition 1.1.2. *A **curve** is the set $\{\mathbf{r}(t) : t \in I\} \subset \mathbb{R}^n$ of values attained by a continuous vector-valued function. We refer to $\mathbf{r}$ as a **parametrization** of the curve.*

With our previous interpretation of vector valued functions, we interpret the curve $\{\mathbf{r}(t) : t \in I\}$ as the path travelled by a particle or object over time, but without any information on where the particle was at any given time $t \in I$. Note that the same curve may have many different parametrizations, corresponding to the object travelling the same path but at different speed and possibly in different directions.

Definition 1.1.3. *The vector-valued function $\mathbf{r} : I \rightarrow \mathbb{R}^n$ is **differentiable** on an open interval I if and only if all component functions $x(t), y(t)$ and $z(t)$ are differentiable*

on I . In this case, we write

$$\frac{d\mathbf{r}}{dt} = \mathbf{r}'(t) := x'(t)\mathbf{i} + y'(t)\mathbf{j} + z'(t)\mathbf{k}.$$

Similarly, we will say that $\mathbf{r}$ is C^k on an interval I if and only if all of its component functions are C^k on that interval.

Following our interpretation, the function $\mathbf{r}'(t) = \mathbf{v}(t)$ is the (instantaneous) velocity of our moving object at time t , defined as the limit of the **average velocity** on the interval $(t, t + \Delta t)$:

$$\frac{\mathbf{r}(t + \Delta t) - \mathbf{r}(t)}{\Delta t} \rightarrow \mathbf{v}(t) \text{ as } \Delta t \rightarrow 0.$$

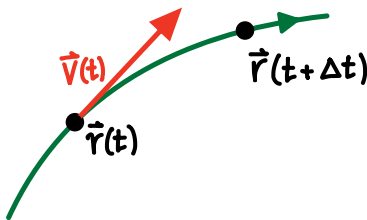


Figure 1.1: The particle at times t and $t + \Delta t$, as well as the instantaneous velocity at time t . The green arrow indicates the direction of motion of the particle along its associated curve.

The function $\mathbf{v}$ is, itself, vector-valued, so that it has both *magnitude* and *direction*. The magnitude of instantaneous velocity, given by $v(t) := |\mathbf{v}(t)|$, is the **speed** at time t . Further, the **acceleration** at time t is given by the formula $\mathbf{a}(t) = \mathbf{v}'(t) = \mathbf{r}''(t)$.

1.1.2 Some Examples

The following are examples of curves in $\mathbb{R}^2$ and $\mathbb{R}^3$, with some associated parametrizations.

Example 1.1.4. Let $P_0 := (x_0, y_0, z_0)$ be some point in $\mathbb{R}^3$ and let $\mathbf{v}$ be some (constant) vector. Then, a straight line through P_0 in the direction $\mathbf{v}$ can be parametrized as

$$\mathbf{r}(t) := \vec{OP}_0 + t\mathbf{v} = \mathbf{r}_0 + t\mathbf{v}, \quad t \in \mathbb{R},$$

where $\mathbf{r}_0 = \vec{OP}_0$. With this parametrization, we have $\mathbf{v}(t) = \mathbf{r}'(t) = \mathbf{v}$, so that the velocity of $\mathbf{r}$ is constant at all times $t \in \mathbb{R}$.

However, the same line can also be parametrized as:

$$\mathbf{r}_2(t) = \mathbf{r}_0 + 2t\mathbf{v} \tag{1.1.1}$$

$$\mathbf{r}_3(t) = \mathbf{r}_0 + t^3 \mathbf{v} \quad (1.1.2)$$

Both (1.1.1) and (1.1.2) describe the same curve. However, in (1.1.1), the velocity is now $2\mathbf{v}$ (so double that of our original $\mathbf{r}$); whereas, in (1.1.2), the velocity is $3t^2\mathbf{v}$, and so is variable in time $t \in \mathbb{R}$.

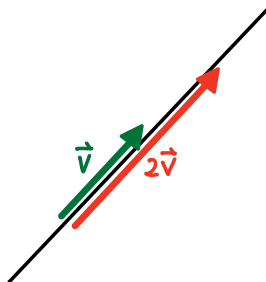


Figure 1.2: Two different velocities along the same line, discussed in Example 1.1.4.

Example 1.1.5. The function $\mathbf{r}(t) = \cos(t)\mathbf{i} + \sin(t)\mathbf{j}$ gives a parametrization of the unit circle $x^2 + y^2 = 1$ in the xy -plane. The velocity and acceleration are

$$\mathbf{v}(t) := -\sin(t)\mathbf{i} + \cos(t)\mathbf{j}$$

$$\mathbf{a}(t) = -\cos(t)\mathbf{i} - \sin(t)\mathbf{j}$$

Example 1.1.6. The equation

$$\mathbf{r}(t) = \cos(t)\mathbf{i} + \sin(t)\mathbf{j} + t\mathbf{k},$$

describes a helix curve. This curve wraps around the z -axis, and is contained in the surface of the cylinder $x^2 + y^2 = 1$. The velocity, speed and acceleration are given by

$$\mathbf{v}(t) = -\sin(t)\mathbf{i} + \cos(t)\mathbf{j} + \mathbf{k}$$

$$v(t) := |\mathbf{v}(t)| = \sqrt{\sin^2(t) + \cos^2(t) + 1^2} = \sqrt{2}$$

$$\mathbf{a}(t) := -\cos(t)\mathbf{i} - \sin(t)\mathbf{j}$$

Example 1.1.7. The function $\mathbf{r}(t) = \cos(t)\mathbf{i} + \sin(t)\mathbf{j} + t^3\mathbf{k}$ parametrizes a “deformed” helix, where the loops (corresponding to full revolutions about the z -axis) are compressed in the z -direction for small $|z|$ and elongated for large values of $|z|$.

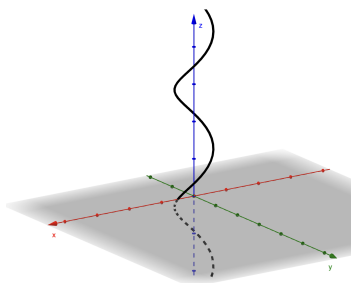


Figure 1.3: The helical curve described in Example 1.1.6.

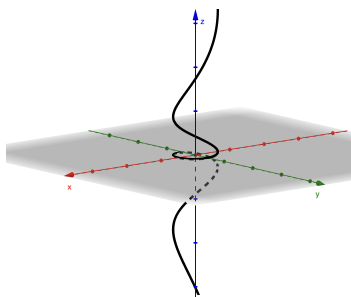


Figure 1.4: The deformed helix in Example 1.1.7.

We have the following differentiation rules for vector-valued functions. These follow easily from the single-variable differentiation rules applied to the component functions.

Proposition 1.1.8. *Suppose that $\mathbf{u}, \mathbf{v} : I \rightarrow \mathbb{R}^3$ are vector-valued functions and $\lambda : I \rightarrow \mathbb{R}$ is scalar-valued, all of whom are differentiable. Then,*

$$\begin{aligned} (\mathbf{u} + \mathbf{v})' &= \mathbf{u}' + \mathbf{v}' \\ (\lambda \mathbf{u})' &= \lambda' \mathbf{u} + \lambda \mathbf{u}' \\ (\mathbf{u} \cdot \mathbf{v})' &= \mathbf{u}' \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{v}' \\ (\mathbf{u} \times \mathbf{v})' &= \mathbf{u}' \times \mathbf{v} + \mathbf{u} \times \mathbf{v}' \end{aligned}$$

Further, if $s : J \rightarrow I$, we have the following **vector-valued Chain Rule**

$$\frac{d}{dt} [\mathbf{u}(s(t))] = s'(t) \mathbf{u}'(s(t)).$$

We often use this identity when **reparametrizing curves** (more on this to come).

Worked Problems

In Questions 1 and 2, we parametrize curves at the intersection of two surfaces. In each case, it helps that one of those surfaces projects to a simple curve (circle or line) in

the xy -plane, and so does the 3-dimensional curve we need to parametrize. Therefore we can parametrize $x(t)$ and $y(t)$ first (we already know how to do it for both the line and the circle), and then we use the equation of the other surface to find $z(t)$. In general, if no such simplifications are available, parametrizing an intersection curve can be significantly more difficult.

Question 4 is important: we will use this equivalence many times in this class.

1. Parametrize the curve of intersection of the cylinder $x^2 + y^2 = 4$ and the plane $z = 3x + 5y$.

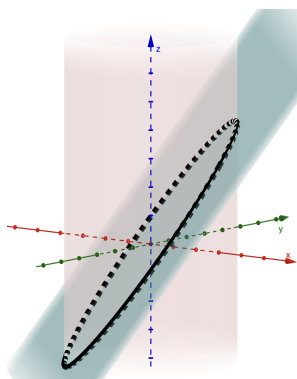


Figure 1.5: The curve described in Worked Problem 1.

Let $x(t) = 2 \cos t$ and $y(t) = 2 \sin t$, so that

$$x^2 + y^2 = 4(\cos^2 t + \sin^2 t) = 4.$$

Then, for any $t, z \in \mathbb{R}$, the point $(x(t), y(t), z)$ lies on the cylinder $x^2 + y^2 = 4$. For the point to also lie in the plane $z = 3x + 5y$, we let

$$z(t) = x(t) + y(t) = 6 \cos t + 10 \sin t.$$

Our parametrization is then

$$\mathbf{r}(t) := 2 \cos t \mathbf{i} + 2 \sin t \mathbf{j} + (6 \cos t + 10 \sin t) \mathbf{k},$$

and we choose $t \in [0, 2\pi]$, which corresponds to one full revolution of the curve.

2. Parametrize the curve of intersection of the paraboloid $z = x^2 + y^2$ and the plane $x = 2y$.

Set $y(t) = t$. Then from the equation of the plane, $x(t) = 2y(t) = 2t$, and from the equation of the paraboloid, $z(t) = x^2(t) + y^2(t) = 5t^2$. Our parametrization is then

$$\mathbf{r}(t) = 2t \mathbf{i} + t \mathbf{j} + 5t^2 \mathbf{k}, \quad t \in \mathbb{R}.$$

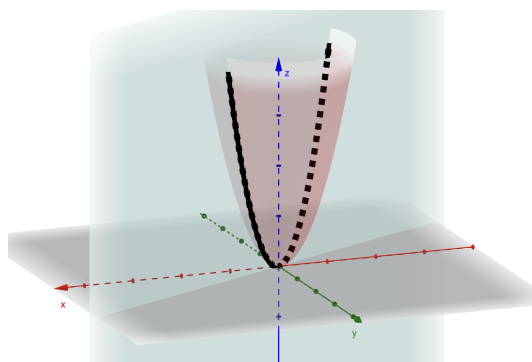


Figure 1.6: The curve described in Worked Problem 2.

3. Find a formula for $\frac{d}{dt}|\mathbf{r}(t)|$.

By the Chain Rule for scalar functions, we have

$$\begin{aligned}
 \frac{d}{dt}|\mathbf{r}(t)| &= \frac{d}{dt}\sqrt{\mathbf{r} \cdot \mathbf{r}} = \frac{1}{2\sqrt{\mathbf{r} \cdot \mathbf{r}}} \cdot \frac{d}{dt}(\mathbf{r} \cdot \mathbf{r}) \\
 &= \frac{1}{2\sqrt{\mathbf{r} \cdot \mathbf{r}}} (\mathbf{r}' \cdot \mathbf{r} + \mathbf{r} \cdot \mathbf{r}') \\
 &= \frac{2\mathbf{r} \cdot \mathbf{r}'}{2\sqrt{\mathbf{r} \cdot \mathbf{r}}} \\
 &= \frac{\mathbf{r} \cdot \mathbf{r}'}{|\mathbf{r}|}
 \end{aligned}$$

4. Suppose $\mathbf{r} : I \rightarrow \mathbb{R}^n$ is differentiable, where I is an interval. Prove that $|\mathbf{r}|$ is constant if and only if $\mathbf{v} \cdot \mathbf{r} = 0$ for all $t \in I$.

(Geometrically, the condition that $|\mathbf{r}| = a$ for some constant $a \geq 0$ means that the trajectory of $\mathbf{r}(t)$ must lie on the sphere of radius a centered at the origin. The condition that $\mathbf{v} \cdot \mathbf{r} = 0$ means that $\mathbf{r} \perp \mathbf{v}$. Here and below, we use the convention that the zero vector is perpendicular to any vector.)

For any $a \in [0, \infty)$, we have the chain of equivalences

$$|\mathbf{r}| = a \Leftrightarrow |\mathbf{r}|^2 = a^2 \Leftrightarrow \mathbf{r} \cdot \mathbf{r} = a^2$$

So, $|\mathbf{r}|$ is constant if and only if $\mathbf{r} \cdot \mathbf{r}$ is constant, if and only if $\frac{d}{dt}(\mathbf{r} \cdot \mathbf{r}) = 0$. However,

$$\frac{d}{dt}(\mathbf{r} \cdot \mathbf{r}) = \mathbf{r}' \cdot \mathbf{r} + \mathbf{r} \cdot \mathbf{r}' = 2\mathbf{r}' \cdot \mathbf{r} = 2\mathbf{v} \cdot \mathbf{r}$$

Combining the above, we get that $|\mathbf{r}|$ is constant if and only if $\mathbf{v} \cdot \mathbf{r} = 0$ for all $t \in I$.

Remark 1.1.1. (a) By the same argument, $|\mathbf{v}|$ is constant if and only if $\mathbf{a} \cdot \mathbf{v} = 0$ for all t .

(b) If instead of $\mathbf{r} \cdot \mathbf{v} = 0$ we have $\mathbf{r} \cdot \mathbf{v} > 0$ for all t , then

$$\frac{d}{dt}(|\mathbf{r}|^2) = 2\mathbf{r} \cdot \mathbf{v} > 0$$

and so $|\mathbf{r}|$ is increasing in t . Similarly, if $\mathbf{r} \cdot \mathbf{v} < 0$, then $|\mathbf{r}|$ is decreasing in t .

Practice Problems

- Let C be the curve in the xy -plane that consists of the half-line $y = -x$, $x \leq 0$, and the half-line $y = x$, $x \geq 0$. Give an example of a vector-valued function $\mathbf{r} : \mathbb{R} \rightarrow \mathbb{R}^2$ which parametrizes C and such that $\mathbf{r}(t)$ is differentiable for all t .
- Suppose that $\mathbf{r} : [a, b] \rightarrow \mathbb{R}^3$ is a parametrized curve such that $\mathbf{r}$ and $\mathbf{r}'$ are continuous on $[a, b]$ and $\mathbf{r}'(t) \neq 0$. Does there have to exist a $t_0 \in (a, b)$ such that $\mathbf{r}(b) - \mathbf{r}(a) = (b - a)\mathbf{r}'(t_0)$? If yes, prove it. If no, give a counterexample.
- Let $A = (a_1, a_2)$ and $B = (b_1, b_2)$ be two points in $\mathbb{R}^2$ such that $|AB| = 1$. Let also $\mathbf{a} = \overrightarrow{OA} = a_1\mathbf{i} + a_2\mathbf{j}$.

(a) Find an orthogonal 2×2 matrix Q such that

- $\det(Q) = 1$,
- the mapping $L : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, defined by

$$L \begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{a} + Q \begin{pmatrix} x \\ y \end{pmatrix}$$

$$\text{satisfies } L \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \text{ and } L \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

(b) Let $\mathbf{r} : [0, 1] \rightarrow \mathbb{R}^2$ be given by $\mathbf{r}(t) = t\mathbf{i} + g(t)\mathbf{j}$, where $g : [0, 1] \rightarrow \mathbb{R}$ is continuous on $[0, 1]$ and satisfies $g(0) = g(1) = 0$. This parametrizes a curve $\mathcal{C}$.

Let L be as in (a). Then $L(\mathbf{r}(t))$ parametrizes the curve obtained from $\mathcal{C}$ by translating and rotating it so that the initial point is moved to A and the terminal point is moved to B . (This is a fact from linear algebra that you do not have to prove.)

Evaluate $L(\mathbf{r}(t))$ in terms of components. Prove that we may choose a parametrization $\mathbf{w} : [0, 1] \rightarrow \mathbb{R}^2$ of the line segment $\overline{AB}$, and a unit vector $\mathbf{n}$ perpendicular to $\overrightarrow{AB}$, so that

$$L(\mathbf{r}(t)) = \mathbf{w}(t) + g(t)\mathbf{n}.$$

(c) Let $\mathcal{C}$ be the curve in $\mathbb{R}^2$ parametrized by $\mathbf{r}(t) = t\mathbf{i} + \sin(\pi t)\mathbf{j}$, $t \in [0, 1]$. Let $\mathcal{C}_1$ be the curve in $\mathbb{R}^2$ obtained from $\mathcal{C}$ by translating and rotating it in the plane so that the initial point is moved to $(3, 4)$ and the terminal point is moved to $(\frac{7}{2}, 4 + \frac{\sqrt{3}}{2})$. Parametrize $\mathcal{C}_1$.

1.2 Arc Length for Curves

1.2.1 Definitions

In the rest of Chapter 1, we will assume that $\mathbf{r} : I \rightarrow \mathbb{R}^3$ is continuous unless explicitly stated otherwise. The interval I may be open, closed, finite or infinite as needed. We will also need additional terminology, as follows.

Definition 1.2.1. (i) A parametrization $\mathbf{r} : (a, b) \rightarrow \mathbb{R}^n$ of a curve is **smooth** if $\mathbf{r}(t)$ is C^1 and, additionally, $\mathbf{r}'(t) \neq \mathbf{0}$ for all $t \in (a, b)$.

(ii) For a closed interval $[a, b]$, a parametrization $\mathbf{r} : [a, b] \rightarrow \mathbb{R}^n$ of a curve is **smooth** if $\mathbf{r}$ is smooth on (a, b) , continuous on $[a, b]$, and the one-sided limits $\lim_{t \rightarrow a^+} \mathbf{r}'(t)$ and $\lim_{t \rightarrow b^-} \mathbf{r}'(t)$ exist.

(iii) A curve in $\mathbb{R}^n$ is **smooth** if it admits a smooth parametrization.

(iv) A curve in $\mathbb{R}^n$ is **piecewise smooth** if it consists of finitely many curves $\mathbf{r}_j : [a_j, b_j] \rightarrow \mathbb{R}^n$, $j = 1, 2, \dots, k$, such that each $\mathbf{r}_j$ is smooth on $[a_j, b_j]$, and the curves are joined to each other at endpoints so that $\mathbf{r}_j(b_j) = \mathbf{r}_{j+1}(a_{j+1})$ for $j = 1, \dots, k - 1$.

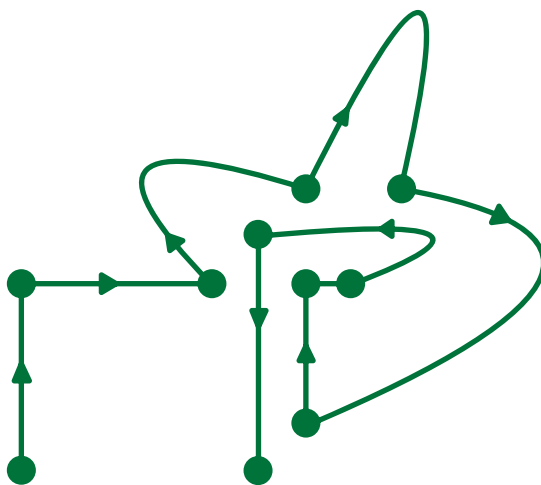


Figure 1.7: An example of a piecewise smooth curve.

This is purpose-specific terminology that is not universal in mathematics and does not always match terminology used in other contexts or in other textbooks. First, we

only require one derivative to exist and be continuous. (In mathematical analysis, for example, a “smooth” function is usually defined as a function whose all derivatives of all orders are continuous.) Second, we impose the additional requirement that $\mathbf{r}'$ should never be zero. This is because we want the underlying geometric object (in this case, the curve) to be smooth in an intuitive sense. Smooth geometric objects are defined here in terms of C^1 functions, but with further constraints added to prohibit geometric non-smoothness phenomena. The example below shows what could go wrong if we remove the condition that $\mathbf{r}'(t) \neq \mathbf{0}$ for a curve. We will see a similar issue later on in connection with parametrized surfaces.

Example 1.2.2. Let $\mathbf{u}$ and $\mathbf{w}$ be nonzero vectors in $\mathbb{R}^2$. For each $t \in \mathbb{R}$, define

$$\mathbf{r}(t) := \begin{cases} t^3 \mathbf{u}, & t \geq 0 \\ t^3 \mathbf{w}, & t \leq 0 \end{cases}$$

Then, the associated curve consists of two half-lines meeting at the origin. Moreover, $\mathbf{r}'(t)$ is defined for all t , with $\mathbf{r}'(0) = \mathbf{0}$ (see Practice Problems in Section 1.1), even though the curve is not geometrically “smooth” if $\mathbf{u}$ and $\mathbf{v}$ are not parallel.

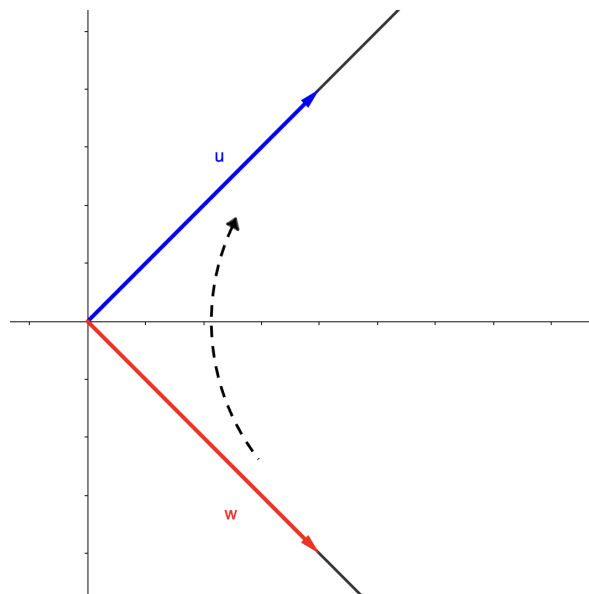


Figure 1.8: The vector-valued function defined in Example 1.2.2. In this picture, we have chosen $\mathbf{u} = \langle 1, 1 \rangle$, $\mathbf{w} = \langle 1, -1 \rangle$. The dashed arc indicates the direction of the parametrization.

If we think of $\mathbf{r}(t)$ as the position of a moving object at time t , the above example corresponds to the object moving in the direction of $\mathbf{w}$ for $t < 0$, slowing down as $t \nearrow 0$, coming to a full stop at $t = 0$, then starting out again in a new (possibly different) direction. In such cases, the trajectory may have a sharp kink even though

$\mathbf{r}(t)$ is still a differentiable function of t . The condition $\mathbf{r}'(0) \neq \mathbf{0}$ in Definition 1.2.1 rules out such examples.

We have already noted that the same curve can be parametrized in many different ways. A smooth curve such as a line will also have “hostile” parametrizations where $\mathbf{r}'(t)$ either does not exist at some points (the speed of the moving object changes instantaneously from one value to another) or takes zero values (the object comes to a full stop and then starts moving again). This is fine: for a curve to be smooth, just one smooth parametrization is enough.

Throughout Chapter 1, we will also assume that higher order derivatives of $\mathbf{r}(t)$ exist as needed. Curves that are not smooth (for example, fractal curves) exist and are important, but the methods used to study them are very different from those covered in this course.

Definition 1.2.3. The curve $\mathbf{r} : [a, b] \rightarrow \mathbb{R}^3$ is **closed** if $\mathbf{r}(a) = \mathbf{r}(b)$; it is **simple and closed** if it has no other self-intersections.

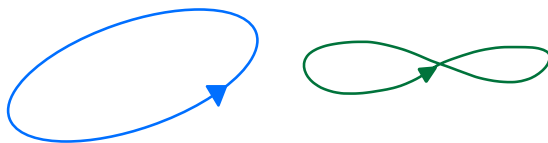


Figure 1.9: Two closed curves in the plane; the curve on the left is simple and closed, the one on the right is closed but not simple.

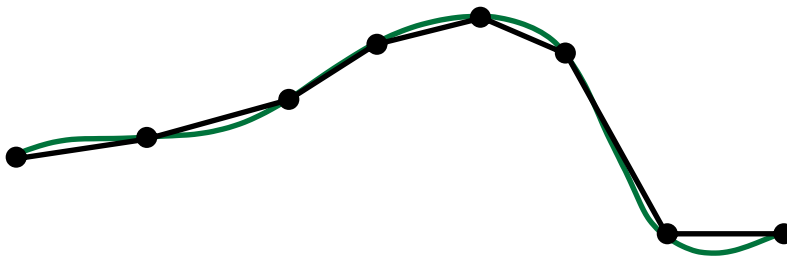
1.2.2 Arc Length of a Curve

Suppose that some curve $\mathcal{C} \subset \mathbb{R}^n$ is parametrized by a continuous vector-valued function $\mathbf{r}(t)$ for $a \leq t \leq b$. Further suppose that $\mathbf{r}$ is one-to-one so that, as $\mathbf{r}(t)$ travels from $\mathbf{r}(a)$ to $\mathbf{r}(b)$, it never covers the same segment of the curve more than once. Now, subdivide the interval $a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b$, and define $\mathbf{r}_i = \mathbf{r}(t_i)$ and let

$$S_{t_0, \dots, t_n} := \sum_{i=1}^n |\mathbf{r}_i - \mathbf{r}_{i+1}|.$$

Then, $S_{t_0, \dots, t_n}$ approximates the length of the curve $\mathcal{C}$ by the sum of the lengths of connected line segments whose endpoints lie on $\mathcal{C}$.

Moreover, as more points are added to the subdivision $t_0, \dots, t_n$, the associated length $S_{t_0, \dots, t_n}$ is non-decreasing. You can prove this using the *triangle inequality*: a line segment from P to Q directly is shorter than the triangular path from P to a third point R and then from R to Q .

Figure 1.10: A curve $\mathcal{C}$ approximated by line segments.

Definition 1.2.4. Let $\mathcal{C}$ be a curve parametrized by a continuous vector-valued function $\mathbf{r}(t)$ for $t \in I$, where I is an interval.

- (i) Assume that I is compact, so that $I = [a, b]$ for some finite $a, b \in \mathbb{R}$. We say that $\mathcal{C}$ is **rectifiable** if there is a constant $K > 0$ such that, with the above notation, we have $S_{t_0, \dots, t_n} \leq K$ for all n and for all possible choices of partitions. The **smallest such** K is called the **length** of $\mathcal{C}$.
- (ii) Assume that I is not compact (hence either unbounded or open at one or both ends). We say that $\mathcal{C}$ is rectifiable if, for any compact interval $[a, b] \subset I$, the curve obtained by restricting $\mathcal{C}$ to the values of $t \in [a, b]$ is rectifiable.

Many familiar curves such as a circle, ellipse, or line segment are rectifiable. There are also many curves of infinite length (line, parabola, helix, etc.) that are rectifiable because any bounded part of them is rectifiable. Fractal curves are often unrectifiable. The Koch snowflake curve (see Figure 1.11) provides one such example.

1.2.3 Computing the Length of the Curve

We have the following formula for computing the length of curves parametrized by continuously differentiable vector-valued functions.

Theorem 1.2.5. If $\mathbf{r}(t)$ has a continuous derivative $\mathbf{v}(t)$, then the length of the curve from $t = a$ to $t = b$ is

$$S = \int_a^b |\mathbf{r}'(t)| dt = \int_a^b |\mathbf{v}(t)| dt = \int_a^b v(t) dt.$$

Sketch of Proof. If we take some partition t_i and $\mathbf{r}_i$, let $\Delta \mathbf{r}_i := \mathbf{r}_i - \mathbf{r}_{i-1}$ so that

$$S_{t_0, \dots, t_n} = \sum_{i=1}^n |\Delta \mathbf{r}_i| \tag{1.2.1}$$

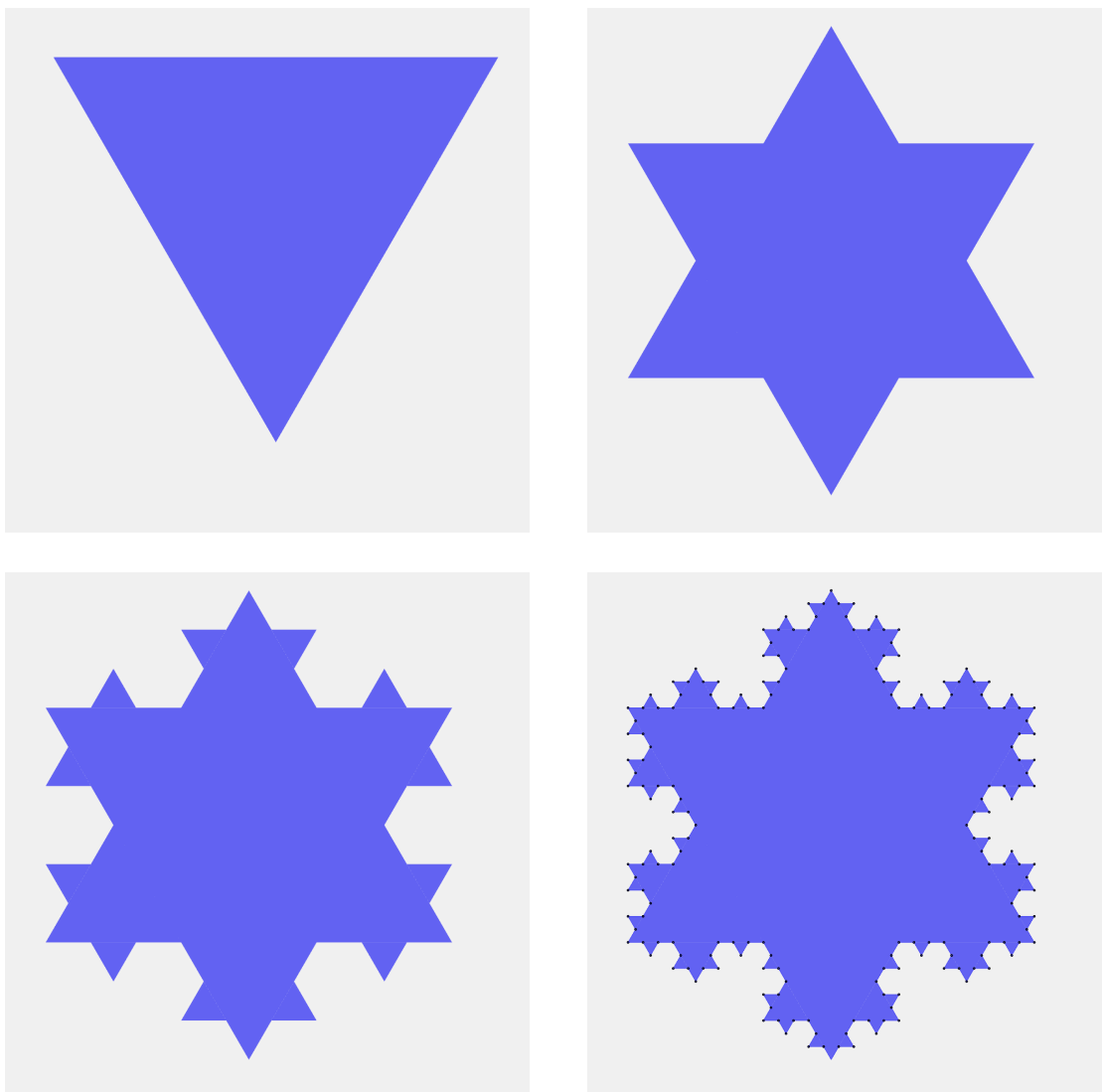


Figure 1.11: The Koch snowflake is an example of a *non-rectifiable* curve. This curve is the boundary of a region (pictured above) with finite area – and yet, the curve itself has “infinite” length. The Koch snowflake is constructed via an iterative process: a simple triangular pattern is repeated many times to create a more complicated pattern. The first four steps of this pattern-copying process are shown above. Tools from this class, as well as classes such as MATH 320 and MATH 321, will give you the tools to begin exploring more exotic geometric objects such as this curve!

It is tempting to write $\Delta \mathbf{r}_i = \mathbf{r}'(t_i^*) \Delta t_i$ for some $t_i^* \in (t_{i-1}, t_i)$ by the Mean Value Theorem, then take the limit of the Riemann sums. This does not quite work because there is no Mean Value Theorem for vector-valued functions (see Practice Problems in Section 1.1). One workaround is to write the difference $\Delta \mathbf{r}_i$ in terms of its component

functions:

$$\Delta \mathbf{r}_i = \Delta x_i \mathbf{i} + \Delta y_i \mathbf{j} + \Delta z_i \mathbf{k}.$$

Then equation (1.2.1) can be re-written as

$$\begin{aligned} S_{t_0, \dots, t_n} &= \sum_{i=1}^n \sqrt{|\Delta x_i|^2 + |\Delta y_i|^2 + |\Delta z_i|^2} \\ &= \sum_{i=1}^n \left[\sqrt{\left| \frac{\Delta x_i}{\Delta t_i} \right|^2 + \left| \frac{\Delta y_i}{\Delta t_i} \right|^2 + \left| \frac{\Delta z_i}{\Delta t_i} \right|^2} \right] \cdot \Delta t_i \\ &= \sum_{i=1}^n \left[\sqrt{|x'(t_i^{(1)})|^2 + |y'(t_i^{(2)})|^2 + |z'(t_i^{(3)})|^2} \right] \cdot \Delta t_i \end{aligned}$$

for some points $t_i^{(1)}, t_i^{(2)}$ and $t_i^{(3)}$ in $I_j := (t_{i-1}, t_i)$; at the last step, we applied the Mean Value Theorem to the (scalar) component functions. This is still not quite a Riemann sum since $t_i^{(1)}, t_i^{(2)}, t_i^{(3)}$ need not be equal. However, since x', y' and z' are continuous, this does not make much difference and we can still conclude (as in single-variable calculus) that the above converges to

$$\int_a^b \sqrt{|x'(t)|^2 + |y'(t)|^2 + |z'(t)|^2} dt = \int_a^b |\mathbf{r}'(t)| dt,$$

where the convergence is over refinements of partitions satisfying $\max_i \Delta t_i \rightarrow 0$.

□

1.2.4 Arc-Length Parametrization

Let $\mathbf{r} : [a, b] \rightarrow \mathbb{R}^3$ be smooth, and recall that $v(t) = |\mathbf{v}(t)|$. Define, for $t \in [a, b]$,

$$s(t) := \int_a^t |\mathbf{r}'(\tau)| d\tau = \int_a^t v(\tau) d\tau.$$

This is the length of the curve from $\mathbf{r}(a)$ to $\mathbf{r}(t)$. For this reason, we call $s(t)$ the **arc-length function**. By the Fundamental Theorem of Calculus, we have

$$\begin{aligned} \frac{ds}{dt} &= \frac{d}{dt} \left(\int_a^t v(\tau) d\tau \right) = v(t) \\ \Rightarrow ds &= v(t) dt = |\mathbf{r}'(t)| dt \end{aligned}$$

We therefore call the quantity $ds = |\mathbf{r}'(t)| dt$ the **arc-length element** associated to $\mathbf{r}$.

Now, suppose we can solve for t as a function of s ; then, we can use this change of variables to reparametrize the curve in terms of arc length. That is, we want to write

$\mathbf{r} = \mathbf{r}(t(s))$ as our new parametrization. The advantage of this parametrization is that

$$\left| \frac{d\mathbf{r}}{ds} \right| = \left| \frac{d\mathbf{r}}{dt} \cdot \frac{dt}{ds} \right| = \left| \frac{d\mathbf{r}}{dt} \right| \cdot \underbrace{\left| \frac{1}{ds/dt} \right|}_{(*)} = v(t) \cdot \frac{1}{v(t)} = 1, \quad (1.2.2)$$

where $(*)$ follows from the Implicit Function Theorem and our assumption that $\frac{ds}{dt} \neq 0$. In other words, the arc-length parametrization describes an object moving along the given curve with constant speed 1.

Remark 1.2.1. *Even if we cannot solve for t as a function of s explicitly, it is still useful to know that such parametrization exists. We will use it in what follows to define various natural quantities associated to smooth curves.*

Example 1.2.6. *A very important special case is when our curve $\mathcal{C}$ is the graph of some function $y = f(x)$ for $a \leq x \leq b$ in the plane. We may then parametrize our curve by the vector-valued function*

$$\mathbf{r}(t) := t\mathbf{i} + f(t)\mathbf{j}, \quad a \leq t \leq b,$$

so that the length is given by the formula

$$s(t) := \int_a^b |\mathbf{r}'(x)| dx = \int_a^b \sqrt{1 + |f'(x)|^2} dx$$

Worked Problems

1. Consider the helix

$$\mathbf{r}(t) := a \cos(t)\mathbf{i} + a \sin(t)\mathbf{j} + bt\mathbf{k}, \quad t \in \mathbb{R}.$$

where $a, b > 0$ are some real numbers.

(a) Find the length of part of the helix where $0 \leq t \leq T$ for some $T > 0$.

(b) Reparametrize the helix in terms of arc length.

(a) To find the length, we first need to determine $|\mathbf{r}'(t)|$. Since $\mathbf{r}'(t) = -a \sin t \mathbf{i} + a \cos t \mathbf{j} + b \mathbf{k}$, we have

$$|\mathbf{r}'(t)| = \sqrt{a^2 \sin^2 t + a^2 \cos^2 t + b^2} = \sqrt{a^2 + b^2},$$

so that

$$s(T) := \int_0^T \sqrt{a^2 + b^2} d\tau = (\sqrt{a^2 + b^2})T.$$

(b) From (a), we have $v(t) := |\mathbf{r}'(t)| = \sqrt{a^2 + b^2}$. Let us use $t_0 = 0$ as our starting point for measuring length, so that

$$s(t) := \int_0^t v(\tau) d\tau = \int_0^t \sqrt{a^2 + b^2} d\tau = (\sqrt{a^2 + b^2})t$$

So, we see that $t = t(s) = \frac{s}{\sqrt{a^2 + b^2}}$. Performing this change of variables gives us the reparametrization

$$\mathbf{r}(s) = a \cos\left(\frac{s}{\sqrt{a^2 + b^2}}\right) \mathbf{i} + a \sin\left(\frac{s}{\sqrt{a^2 + b^2}}\right) \mathbf{j} + b \left(\frac{s}{\sqrt{a^2 + b^2}}\right) \mathbf{k}.$$

2. A curve in $\mathbb{R}^3$ is given by

$$\mathbf{r}(t) := 2t\mathbf{i} + t^2\mathbf{j} + \frac{t^3}{3}\mathbf{k}. \quad (1.2.3)$$

Find the length of the curve from $t = 1$ to $t = 2$.

We have $\mathbf{r}'(t) = 2\mathbf{i} + 2t\mathbf{j} + t^2\mathbf{k}$. Therefore

$$v(t) := |\mathbf{r}'(t)| = \sqrt{4 + 4t^2 + t^4} = \sqrt{(2 + t^2)^2} = |t^2 + 2| = t^2 + 2,$$

since $t^2 + 2 \geq 0$. Hence, the length from $t = 1$ to $t = 2$ is given by

$$\int_1^2 (t^2 + 2) dt = \left[\frac{t^3}{3} + 2t \right]_{t=1}^2 = \frac{8}{3} + 4 - \frac{1}{3} - 2 = \frac{13}{3}.$$

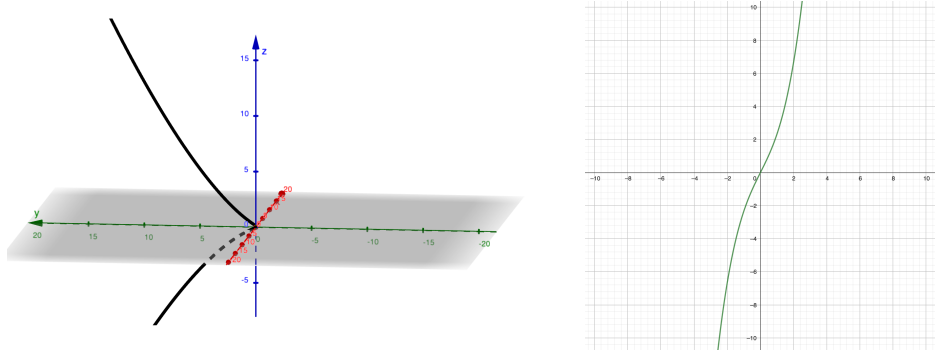


Figure 1.12: The curve in (1.2.3) and its associated arc-length function $s(t)$.

Remark 1.2.2. In the last example, we have $v(t) = t^2 + 2$, so that the arc-length function (measuring from $t = 0$ as a base point) is

$$s(t) := \int_0^t (\tau^2 + 2) d\tau = \frac{t^3}{3} + 2t.$$

The function $t \rightarrow s(t)$ is one-to-one (why?), therefore it indeed has an inverse function $t = t(s)$. However, to find such a function, we would need to solve the cubic equation appearing previously—a not so easy task. This is typical: in fact, many common arc-length functions cannot be expressed in terms of elementary functions. That will not stop us from working with them!

Practice Problems

1. Let C be the curve $x = e^{2t}$, $y = 2e^t$, $z = t$ for $t \in \mathbb{R}$. Find the length of the curve from $t = 0$ to $t = 1$.
2. Let $\mathbf{r}(t) = e^t \cos(2t)\mathbf{i} + e^t \sin(2t)\mathbf{j}$ be a parametrization of a curve C in $\mathbb{R}^2$.
 - (a) Let $s(b)$ be the length of the curve $\mathbf{r}(t)$ from $t = -\infty$ to $t = b$. (This is defined as the limit as $a \rightarrow -\infty$ of the length of $\mathbf{r}(t)$ from $t = a$ to $t = b$.) Find $s(b)$ as a function of b .
 - (b) Use part (a) to give an arc-length parametrization of C .
3. The paraboloid $z = x^2 + y^2 + 1$ intersects the surface $x^2 - 2x + y^2 = 0$ in a curve. Prove that the length of this curve is less than 15.

1.3 Unit Tangent and Curvature

This section seeks to answer the following.

Question 1.3.1. *How quickly does a given curve change direction?*

For curves, *direction* is represented by a unit tangent vector. (We use the *unit* tangent vector because we are only interested in its direction, not magnitude.) Moreover, we want to make sure that we have a geometric definition for the phrase “how quickly”. We say *geometric* because we want this definition to depend solely upon the shape of the curve, and not upon the choice of parametrization (which is a more analytic object). For this reason, we will use arc-length parametrization, so that motion along the curve has constant speed equal to one.

Remark 1.3.1. *Even when it may be difficult or impossible to calculate the arc-length parametrization, we can still use it to define new concepts. Eventually, we will have formulas for the same concepts that do not require us to know an explicit formula for arc length.*

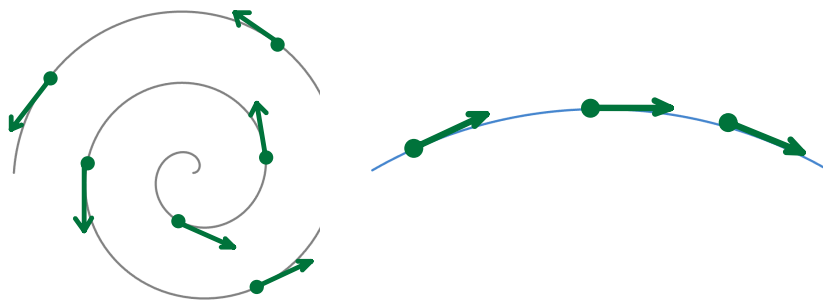


Figure 1.13: Two curves; the right-hand curve seems to be changing direction more slowly than the left-hand curve. The vector overlaid on the points of the curve is the tangent vector at that point.

1.3.1 Key Ideas and Definitions

We assume $\mathbf{r} : I \rightarrow \mathbb{R}^3$ is a smooth parametrization of a curve, for some interval I . In the sequel, whenever we work both with the initial parametrization $\mathbf{r}(t)$ and the arc-length reparametrization of the same curve, we will use $\mathbf{r}'$, $\mathbf{r}''$, etc. to denote derivatives with respect to t , and $d\mathbf{r}/ds$, $d^2\mathbf{r}/ds^2$, etc. to denote differentiation in s . In the rest of this chapter, we will always assume that all the needed derivatives exist and are continuous.

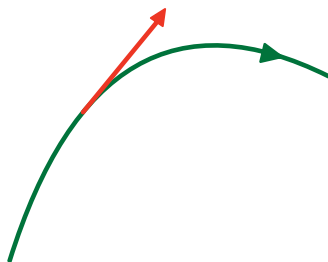


Figure 1.14: A tangent vector to some curve.

We define $\mathbf{T}$ to be the **unit tangent vector** to the curve. This is a geometric object depending (up to the choice of direction) only on the shape of the curve. In terms of the parametrization $\mathbf{r}$, we have

$$\mathbf{T}(t) := \frac{\mathbf{v}(t)}{|\mathbf{v}(t)|} = \frac{\mathbf{r}'(t)}{|\mathbf{r}'(t)|}, \text{ so that } |\mathbf{T}(t)| = 1, \quad (1.3.1)$$

with direction matching the direction of increasing t . If we use arc-length parametrization s , then $|\frac{d\mathbf{r}}{ds}| = 1$, so that

$$\mathbf{T}(s) = \frac{d\mathbf{r}/ds}{|d\mathbf{r}/ds|} = \frac{d\mathbf{r}}{ds}.$$

The **curvature** $\kappa := \kappa(s)$ is defined relative to the arc-length parametrization as

$$\kappa(s) := \left| \frac{d\mathbf{T}(s)}{ds} \right|,$$

assuming that the derivative exists. In words, $\kappa(s)$ measures the rate of change of $\mathbf{T}$ if we move along the curve with unit speed. If $\mathbf{r}$ is any parametrization of the curve, we may use the chain rule to write $\kappa(s) = \kappa(s(t))$ as

$$\begin{aligned} \kappa(t) &= \left| \frac{d\mathbf{T}}{ds} \right| = \left| \frac{d\mathbf{T}}{dt} \cdot \frac{dt}{ds} \right| = \frac{|d\mathbf{T}/dt|}{|ds/dt|} = \frac{|\mathbf{T}'(t)|}{|s'(t)|} \\ &= \frac{|\mathbf{T}'(t)|}{|\mathbf{r}'(t)|}. \end{aligned}$$

1.3.2 Examples, and a Useful Formula for Curvature

Example 1.3.2. Consider the straight line $\mathbf{r}(t) := \mathbf{r}_0 + t\mathbf{v}$. Then,

$$\mathbf{T} := \frac{\mathbf{v}}{|\mathbf{v}|} \text{ is constant,}$$

and so $\kappa = \left| \frac{d\mathbf{T}}{ds} \right| = 0$.

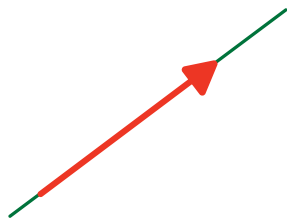


Figure 1.15: A line segment in the direction of $\mathbf{v}$.

Example 1.3.3. Consider the circle $\mathbf{r}(t) = a \cos(t)\mathbf{i} + a \sin(t)\mathbf{j}$ for $a > 0$. To compute the tangent vector $\mathbf{T}$, we first find

$$\mathbf{r}'(t) := -a \sin(t)\mathbf{i} + a \cos(t)\mathbf{j}, \quad |\mathbf{r}'(t)| = a,$$

so that

$$\mathbf{T}(t) = \frac{\mathbf{r}'(t)}{|\mathbf{r}'(t)|} = -\sin(t)\mathbf{i} + \cos(t)\mathbf{j}.$$

To compute the curvature, we need

$$\mathbf{T}'(t) = -\cos(t)\mathbf{i} - \sin(t)\mathbf{j}, \quad |\mathbf{T}'(t)| = 1,$$

so that

$$\kappa = \frac{|\mathbf{T}'(t)|}{|\mathbf{r}'(t)|} = \frac{1}{a}.$$

In particular, we see that as $a \rightarrow 0$, we must have $\kappa \rightarrow \infty$, so that large values of curvature correspond to small values of radius $a > 0$.

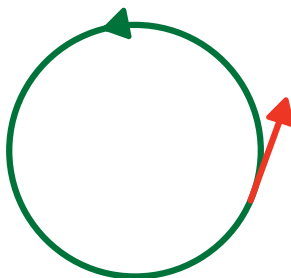


Figure 1.16: A circle with associated tangent vector.

Remark 1.3.2. For general curves, the quantity $\rho := \frac{1}{\kappa}$ is called the **radius of curvature**. It is the radius of the circle (in the three-dimensional space) that best approximates the curve at the given point – more on this later. If the curve is an actual circle, then $\rho = a$ is its radius.

Example 1.3.4. Take the helix $\mathbf{r}(t) = a \cos(t)\mathbf{i} + a \sin(t)\mathbf{j} + bt\mathbf{k}$ where both $a, b > 0$. Then,

$$\mathbf{r}'(t) = -a \sin(t)\mathbf{i} + a \cos(t)\mathbf{j} + b\mathbf{k}, \quad |\mathbf{r}'(t)| = \sqrt{a^2 + b^2}.$$

so that

$$\begin{aligned} \mathbf{T}(t) &:= \frac{\mathbf{r}'(t)}{|\mathbf{r}'(t)|} = -\frac{a}{\sqrt{a^2 + b^2}} \sin(t)\mathbf{i} + \frac{a}{\sqrt{a^2 + b^2}} \cos(t)\mathbf{j} + \frac{b}{\sqrt{a^2 + b^2}} \mathbf{k}, \\ \mathbf{T}'(t) &:= -\frac{a}{\sqrt{a^2 + b^2}} \cos(t)\mathbf{i} - \frac{a}{\sqrt{a^2 + b^2}} \sin(t)\mathbf{j}. \end{aligned}$$

Therefore

$$\kappa(t) := \frac{|\mathbf{T}'(t)|}{|\mathbf{r}'(t)|} = \frac{\frac{a}{\sqrt{a^2 + b^2}}}{\sqrt{a^2 + b^2}} = \frac{a}{a^2 + b^2}$$

and so our curvature is constant.

Note that if either a is fixed and $b \rightarrow \infty$ (the helix gets stretched out in the z -direction), or if $a \rightarrow \infty$ as b is fixed (the vertical speed is constant but the radius of the cylinder around which the helix is wrapped increases to infinity), we have $\kappa \rightarrow 0$.

We conclude this section with a formula for computing curvature which is often more practical than the formula given by the definition. As mentioned above, we use primes to denote differentiation in t , for example $\mathbf{T}' = \frac{d\mathbf{T}}{dt}$. Differentiation in s will always be written as $\frac{d}{ds}$.

Theorem 1.3.5. *If $\mathbf{r}$ is twice differentiable and $\mathbf{r}' \neq 0$, then*

$$\kappa(t) = \frac{|\mathbf{r}'(t) \times \mathbf{r}''(t)|}{|\mathbf{r}'(t)|^3} \quad (1.3.2)$$

Proof. We begin by computing the cross product $\mathbf{r}' \times \mathbf{r}''$ in terms of $\mathbf{T}$ and the arc-length parametrization. Recall that $\mathbf{T} := \frac{\mathbf{r}'(t)}{|\mathbf{r}'(t)|}$ and $|\mathbf{r}'(t)| = \frac{ds}{dt}$. Therefore

$$\mathbf{r}'(t) = |\mathbf{r}'(t)|\mathbf{T}(t) = \frac{ds}{dt}(t)\mathbf{T}(t). \quad (1.3.3)$$

To find $\mathbf{r}''(t)$, we differentiate the above with respect to t , obtaining

$$\mathbf{r}'' = \frac{d}{dt} \left(\frac{ds}{dt} \mathbf{T} \right) = \frac{d^2s}{dt^2} \mathbf{T} + \frac{ds}{dt} \mathbf{T}'. \quad (1.3.4)$$

Using (1.3.3) and (1.3.4), we find that

$$\begin{aligned} \mathbf{r}' \times \mathbf{r}'' &= \left(\frac{ds}{dt} \mathbf{T} \right) \times \left(\frac{d^2s}{dt^2} \mathbf{T} + \frac{ds}{dt} \mathbf{T}' \right) \\ &= \frac{ds}{dt} \frac{d^2s}{dt^2} (\mathbf{T} \times \mathbf{T}) + \left(\frac{ds}{dt} \right)^2 \mathbf{T} \times \mathbf{T}' \end{aligned}$$

From the definition of the cross product, we know that $\mathbf{T} \times \mathbf{T} = 0$. Therefore

$$\mathbf{r}' \times \mathbf{r}'' = \left(\frac{ds}{dt} \right)^2 \mathbf{T} \times \mathbf{T}' \quad (1.3.5)$$

Now, we know that $|\mathbf{T}| = 1$ for all values of t ; this further tells us that $\mathbf{T} \perp \mathbf{T}'$ for all values of t (Worked Problems, Section 1.1). Consequently, we see that

$$|\mathbf{T} \times \mathbf{T}'| = |\mathbf{T}| \cdot |\mathbf{T}'| \sin\left(\frac{\pi}{2}\right) = |\mathbf{T}'|. \quad (1.3.6)$$

Putting together equations (1.3.5) and (1.3.6) shows that

$$|\mathbf{r}' \times \mathbf{r}''| = \left(\frac{ds}{dt} \right)^2 |\mathbf{T}'| = |\mathbf{r}'|^2 |\mathbf{T}'| \Rightarrow |\mathbf{T}'| = \frac{|\mathbf{r}' \times \mathbf{r}''|}{|\mathbf{r}'|^2}.$$

The definition of curvature then gives that

$$\kappa := \frac{|\mathbf{T}'|}{|\mathbf{r}'|} = \frac{|\mathbf{r}' \times \mathbf{r}''|/|\mathbf{r}'|^2}{|\mathbf{r}'|} = \frac{|\mathbf{r}' \times \mathbf{r}''|}{|\mathbf{r}'|^3}.$$

□

Worked Problems

1. Compute the curvature $\kappa(t)$ if $\mathbf{r}(t) = \langle 2t, t^2, \frac{t^3}{3} \rangle$ using Theorem 1.3.5.

We first calculate the following

$$\begin{aligned}\mathbf{r}'(t) &= \langle 2, 2t, t^2 \rangle, & |\mathbf{r}'(t)| &= \sqrt{4 + 4t^2 + t^4} = t^2 + 2. \\ \mathbf{r}''(t) &= \langle 0, 2, 2t \rangle\end{aligned}$$

The associated cross product is then given by:

$$\begin{aligned}\mathbf{r}' \times \mathbf{r}'' &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2 & 2t & t^2 \\ 0 & 2 & 2t \end{vmatrix} = (4t^2 - 2t^2)\mathbf{i} - 4t\mathbf{j} + 4\mathbf{k} \\ &= 2t^2\mathbf{i} - 4t\mathbf{j} + 4\mathbf{k},\end{aligned}$$

which has length

$$|\mathbf{r}'(t) \times \mathbf{r}''(t)| = 2\sqrt{t^4 + 4t^2 + 4} = 2(t^2 + 2).$$

The curvature is then given by the formula

$$\kappa(t) = \frac{2(t^2 + 2)}{(t^2 + 3)^3} = \frac{2}{(t^2 + 2)^2}$$

2. Let $\mathcal{C}$ be the curve at the intersection of the cylinder $x^2 + y^2 = 1$ in $\mathbb{R}^3$ with the plane $z = ax$ for some $a > 0$. Parametrize the curve, and find its curvature as a function of the parameter.

The curve can be parametrized as $\mathbf{r}(t) = \cos t\mathbf{i} + \sin t\mathbf{j} + a \cos t\mathbf{k}$, $0 \leq t \leq 2\pi$. Then

$$\begin{aligned}\mathbf{r}'(t) &= -\sin t\mathbf{i} + \cos t\mathbf{j} - a \sin t\mathbf{k} \\ \mathbf{r}''(t) &= -\cos t\mathbf{i} - \sin t\mathbf{j} - a \cos t\mathbf{k} \\ |\mathbf{r}'| &= \sqrt{\sin^2 t + \cos^2 t + a^2 \sin^2 t} = \sqrt{1 + a^2 \sin^2 t} \\ \mathbf{r}'(t) \times \mathbf{r}''(t) &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -\sin t & \cos t & -a \sin t \\ -\cos t & -\sin t & -a \cos t \end{vmatrix} = (-a \cos^2 t - a \sin^2 t)\mathbf{i} \\ &\quad - (a \cos t \sin t - a \sin t \cos t)\mathbf{j} + (\sin^2 t + \cos^2 t)\mathbf{k} = -a\mathbf{i} + \mathbf{k}\end{aligned}$$

$$\begin{aligned}|\mathbf{r}'(t) \times \mathbf{r}''(t)| &= \sqrt{a^2 + 1} \\ \kappa &= \frac{|\mathbf{r}'(t) \times \mathbf{r}''(t)|}{|\mathbf{r}'|^3} = \frac{\sqrt{a^2 + 1}}{(1 + a^2 \sin^2 t)^{3/2}}\end{aligned}$$

Practice Problems

- Let $\mathcal{C}$ be the curve $x = \frac{t^3}{3} - 2t$, $y = \frac{t^3}{3} + 2t$, $z = \sqrt{2}t^2$ for $t \in \mathbb{R}$.
 - Find the arc length of the curve from $t = 0$ to $t = 1$.
 - Find the curvature at the point where $t = 0$.
- Let $\mathbf{r}(t) = t\mathbf{i} + \frac{t^2}{2}\mathbf{j}$, $t \in \mathbb{R}$ (a parabola in $\mathbb{R}^2$), and let $\mathbf{T}(t)$ be the unit tangent vector to $\mathbf{r}(t)$. Suppose that $\mathbf{r}(u)$, $u \in I$ (for some interval I), is a smooth reparametrization of the parabola such that

$$\left| \frac{d\mathbf{T}}{du} \right| = 1.$$

Find $\frac{du}{dt}$ as a function of t . Prove that the interval I is finite, and find the length of it.

- Let $\mathbf{r} : \mathbb{R} \rightarrow \mathbb{R}^3$ be given by

$$\mathbf{r}(t) = \cos(t)\mathbf{i} + \sin(t)\mathbf{j} + f(t)\mathbf{k},$$

where $f : \mathbb{R} \rightarrow \mathbb{R}$ satisfies $f(t) = f(t + 2\pi)$. Assume that f and all its derivatives are continuous. Thus $\mathbf{r}(t)$ parametrizes a closed smooth curve that lies on the cylinder $x^2 + y^2 = 1$, with each t -interval of length 2π covering the entire curve. Prove that there is a point $\mathbf{r}(t_0)$ on the curve where $\kappa \leq 1$.

- Let $\mathbf{r}(t)$ be a smooth vector-valued function, and suppose that there is a point P that lies on every normal plane to $\mathbf{r}(t)$. Prove that the curve lies on the surface of a sphere.

1.4 The Principal Normal and The Osculating Plane

In this section, we introduce many new vector-valued functions and definitions associated to a curve $\mathbf{r}(t)$ in three-dimensional space. Each of these will have a natural geometric interpretation, motivated by the following general question.

Question 1.4.1. *Given a curve in $\mathbb{R}^3$, how can we measure the “three-dimensionality” of the that curve?*

We will see that the **osculating plane** will be the plane of “best fit” for the curve at any given point, whereas a quantity called **torsion** will measure how far the curve is from fitting into said plane. These two objects give an answer to Question 1.4.1. Here and in the rest of this chapter, we will assume that $\mathbf{r}$ has as many derivatives as we need.

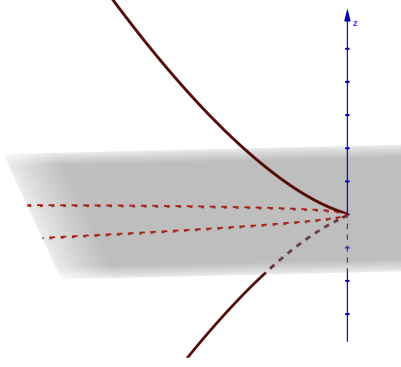


Figure 1.17: A curve in $\mathbb{R}^3$, together with its osculating plane (given by the equation $z = 0$) and its projection (i.e. *shadow*) onto that plane.

1.4.1 Unit Principal Normal

Recall that the unit tangent vector is given by $\mathbf{T} = \frac{d\mathbf{r}}{ds}$ in the arclength parametrization. It has the properties

$$\mathbf{T} = \frac{\mathbf{r}'}{|\mathbf{r}'|}, \quad |\mathbf{T}| = 1.$$

Assuming that the curvature $\kappa = |\frac{d\mathbf{T}}{ds}|$ is nonzero, we define

$$\mathbf{N}(s) := \frac{d\mathbf{T}/ds}{|d\mathbf{T}/ds|} \quad (\text{Normal Vector}). \quad (1.4.1)$$

This is a new vector-valued function, normalized so that $\mathbf{N}(s)$ is always a unit vector. Geometrically, $\mathbf{N}$ indicates the direction in which $\mathbf{T}(s)$ is bending as we travel along the curve. We call $\mathbf{N}$ the **unit principal normal vector** or just the **principal normal** to $\mathbf{T}$.

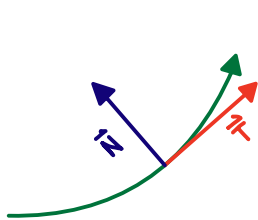
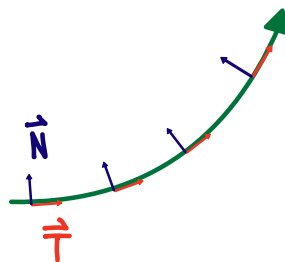
Proposition 1.4.2. *We always have $\mathbf{N} \perp \mathbf{T}$.*

Proof. We know that $|\mathbf{T}(t)| = 1$ for all t , and we proved earlier (Section 1.1) that this implies $\mathbf{T}(t) \perp \mathbf{T}'(t)$. Hence, by (1.4.1), we have $\mathbf{N} \perp \mathbf{T}$. $\square$

In order to actually compute $\mathbf{N}$, it is typically easier to use the original $\mathbf{r}(t)$ parametrization rather than arc length. Our formula for this is

$$\mathbf{N}(t) = \frac{d\mathbf{T}/ds}{|d\mathbf{T}/ds|} = \frac{\frac{d\mathbf{T}}{dt} \cdot \frac{dt}{ds}}{|\frac{d\mathbf{T}}{dt} \cdot \frac{dt}{ds}|} = \frac{\mathbf{T}'(t)}{|\mathbf{T}'(t)|}.$$

Notice that, since $\frac{ds}{dt} := |\mathbf{r}'(t)| > 0$, we always have $\frac{dt}{ds} = |\frac{dt}{ds}|$.

Figure 1.18: $\mathbf{T}$ and $\mathbf{N}$ at a pointFigure 1.19: $\mathbf{T}$ and $\mathbf{N}$ along the curve

1.4.2 The Osculating Plane and Circle

Equipped with our definitions of the tangent vector $\mathbf{T}(t)$ and the normal vector $\mathbf{N}(t)$ to a curve $\mathbf{r}(t)$, we define natural geometric objects to answer Question 1.4.1.

The **osculating plane** at a point $\mathbf{r}(t)$ is the plane passing through $\mathbf{r}(t)$ and parallel to the vectors $\mathbf{T}$ and $\mathbf{N}$. This is the plane that offers the “best fit” (among planes) for the curve $\mathbf{r}$ at the given point, in the following sense. Assume that $\kappa(t) \neq 0$, and let t_1, t_2, t_3 be distinct and close enough to t . It can then be proved that $\mathbf{r}(t_1), \mathbf{r}(t_2), \mathbf{r}(t_3)$ are not collinear, so that there is a unique plane Π passing through all three points. Further, if we take the limit as $t_1, t_2, t_3 \rightarrow t$, then Π converges to the osculating plane at $\mathbf{r}(t)$. See the Practice Problems for a specific example of this.

Given a parametrization $\mathbf{r}(t)$, we can find the equation of the osculating plane using the formula (1.3.5). We recall that formula here:

$$\mathbf{r}' \times \mathbf{r}'' = \left(\frac{ds}{dt}\right)^2 \mathbf{T} \times \mathbf{T}'.$$

The osculating plane, by the definition above, should be perpendicular to both $\mathbf{T}$ and $\mathbf{N} = \frac{\mathbf{T}'}{|\mathbf{T}'|}$. Hence the normal to it should be parallel to the cross product $\mathbf{T} \times \mathbf{T}'$, therefore also to $\mathbf{r}' \times \mathbf{r}''$. This offers a way to find the equation of the osculating plane directly by computing $\mathbf{r}' \times \mathbf{r}''$ and using it as a normal vector. We also note the physical interpretation of this: at any given point on the curve, the osculating plane is parallel both to the velocity $\mathbf{r}'$ and the acceleration $\mathbf{r}''$. We will have more to say about this in Section 1.6.

Within the osculating plane spanned by $\mathbf{T}(t)$ and $\mathbf{N}(t)$, we also construct the **osculating circle**. The osculating circle is tangent to the curve at the given point $\mathbf{r}(t)$, with radius $\rho = \frac{1}{\kappa}$ and center at $\mathbf{r}(t) + \rho\mathbf{N}(t)$.

The osculating circle has the same tangent vector and osculating plane as the curve $\mathbf{r}(t)$. It also has the same *curvature* as $\mathbf{r}(t)$ for the given value of t (see Worked Problem 1 below). Geometrically, it is the circle that gives the best approximation (among all circles) to the curve at the given point.

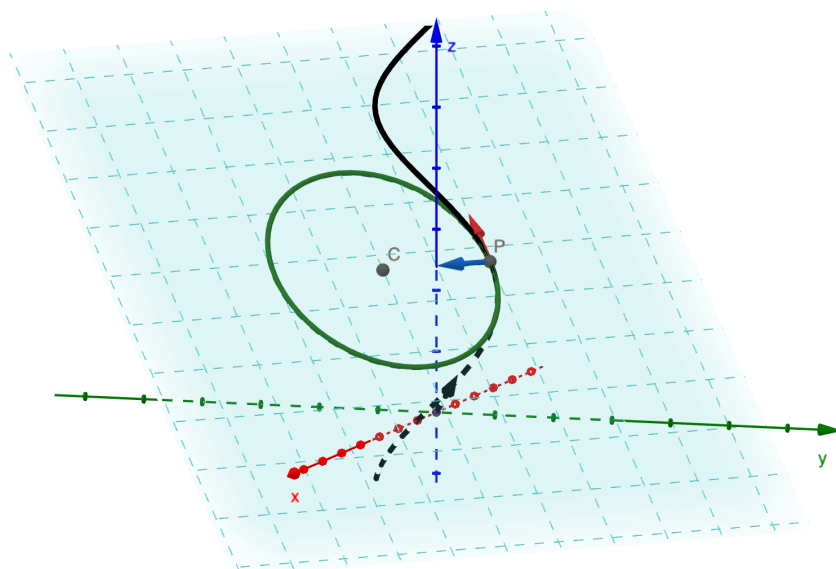


Figure 1.20: The osculating plane and osculating circle for the helix. The red vector is the tangent vector and the blue is the normal vector. See Worked Problem 2 for the calculation.

Worked Problems

1. Calculate $\mathbf{T}(t)$, $\mathbf{N}(t)$, and the osculating plane of the circle $\mathbf{r}(t) = a \cos(t)\mathbf{i} + a \sin(t)\mathbf{j}$. Then, show that $\mathbf{r}(t)$ is its own osculating circle.

First, we calculate $\mathbf{T}(t)$ and $\mathbf{N}(t)$ for arbitrary t . From an earlier example, the unit tangent vector is

$$\mathbf{T}(t) = -\sin(t)\mathbf{i} + \cos(t)\mathbf{j}.$$

Thus,

$$\mathbf{T}'(t) = -\cos(t)\mathbf{i} - \sin(t)\mathbf{j} \Rightarrow |\mathbf{T}'(t)| = 1.$$

Therefore, the principal unit normal vector at time t is given by the formula

$$\mathbf{N}(t) = \frac{\mathbf{T}'(t)}{|\mathbf{T}'(t)|} = -\cos(t)\mathbf{i} - \sin(t)\mathbf{j} = -\frac{1}{a}\mathbf{r}(t). \quad (1.4.2)$$

In particular, we see that $\mathbf{N}$ is a unit vector directed towards the center of the circle $\mathbf{r}(t)$.

The osculating plane spanned by $\mathbf{T}$ and $\mathbf{N}$ is given by the xy -plane. This follows since the vectors $\langle -\sin(t), \cos(t) \rangle$ and $-\langle \cos(t), \sin(t) \rangle$ form a basis for $\mathbb{R}^2$. Moreover, since the curvature of the circle is given by $\kappa = \frac{1}{a}$, we know that $\rho = a$ (so that

the osculating circle has radius equal to the circle $\mathbf{r}(t)$). We also see from (1.4.2) that the center of the osculating circle is at $\mathbf{r}(t) + \rho\mathbf{N}(t) = \mathbf{0}$. Since the osculating circle has the same center and radius as $\mathbf{r}(t)$, and lies in the same plane, this proves that $\mathbf{r}$ is its own osculating circle.

2. Find $\mathbf{T}(t)$, $\mathbf{N}(t)$, and the center and radius of the osculating circle at a general point $\mathbf{r}(t)$ of the helix

$$\mathbf{r}(t) = a \cos(t)\mathbf{i} + a \sin(t)\mathbf{j} + bt\mathbf{k}, \quad a > 0, b \neq 0,$$

*Does the helix have any points where the osculating plane is parallel to the xy -plane?
Does the center of the osculating circle lie on the z -axis?*

From Section 1.3, we already know that

$$\begin{aligned} \mathbf{T}(t) &= -\frac{a}{\sqrt{a^2 + b^2}} \sin(t)\mathbf{i} + \frac{a}{\sqrt{a^2 + b^2}} \cos(t)\mathbf{j} + \frac{b}{\sqrt{a^2 + b^2}} \mathbf{k} \\ \mathbf{T}'(t) &= -\frac{a}{\sqrt{a^2 + b^2}} \cos(t)\mathbf{i} - \frac{a}{\sqrt{a^2 + b^2}} \sin(t)\mathbf{j} \\ |\mathbf{T}'(t)| &= \frac{a}{\sqrt{a^2 + b^2}}. \\ \kappa(t) &= \frac{a}{a^2 + b^2} \end{aligned}$$

Therefore, the unit principal normal vector is

$$\mathbf{N}(t) = \frac{\mathbf{T}'(t)}{|\mathbf{T}'(t)|} = -\cos(t)\mathbf{i} - \sin(t)\mathbf{j},$$

the same as for the circle (Worked Problem 1 above). The radius of curvature is given by

$$\rho = \frac{1}{\kappa} = \frac{a^2 + b^2}{a},$$

and the center of the osculating circle is located at

$$\begin{aligned} \mathbf{r}(t) + \rho\mathbf{N}(t) &= a \cos(t)\mathbf{i} + a \sin(t)\mathbf{j} + bt\mathbf{k} + \frac{a^2 + b^2}{a}(-\cos(t)\mathbf{i} - \sin(t)\mathbf{j}) \\ &= -\frac{b^2}{a} \cos(t)\mathbf{i} - \frac{b^2}{a} \sin(t)\mathbf{j} + bt\mathbf{k}. \end{aligned}$$

In particular, the center of the osculating circle *does not* lie on the z -axis unless $b = 0$. Further, because of the factor of $\frac{b}{\sqrt{a^2 + b^2}}$ appearing on the $\mathbf{k}$ term in $\mathbf{T}$, the osculating plane is never parallel to the xy -plane.

3. If $\mathbf{r}(t) = t\mathbf{i} + at^2\mathbf{j}$ for some $a > 0$ (a parametrization of the parabola $y = ax^2$ in $\mathbb{R}^2$), find $\mathbf{T}(t)$, $\mathbf{N}(t)$ and $\kappa(t)$ at a general point of the curve.

We have $\mathbf{r}'(t) = \mathbf{i} + 2at\mathbf{j}$, $|\mathbf{r}'(t)| = \sqrt{1 + 4a^2t^2}$, and

$$\begin{aligned}\mathbf{T}(t) &= (1 + 4a^2t^2)^{-1/2} \mathbf{i} + 2at(1 + 4a^2t^2)^{-1/2} \mathbf{j}, \\ \mathbf{T}'(t) &= -\frac{1}{2}(1 + 4a^2t^2)^{-3/2}(8a^2t)\mathbf{i} + \left[2a(1 + 4a^2t^2)^{-1/2} - 2at\frac{8a^2t}{2(1 + 4a^2t^2)^{3/2}}\right] \mathbf{j} \\ &= -\frac{4a^2t}{(1 + 4a^2t^2)^{3/2}} \mathbf{i} + \frac{2a(1 + 4a^2t^2) - 8a^3t^2}{(1 + 4a^2t^2)^{3/2}} \mathbf{j} \\ &= -\frac{4a^2t}{(1 + 4a^2t^2)^{3/2}} \mathbf{i} + \frac{2a}{(1 + 4a^2t^2)^{3/2}} \mathbf{j}, \\ |\mathbf{T}'(t)| &= \frac{2a}{(1 + 4a^2t^2)^{3/2}} \sqrt{4a^2t^2 + 1} = \frac{2a}{1 + 4a^2t^2}, \\ \mathbf{N}(t) &= \frac{\mathbf{T}'(t)}{|\mathbf{T}'(t)|} = -\frac{2at}{\sqrt{1 + 4a^2t^2}} \mathbf{i} + \frac{1}{\sqrt{1 + 4a^2t^2}} \mathbf{j}, \\ \kappa(t) &= \frac{|\mathbf{T}'(t)|}{|\mathbf{r}'(t)|} = \frac{2a/(1 + 4a^2t^2)}{\sqrt{1 + 4a^2t^2}} = \frac{2a}{(1 + 4a^2t^2)^{3/2}}.\end{aligned}$$

Practice Problems

1. Let

$$\mathbf{r}(t) = \langle \sqrt{2}t^{3/2}, (t+2)^{3/2}, (t-2)^{3/2} \rangle, \text{ for } t \in (2, \infty).$$

Find the unit tangent vector $\mathbf{T}(t)$, the principal normal vector $\mathbf{N}(t)$, and the curvature $\kappa(t)$. Find the limit of $\mathbf{N}(t)$ as $t \rightarrow 2^+$. (No need for an $\epsilon - \delta$ proof here, it suffices to use the general rules for limits.)

(The calculations in this question, as well as in Worked Problem 3 above, are a little bit messy. Later on, we will see another way of finding $\mathbf{N}$ that leads to simpler calculations in many cases.)

2. Let $\mathbf{r}(t) = t\mathbf{i} + t^2\mathbf{j} + t^3\mathbf{k}$ for $t \in \mathbb{R}$.

(a) Find the equation of the osculating plane at the point $t = 0$.

(b) For $s, t \in \mathbb{R}$ such that $s, t, 0$ are all distinct, find the unit vector $\mathbf{n}(t, s)$ normal to the plane that passes through the points $\mathbf{r}(t)$, $\mathbf{r}(s)$, $\mathbf{r}(0)$. (There are two such vectors, pointing in opposite directions; choose the one that has a positive $\mathbf{k}$ -component.) Prove that

$$\lim_{s, t \rightarrow 0, s \neq t} \mathbf{n}(t, s)$$

is perpendicular to the osculating plane found in (a).

3. Let $\mathbf{r}(t) = t\mathbf{i} + at^2\mathbf{j}$ for some $a > 0$.

(a) Find the center Q and radius ρ of the osculating circle at the point $(0, 0)$.

(b) For $t > 0$, find the center $Q(t)$ and radius $R(t)$ of the circle in the xy -plane that passes through the points $\mathbf{r}(t)$, $\mathbf{r}(-t)$, $\mathbf{r}(0)$. Prove that $Q(t) \rightarrow Q$ and $R(t) \rightarrow \rho$ as $t \rightarrow 0$, where Q and ρ are the answers in (a).)

4. Let $\mathbf{r}$ be a parametrized curve in $\mathbb{R}^3$ such that $\mathbf{r}$ and its derivatives of all orders are continuous, $\mathbf{r}'(t) \neq 0$ for t in some interval I , and $|\mathbf{r}(t)| = r$ for all $t \in I$ (i.e. the curve lies on the sphere S of radius r centered at the origin O). Let $P = \mathbf{r}(t)$ be a point on the curve.

(a) Prove that $\rho = r \cos \theta$, where θ is the angle between $\vec{PO}$ and $\mathbf{N}$. Use this to prove that $\kappa \geq 1/r$.

(Hint: Use arclength parametrization. We have $\mathbf{r}(s) \cdot \frac{d\mathbf{r}}{ds} = 0$ (why?). If we differentiate this in s and use that $|\frac{d\mathbf{r}}{ds}| = 1$, what does it say about the dot product $\vec{PO} \cdot \mathbf{N}$?)

(b) Prove that for any point $P = \mathbf{r}(t)$ on the curve, the osculating circle for the curve at that point lies on the sphere S .

5. Let $\mathbf{r}$ be a closed, simple parametrized curve in $\mathbb{R}^3$. We use a “wrap-around” parametrization: if $\mathbf{r}(t)$ for $t \in [0, a]$ parametrizes one full loop around C returning to the starting point, we extend $\mathbf{r}$ to a periodic function $\mathbf{r} : \mathbb{R} \rightarrow \mathbb{R}^3$ so that $\mathbf{r}(t + a) = \mathbf{r}(t)$ for all t . (This is similar to parametrizing the unit circle by $\mathbf{r}(t) = \cos(t)\mathbf{i} + \sin(t)\mathbf{j}$, with period 2π .)

Assume that $\mathbf{r}$ and its derivatives of all orders are continuous, $\mathbf{r}'(t) \neq 0$, and $|\mathbf{r}(t)| \leq 1$ for all t (i.e. the curve lies in the closed ball of radius 1 centered at the origin O). Prove that there is a point $\mathbf{r}(t_0)$ on the curve where $\kappa(t_0) \geq 1$.

(Hint: Let $f(t) = |\mathbf{r}(t)|^2$, then $f(t)$ attains its maximum at some point $t = t_0$. What does this say about $f'(t_0)$ and $f''(t_0)$? It may be helpful to use arclength parametrization, starting from $t = t_0$ so that $s(t_0) = 0$.)

1.5 Torsion

In this section, we define the **Frenet Frame**, an orthonormal basis for $\mathbb{R}^3$ that is intrinsic to a smooth curve at a given point and changes as we move along that curve.

We also introduce the notion of **torsion**: a measurement of how quickly the “direction” of the osculating plane changes along a curve $\mathcal{C} \subset \mathbb{R}^3$.

Although the above gives a heuristic explanation of torsion, it is somewhat wrong—after all, a plane does not have a “direction”. However, we bypass this by instead measuring the rate of change of the direction of the unit normal vector to the osculating plane.

Torsion is a scalar quantity which is signed; the $\pm$ sign depends upon whether the curve is right-handed or left-handed.

1.5.1 Binormal Vector and Normal Plane

The **binormal vector** is defined as

$$\boxed{\mathbf{B} := \mathbf{T} \times \mathbf{N} \quad (\text{Binormal Vector}).} \quad (1.5.1)$$

This tells us that $\mathbf{B}$ is perpendicular to both $\mathbf{T}$ and $\mathbf{N}$. Since the tangent and normal vectors span the osculating plane, this means that $\mathbf{B}$ is perpendicular to the osculating plane. This will be useful in finding equations of osculating planes.

The **normal plane** is defined as the plane spanned by $\mathbf{N}$ and $\mathbf{B}$. This plane is always perpendicular to $\mathbf{T}$ (intuitively, perpendicular to the curve at the given point).

The vectors $\mathbf{T}, \mathbf{N}, \mathbf{B}$ together form a basis of orthogonal vectors for $\mathbb{R}^3$. This collection of vectors—defined relative to some curve $\mathbf{r}$, and changing as we move along the curve—is called the **Frenet frame** for $\mathbf{r}$.

1.5.2 Torsion: Deriving a Formula

Consider the vector-valued function $\frac{d\mathbf{B}}{ds}$, with the derivative taken with respect to ar-length parametrization. We claim that this vector is perpendicular to both $\mathbf{B}$ and $\mathbf{T}$.

Since $|\mathbf{B}| = 1$ for all s , it follows that $\frac{d\mathbf{B}}{ds} \perp \mathbf{B}$ as discussed previously (see Worked Problems in Section 1.1). For $\mathbf{T}$, we compute the derivative of $\mathbf{B}$:

$$\begin{aligned} \frac{d\mathbf{B}}{ds} &= \frac{d}{ds}(\mathbf{T} \times \mathbf{N}) = \frac{d\mathbf{T}}{ds} \times \mathbf{N} + \mathbf{T} \times \frac{d\mathbf{N}}{ds} \\ &= \underbrace{\kappa(\mathbf{N} \times \mathbf{N})}_{=0} + \underbrace{\mathbf{T} \times \frac{d\mathbf{N}}{ds}}_{\perp \mathbf{T}} \end{aligned}$$

Since $\frac{d\mathbf{B}}{ds} \perp \mathbf{T}$ and also $\frac{d\mathbf{B}}{ds} \perp \mathbf{B}$, we know that $\frac{d\mathbf{B}}{ds}$ is parallel to the unit normal vector of the plane spanned by $\mathbf{T}$ and $\mathbf{B}$ —which is the vector $\mathbf{N}$!

In particular, for each s , there exists some scalar quantity $\tau = \tau(s)$ such that

$$\boxed{\frac{d\mathbf{B}}{ds} = -\tau\mathbf{N}.} \quad (1.5.2)$$

This scalar quantity $\tau = \tau(s)$ is called the **torsion** of $\mathbf{r}$ at s . (The letter τ (i.e. “tau”) comes from the Greek alphabet, and rhymes with “cow”.)

1.5.3 Torsion: A Worked Example

Let us consider again the helix

$$\mathbf{r}(t) = a \cos(t)\mathbf{i} + a \sin(t)\mathbf{j} + b t \mathbf{k}, \quad a > 0.$$

From earlier work, we know that

$$\begin{aligned} \mathbf{T}(t) &= -\frac{a}{\sqrt{a^2 + b^2}} \sin(t)\mathbf{i} + \frac{a}{\sqrt{a^2 + b^2}} \cos(t)\mathbf{j} + \frac{b}{\sqrt{a^2 + b^2}} \mathbf{k}, \\ \mathbf{N}(t) &= -\cos(t)\mathbf{i} - \sin(t)\mathbf{j}, \end{aligned}$$

It follows that

$$\begin{aligned} \mathbf{B}(t) &= \mathbf{T} \times \mathbf{N} = \frac{1}{\sqrt{a^2 + b^2}} \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -a \sin(t) & a \cos(t) & b \\ -\cos(t) & -\sin(t) & 0 \end{vmatrix} \\ &= \frac{1}{\sqrt{a^2 + b^2}} (b \sin(t)\mathbf{i} - b \cos(t)\mathbf{j} + a \mathbf{k}) \end{aligned}$$

We also know from previous examples that the curvature is given by

$$\kappa = \frac{a}{a^2 + b^2} \quad (1.5.3)$$

We now calculate the torsion by comparing the left-hand and right-hand sides of (1.5.2). From the left-hand side, we rewrite

$$\begin{aligned} \frac{d\mathbf{B}}{ds} &= \frac{d\mathbf{B}/dt}{ds/dt} = \frac{\mathbf{B}'(t)}{\sqrt{a^2 + b^2}} \\ &= \frac{1}{\sqrt{a^2 + b^2}} \cdot \underbrace{\frac{1}{\sqrt{a^2 + b^2}} (b \cos(t)\mathbf{i} + b \sin(t)\mathbf{j})}_{\mathbf{B}'(t)} \end{aligned}$$

Simplifying leads to the expression

$$\frac{d\mathbf{B}}{ds} = \frac{-b}{a^2 + b^2} (-\cos(t)\mathbf{i} - \sin(t)\mathbf{j}) = \frac{-b}{a^2 + b^2} \mathbf{N},$$

and our torsion equation (1.5.2) then gives

$$\tau = \frac{b}{a^2 + b^2}. \quad (1.5.4)$$

Remark 1.5.1. *From equations (1.5.3) and (1.5.4), we see that both the curvature and the torsion of the helix are constant and non-zero for all parameters t . This is one reason why we use the helix for so many examples—because these quantities (κ and τ) are well-behaved.*

1.5.4 The Frenet-Serret Formulas.

Notice that the vectors $\mathbf{T}' = \frac{d\mathbf{T}}{ds}$ and $\mathbf{B}' = \frac{d\mathbf{B}}{ds}$ both satisfy nice formulas in terms of the Frenet frame:

$$\frac{d\mathbf{T}}{ds} = \kappa\mathbf{N}, \quad \frac{d\mathbf{B}}{ds} = -\tau\mathbf{N}.$$

This leads one to ask whether there is a similar formula for $\frac{d\mathbf{N}}{ds}$? We derive such a formula here.

Recall: the Frenet frame $\mathbf{T}, \mathbf{N}, \mathbf{B}$ consists of mutually perpendicular unit vectors. By definition, we let $\mathbf{B} = \mathbf{T} \times \mathbf{N}$; however, by mutual orthogonality, we must also have

$$\mathbf{N} = \mathbf{B} \times \mathbf{T}, \quad \mathbf{T} = \mathbf{N} \times \mathbf{B}.$$

Hence, similar to our calculations for torsion, we have

$$\begin{aligned} \frac{d\mathbf{N}}{ds} &= \frac{d}{ds}(\mathbf{B} \times \mathbf{T}) = \frac{d\mathbf{B}}{ds} \times \mathbf{T} + \mathbf{B} \times \frac{d\mathbf{T}}{ds} \\ (\text{Definition of } \tau \text{ and } \kappa) &= (-\tau\mathbf{N}) \times \mathbf{T} + \mathbf{B} \times (\kappa\mathbf{N}) \\ &= \tau(\underbrace{\mathbf{T} \times \mathbf{N}}_{=\mathbf{B}}) - \kappa(\underbrace{\mathbf{N} \times \mathbf{B}}_{=\mathbf{T}}). \end{aligned}$$

Therefore, we have the formula

$$\frac{d\mathbf{N}}{ds} = \tau\mathbf{B} - \kappa\mathbf{T}. \quad (1.5.5)$$

This leads to the **Frenet-Serret formulas**, which in matrix form are:

$$\frac{d}{ds} \begin{pmatrix} \mathbf{T} \\ \mathbf{N} \\ \mathbf{B} \end{pmatrix} = \begin{pmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \tau \\ 0 & -\tau & 0 \end{pmatrix} \begin{pmatrix} \mathbf{T} \\ \mathbf{N} \\ \mathbf{B} \end{pmatrix} \quad (1.5.6)$$

1.5.5 Congruence of Curves in $\mathbb{R}^3$

We conclude with a deeper result concerning the curvature and torsion of space curves. This is the *Fundamental Theorem of Space Curves*. The proof is omitted.

Theorem 1.5.1. Suppose $\mathcal{C}_1$ and $\mathcal{C}_2$ are two curves with the same curvature $\kappa(t)$ and torsion $\tau(t)$, and $\kappa(t) \neq 0$. Then, the curves $\mathcal{C}_1$ and $\mathcal{C}_2$ are **congruent**—that is, $\mathcal{C}_2$ can be translated and rotated so as to coincide with $\mathcal{C}_1$.

This allows us to immediately classify when two curves are, in the geometric sense of *congruence*, “copies” of one another. As a nice and simple application, we have the following result.

Corollary 1.5.2. If $\mathcal{C}$ is a curve with κ and τ both nonzero and constant (independent of t), then the curve is congruent to some helix $\mathcal{H}$. If κ is nonzero and constant, but $\tau = 0$ for all t , then the curve is a circle.

Worked Problems

1. Find the equations of the normal and osculating plane for the helix

$$\mathbf{r}(t) = \cos(t)\mathbf{i} + \sin(t)\mathbf{j} + t\mathbf{k},$$

at $t = 0$.

To find the normal and osculating planes at $t = 0$, we need to determine their normal vectors $\mathbf{T}$ and $\mathbf{B}$ at $t = 0$. From the Worked Example above, with $a = b = 1$, we have

$$\begin{aligned}\mathbf{T}(t) &= \frac{1}{\sqrt{2}}\sin(t)\mathbf{i} + \frac{1}{\sqrt{2}}\cos(t)\mathbf{j} + \frac{1}{\sqrt{2}}\mathbf{k}, \\ \mathbf{B}(t) &= \frac{1}{\sqrt{2}}(\sin(t)\mathbf{i} - \cos(t)\mathbf{j} + \mathbf{k}).\end{aligned}$$

At $t = 0$, we have $\mathbf{r} = \langle 1, 0, 0 \rangle$, and

$$\mathbf{T} = \frac{1}{\sqrt{2}}\langle 0, 1, 1 \rangle \quad \mathbf{N} = \langle -1, 0, 0 \rangle, \quad \mathbf{B} = \frac{1}{\sqrt{2}}\langle 0, -1, 1 \rangle$$

The normal plane must be perpendicular to $\mathbf{j} + \mathbf{k}$ (we took a multiple of $\mathbf{T}$ to simplify the calculations), and passes through $(1, 0, 0)$. The equation of the plane is

$$0 \cdot (x - 1) + (y - 0) + (z - 0) = 0 \Leftrightarrow y + z = 0.$$

Similarly, the osculating plane is perpendicular to $-\mathbf{j} + \mathbf{k}$ and passes through $(1, 0, 0)$, so that its equation is

$$0 \cdot (x - 1) - (y - 0) + (z - 0) = 0 \Leftrightarrow -y + z = 0.$$

Hence, our equations for the planes are:

$$\begin{array}{ll}(\text{Normal Plane}) & y + z = 0 \\ (\text{Osculating Plane}) & -y + z = 0\end{array}$$

2. Recall the Frenet-Serret formulas:

$$\begin{aligned}\mathbf{T}'(s) &= \kappa \mathbf{N}, \\ \mathbf{N}'(s) &= -\kappa \mathbf{T} + \tau \mathbf{B}, \\ \mathbf{B}'(s) &= -\tau \mathbf{N},\end{aligned}$$

where κ is the curvature and τ is the torsion. Find a vector-valued function $\mathbf{W}$ such that these formulas can be written as $\mathbf{T}'(s) = \mathbf{W} \times \mathbf{T}$, $\mathbf{N}'(s) = \mathbf{W} \times \mathbf{N}$, $\mathbf{B}'(s) = \mathbf{W} \times \mathbf{B}$.

Let $\mathbf{W} = a\mathbf{T} + b\mathbf{N} + c\mathbf{B}$, where a, b, c are parameters. We are looking for a, b, c such that

$$\begin{aligned}\kappa \mathbf{N} &= \mathbf{W} \times \mathbf{T} = b\mathbf{N} \times \mathbf{T} + c\mathbf{B} \times \mathbf{T} = -b\mathbf{B} + c\mathbf{N}, \\ -\kappa \mathbf{T} + \tau \mathbf{B} &= \mathbf{W} \times \mathbf{N} = a\mathbf{T} \times \mathbf{N} + c\mathbf{B} \times \mathbf{N} = a\mathbf{B} - c\mathbf{T}, \\ -\tau \mathbf{N} &= \mathbf{W} \times \mathbf{B} = a\mathbf{T} \times \mathbf{B} + c\mathbf{N} \times \mathbf{B} = -a\mathbf{N} + b\mathbf{T}.\end{aligned}$$

This means $a = \tau$, $b = 0$, $c = \kappa$, $\mathbf{W} = \tau \mathbf{T} + \kappa \mathbf{B}$.

Practice Problems

1. Let C be the curve $x = e^{2t}$, $y = 2e^t$, $z = t$ for $t \in \mathbb{R}$. Find all points on the curve where the osculating plane is parallel to the plane $x - 6y + 18z = 0$. (If there are no such points, prove it.)
2. Prove that if $\mathbf{r}(t)$ is a smooth curve such that $\kappa \neq 0$ and $\mathbf{r} + \kappa^{-1}\mathbf{N}$ (the center of the osculating circle) is constant, then $\mathbf{r}(t)$ is a circle. (Hint: By the uniqueness theorem for curves (see above), it suffices to prove that κ is constant and the torsion τ is 0.)
3. Let C be a smooth curve parametrized by $\mathbf{r}(t) : \mathbb{R} \rightarrow \mathbb{R}^3$. Assume that $(0,0,0)$ belongs to the osculating plane at some point $\mathbf{r}(t_0)$. Prove that at $t = t_0$,

$$(\mathbf{r} \times \mathbf{r}') \cdot \mathbf{r}'' = 0.$$

1.6 Velocity and Acceleration

We have now constructed and examined the Frenet frame $\mathbf{T}$, $\mathbf{N}$, $\mathbf{B}$ for the past few sections. For this section, we wish to write these vectors in terms of well-known kinematic functions—specifically **velocity** and **acceleration**.

1.6.1 Deriving Equations for Velocity and Acceleration

Assume that $\mathbf{r}(t)$ is twice differentiable, with $\mathbf{r}' \neq 0$. Recall the following formulas

$$\mathbf{v}(t) = \mathbf{r}'(t) \quad (\text{Velocity}) \quad (1.6.1)$$

$$v(t) = |\mathbf{v}'(t)| \quad (\text{Speed}) \quad (1.6.2)$$

$$\mathbf{a}(t) = \mathbf{v}'(t) = \mathbf{r}''(t) \quad (\text{Acceleration}) \quad (1.6.3)$$

Our goal is to derive formulas for $\mathbf{v}$ and $\mathbf{a}$ in terms of the vectors $\mathbf{T}$, $\mathbf{N}$, and $\mathbf{B}$. For velocity, we have the straightforward identity

$$\mathbf{v}(t) = \mathbf{r}'(t) = |\mathbf{r}'(t)|\mathbf{T}(t) = v(t)\mathbf{T}(t). \quad (1.6.4)$$

For acceleration, we proceed in a few steps. First, note that

$$\mathbf{a}(t) = \mathbf{v}'(t) = \frac{d}{dt}(v\mathbf{T}) = \frac{dv}{dt}\mathbf{T} + v\frac{d\mathbf{T}}{dt}, \quad (1.6.5)$$

where we used (1.6.4) in the second equality. Now, recalling that

$$\mathbf{N} = \frac{\mathbf{T}'}{|\mathbf{T}'|} \Rightarrow \mathbf{T}' = |\mathbf{T}'|\mathbf{N},$$

equation (1.6.5) becomes

$$\mathbf{a} = \frac{dv}{dt}\mathbf{T} + v|\mathbf{T}'|\mathbf{N}. \quad (1.6.6)$$

Lastly, we use that

$$\kappa = \frac{|\mathbf{T}'|}{|\mathbf{r}'|} = \frac{|\mathbf{T}'|}{v} \Rightarrow |\mathbf{T}'(t)| = \kappa v(t),$$

to again rewrite (1.6.6) as

$$\boxed{\mathbf{a} = \frac{dv}{dt}\mathbf{T} + v^2\kappa\mathbf{N}.} \quad (1.6.7)$$

We call the scalar-valued functions $a_T = v'(t)$ and $a_N = \kappa v^2(t)$ the **tangential component** and **normal component** of acceleration, respectively. With this notation, we have

$$\mathbf{a}(t) = a_T(t)\mathbf{T}(t) + a_N(t)\mathbf{N}(t).$$

Two observations now follow immediately:

- Acceleration $\mathbf{a}$ has no $\mathbf{B}$ component in the Frenet frame. Geometrically, this means that acceleration is always parallel to the osculating plane (which is spanned by $\mathbf{T}$ and $\mathbf{N}$).
- Normal (or *centripetal*) acceleration is always directed towards the centre of curvature (the centre of the osculating circle); this component of acceleration represents the change in direction.

1.6.2 Alternative Formulas for Normal and Binormal Vectors

Let us derive new formulas for $\mathbf{B}$ and $\mathbf{N}$ in terms of velocity and acceleration.

To start, notice that (1.6.4) and (1.6.7) imply that both $\mathbf{v}$ and $\mathbf{a}$ are parallel to the osculating plane. So, their cross-product should be a scalar multiple of $\mathbf{B}$. In fact:

$$\begin{aligned}\mathbf{v} \times \mathbf{a} &= v\mathbf{T} \times (v'\mathbf{T} + v^2\kappa\mathbf{N}) = v^3\kappa(\mathbf{T} \times \mathbf{N}) \\ &= v^3\kappa\mathbf{B}.\end{aligned}\tag{1.6.8}$$

Since $\mathbf{B}$ is a unit vector, this means that

$$|\mathbf{v} \times \mathbf{a}| = v^3\kappa\tag{1.6.9}$$

(we already knew this from an earlier formula for κ), and that

$$\mathbf{B} = \frac{\mathbf{v} \times \mathbf{a}}{v^3\kappa} = \frac{\mathbf{v} \times \mathbf{a}}{|\mathbf{v} \times \mathbf{a}|}.$$

Now that we have this, we can compute $\mathbf{N}$ from the identity $\mathbf{N} = \mathbf{B} \times \mathbf{T}$. This calculation is often simpler than differentiating $\mathbf{T}$. Additionally, we can find the scalar functions a_T and a_N from the formulas

$$\begin{aligned}a_T &= \mathbf{a} \cdot \mathbf{T} = \frac{\mathbf{a} \cdot \mathbf{v}}{|\mathbf{v}|} \\ a_N &= \mathbf{a} \cdot \mathbf{N} = \sqrt{|\mathbf{a}|^2 - |a_T|^2}.\end{aligned}$$

1.6.3 Optional: A Formula for Torsion

We can also use our functions $\mathbf{v}$ and $\mathbf{a}$ to find an easily computable formula for torsion:

$$\tau = \frac{\mathbf{a}' \cdot (\mathbf{v} \times \mathbf{a})}{|\mathbf{v} \times \mathbf{a}|^2}\tag{1.6.10}$$

Proof. We write $\mathbf{a}'$ in terms of the Frenet frame: $\mathbf{a}' = \lambda\mathbf{T} + \mu\mathbf{N} + \nu\mathbf{B}$, and then pay particular attention to the last component. To begin, we differentiate

$$\begin{aligned}\mathbf{a}'(t) &= \frac{d}{dt}(v'\mathbf{T} + v^2\kappa\mathbf{N}) \\ &= v''\mathbf{T} + v'\underbrace{\mathbf{T}'(t)}_{|\mathbf{T}'|\mathbf{N}} + \frac{d}{dt}(v^2\kappa)\mathbf{N} + \underbrace{v^2\kappa\mathbf{N}'(t)}_{(*)}\end{aligned}$$

The first 3 parts are linear combinations of just $\mathbf{T}$ and $\mathbf{N}$. For the last part $(*)$, we convert to arc length parametrization with $\frac{ds}{dt} = v$, and then use the the Frenet-Serret equations:

$$(*) = v^2\kappa \frac{d\mathbf{N}/ds}{dt/ds} = v^2\kappa v(\tau\mathbf{B} - \kappa\mathbf{T}) = \dots + v^3\kappa\tau\mathbf{B},$$

where $\dots$ denotes some multiple of $\mathbf{T}$ that we are not interested in. So, in particular, we have

$$\mathbf{a}'(t) = \lambda\mathbf{T} + \mu\mathbf{N} + v^3\kappa\tau\mathbf{B},$$

for some scalar-valued functions λ, μ . The exact form of λ, μ does not matter because we will dispose of them in the next step. Recall from (1.6.8) that

$$\mathbf{v} \times \mathbf{a} = v^3\kappa\mathbf{B}.$$

Since the Frenet frame vectors are orthonormal, we get

$$\begin{aligned} \mathbf{a}' \cdot (\mathbf{v} \times \mathbf{a}) &= (\lambda\mathbf{T} + \mu\mathbf{N} + v^3\kappa\tau\mathbf{B}) \cdot (v^3\kappa\mathbf{B}) \\ &= (v^3\kappa\tau)(v^3\kappa) \\ &= (v^3\kappa)^2\tau \\ &= |\mathbf{v} \times \mathbf{a}|^2\tau, \end{aligned}$$

where we used (1.6.9) to replace $v^3\kappa$ by $|\mathbf{v} \times \mathbf{a}|$. Solving for τ , we get (1.6.10). □

Remark 1.6.1. Let $\mathbf{r}_0 = \mathbf{r}(t_0)$ be a point on the curve. For simplicity, we will assume that $t_0 = 0$. Consider the Taylor expansion at this point:

$$\begin{aligned} \mathbf{r}(t) &\approx \mathbf{r}(0) + t\mathbf{r}'(0) + \frac{t^2}{2}\mathbf{r}''(0) + \frac{t^3}{6}\mathbf{r}'''(0) \\ &= \mathbf{r}_0 + t\mathbf{v}(0) + \frac{t^2}{2}\mathbf{a}(0) + \frac{t^3}{6}\mathbf{a}'(0). \end{aligned}$$

Using the formulas we derived for $\mathbf{v}$, $\mathbf{a}$, and $\mathbf{a}'$ (can you find them?), we get

$$\mathbf{r}t \approx \mathbf{r}_0 + \dots + \frac{t^3}{6}v^3\kappa\tau\mathbf{B}(0),$$

where $\dots$ stands for some linear combination of $\mathbf{T}(0)$ and $\mathbf{N}(0)$. Notice that the $\mathbf{r}_0 + \dots$ part lies in the osculating plane Π_0 at $t = 0$, and $\mathbf{B}(0)$ is perpendicular to that plane. Thus the distance from the curve near $\mathbf{r}_0$ to Π_0 is about $v^3\kappa\tau t^3/6$, and this simplifies additionally to $\kappa\tau s^3/6$ if we use the arclength parametrization with $v = 1$. (For more precision, we could use the Taylor expansion with error estimates.) This justifies our earlier (and more vague) statements that τ measures how close the curve stays to a plane.

Worked Problems

1. Find $\mathbf{T}, \mathbf{N}, \mathbf{B}$ and κ if the curve is given by $\mathbf{r}(t) = \langle 2t, t^2, \frac{t^3}{3} \rangle$.

From earlier examples, we know that

$$\mathbf{v} = \mathbf{r}' = \langle 2, 2t, t^2 \rangle \Rightarrow v(t) = t^2 + 2$$

$$\mathbf{a} = \mathbf{v}' = \langle 0, 2, 2t \rangle.$$

We next compute

$$\mathbf{v} \times \mathbf{a} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2 & 2t & t^2 \\ 0 & 2 & 2t \end{vmatrix} = 2t^2\mathbf{i} - 4t\mathbf{j} + 4\mathbf{k}$$

and also

$$|\mathbf{v} \times \mathbf{a}| = \sqrt{4(t^4 + 4t^2 + 4)} = 2(t^2 + 2).$$

We can now use the formulas from this section, to determine that

$$\kappa = \frac{|\mathbf{v} \times \mathbf{a}|}{v^3} = \frac{2(t^2 + 2)}{(t^2 + 2)^3} = \frac{2}{(t^2 + 2)^2}$$

$$\mathbf{T} = \frac{\mathbf{v}}{v} = \frac{1}{t^2 + 2} \langle 2, 2t, t^2 \rangle$$

$$\mathbf{B} = \frac{\mathbf{v} \times \mathbf{a}}{|\mathbf{v} \times \mathbf{a}|} = \frac{1}{t^2 + 2} \langle t^2, -2t, 2 \rangle$$

$$\begin{aligned} \mathbf{N} &= \mathbf{B} \times \mathbf{T} = \frac{1}{(t^2 + 2)^2} \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ t^2 & -2t & 2 \\ 2 & 2t & t^2 \end{vmatrix} \\ &= \frac{1}{(t^2 + 2)^2} [(-2t^3 - 4t)\mathbf{i} - (t^4 - 4)\mathbf{j} + (2t^3 + 4t)\mathbf{k}] \end{aligned}$$

2. Let $\mathbf{r}(t) = \langle \cos t, t, t^2 \rangle$. Find κ , $\mathbf{T}$, $\mathbf{N}$, and $\mathbf{B}$ at $t = 0$.

The advantage of using our new formulas is that as soon as we have computed $\mathbf{v}$ and $\mathbf{a}$, we can plug in $t = 0$ and then do the rest of the calculation with $t = 0$ plugged in. No further differentiations will be needed!

We have $\mathbf{v}(t) = \langle -\sin t, 1, 2t \rangle$ and $\mathbf{a}(t) = \langle -\cos t, 0, 2 \rangle$, so that $\mathbf{v}(0) = \langle 0, 1, 0 \rangle$ and $\mathbf{a}(0) = \langle -1, 0, 2 \rangle$. Then

$$\mathbf{v}(0) \times \mathbf{a}(0) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 0 & 1 & 0 \\ -1 & 0 & 2 \end{vmatrix} = 2\mathbf{i} + \mathbf{k}, \quad |\mathbf{v}(0) \times \mathbf{a}(0)| = \sqrt{4 + 1} = \sqrt{5}$$

and $v(0) = 1$. At time $t = 0$, we have

$$\kappa(0) = \frac{|\mathbf{v}(0) \times \mathbf{a}(0)|}{v^3} = \frac{\sqrt{5}}{1} = \sqrt{5}$$

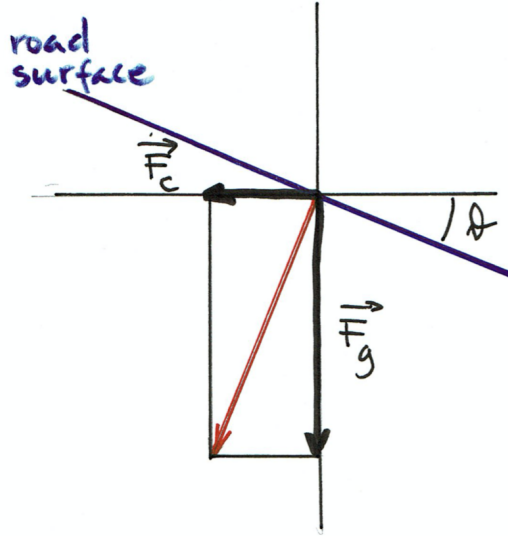
For the Frenet frame, we have

$$\mathbf{T}(0) = \frac{\mathbf{v}(0)}{v(0)} = \langle 0, 1, 0 \rangle = \mathbf{j}$$

$$\mathbf{B}(0) = \frac{\mathbf{v}(0) \times \mathbf{a}(0)}{|\mathbf{v}(0) \times \mathbf{a}(0)|} = \frac{1}{\sqrt{5}}(2\mathbf{i} + \mathbf{k})$$

$$\mathbf{N}(0) = \mathbf{B}(0) \times \mathbf{T}(0) = \frac{1}{\sqrt{5}} \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2 & 0 & 1 \\ 0 & 1 & 0 \end{vmatrix} = \frac{1}{\sqrt{5}}(-\mathbf{i} + 2\mathbf{k})$$

3. (*Banking the road*) A level, curved road has radius of curvature R (not necessarily constant). Assuming that cars will likely travel with constant speed v (in meters per second), at what angle θ should the road be banked so that the resultant of the centrifugal and gravitational forces acting on the vehicle is perpendicular to the surface of the road? Compute θ at a point where $R = 100\text{m}$, assuming that $v = 80\text{km/h}$. What if $R = 200\text{m}$?



The gravitational force acting on the car is $\mathbf{F}_g = -mg\mathbf{k}$, where m is the mass of the car and $g = 9.8 \text{ m/s}^2$ is the gravitational acceleration, and the centrifugal force is $\mathbf{F}_c = -ma_N\mathbf{N}$, where $a_N = v^2\kappa = v^2/R$. We want to choose θ so that

$$\tan \theta = \frac{ma_N}{mg}$$

(see picture). Thus $\tan \theta = a_N/g = v^2/(9.8R)$, with meters and seconds as units.

If $v = 80\text{km/h} \approx 22.2\text{m/s}$ and $R = 100\text{m}$, then, again using meters and seconds as units,

$$\tan \theta \approx \frac{(22.2)^2}{100 \cdot 9.8} \approx 0.5039$$

so that $\theta \approx \tan^{-1}(0.5039) \approx 26.7$ degrees.

If v is the same and $R = 200\text{m}$, then using the same calculation, $\theta \approx \tan^{-1}(0.25195) \approx 14.1$ degrees.

4. (*Acceleration and skidding*) Consider a car of mass m travelling along a curved but level road. To avoid skidding, the road must supply a frictional force $\mathbf{F} = m\mathbf{a}$, where $\mathbf{a}$ is the car's acceleration vector. The estimated magnitude of the frictional force is μmg where μ is the coefficient of friction and $g = 9.8\text{m/s}^2$. Let v be the car's speed in meters per second.

- (a) Show that the car will not skid if the curvature of the road is such that (with $R = 1/\kappa$)

$$(v')^2 + \left(\frac{v^2}{R}\right)^2 \leq (\mu g)^2.$$

Note that braking ($v' < 0$) and speeding up ($v' > 0$) contribute equally to skidding.

We need to have $|m\mathbf{a}| \leq \mu mg$, or equivalently, $|\mathbf{a}| \leq \mu g$, $|\mathbf{a}|^2 \leq (\mu g)^2$. Writing out the components of acceleration

$$\mathbf{a} = v'\mathbf{T} + v^2\kappa\mathbf{N}$$

we get the condition

$$|\mathbf{a}|^2 = (v')^2 + (v^2\kappa)^2 = (v')^2 + \left(\frac{v^2}{R}\right)^2 \leq (\mu g)^2$$

as required.

- (b) Suppose that the approximate radius of curvature along a stretch of a curved highway is $R = 180\text{m}$. How fast can an automobile travel (at constant speed) along that part of the highway if the coefficient of friction is $\mu = 0.5$? What if $R = 100\text{m}$?

If $R = 180\text{m}$, $\mu = 0.5$, and v is constant so that $v' = 0$, we get the condition

$$\frac{v^2}{R} \leq \mu g = 0.5 \cdot 9.8\text{m/s}^2 = 4.9\text{m/s}^2.$$

Plugging in $R = 180\text{m}$, we get

$$v^2 \leq 180\text{m} \cdot 4.9\text{m/s}^2 = 882(\text{m/s})^2$$

so that $v \leq 29.7\text{m/s} \approx 106.9\text{km/h}$.

If $R = 100\text{m}$, $\mu = 0.5$, and v is constant, we get instead

$$v^2 \leq 100\text{m} \cdot 4.9\text{m/s}^2 = 490(\text{m/s})^2$$

so that $v \leq 22.136\text{m/s} \approx 79.7\text{km/h}$.

Practice Problems

- The position of an object is given by a smooth vector-valued function $\mathbf{r}(t)$. Let $\mathbf{v}(t) = \mathbf{r}'(t)$, and $\mathbf{a}(t) = \mathbf{r}''(t)$. Assume that $\mathbf{a}(t) = \mathbf{k} \times \mathbf{v}(t)$ for all t .
 - Prove that the speed v is constant.
 - Assuming that v is nonzero, prove that the angle $\theta(t)$ between $\mathbf{v}(t)$ and $\mathbf{k}$ is constant.
- Assume that $\mathbf{a}(t) = c\mathbf{r}(t)$ for some constant c . (This describes the motion of an object under a central force.)
 - Prove that $\mathbf{w} := \mathbf{r} \times \mathbf{v}$ is constant.
 - Assuming that $\mathbf{w} \neq \mathbf{0}$, prove that the curve $\mathbf{r}(t)$ lies in the plane passing through the origin and perpendicular to $\mathbf{w}$.
- Suppose that a curve $\mathcal{C}$ in $\mathbb{R}^3$ has a smooth parametrization $\mathbf{r}(t)$, the curvature κ is a positive constant, and the torsion τ is zero. Prove, without using the Fundamental Theorem of Calculus for curves, that $\mathcal{C}$ is a circle. (Hint: how would you find the center of the circle? You may also want to use the Frenet-Serret formulas.)
- Let C be a curve on the sphere $\mathcal{S}$ of radius 1 centered at $\mathbf{0}$. Suppose that C has a smooth parametrization $\mathbf{r}(t)$ such that the velocity $\mathbf{r}'(t)$ is never zero and the acceleration $\mathbf{r}''(t)$ is always perpendicular to $\mathcal{S}$ at $\mathbf{r}(t)$. Prove that the curve is contained in a great circle of $\mathcal{S}$ (that is, a circle where $\mathcal{S}$ intersects a plane through $\mathbf{0}$).
- Let $\mathbf{r}(s) : \mathbb{R} \rightarrow \mathbb{R}^3$ be a smooth curve, parametrized by arc length. Assume that $\frac{d^2\mathbf{r}}{ds^2} = \mathbf{r} \times \mathbf{w}$ for all s , where $\mathbf{w}$ is a fixed nonzero vector. Prove that this is only possible when $\frac{d^2\mathbf{r}}{ds^2} = \mathbf{r} \times \mathbf{w} = \mathbf{0}$ for all s .
 (Hint: suppose that there is a value of s where $\mathbf{r} \times \mathbf{w} \neq \mathbf{0}$. Prove first that there are scalar functions $\lambda(s), \mu(s)$ such that $\frac{d\mathbf{r}}{ds} = \lambda\mathbf{r} + \mu\mathbf{w}$ near that point. What does this imply about $\frac{d^2\mathbf{r}}{ds^2}$?)

Chapter 2

Vector Fields

2.1 Introduction to Vector Fields

A **vector field** is a mapping $\mathbf{F} : \mathbb{R}^n \supset D \rightarrow \mathbb{R}^n$ which sends vectors $\mathbf{x} \in D$ to vectors $\mathbf{F}(\mathbf{x})$ in the same space. Here, D is the **domain** of $\mathbf{F}$, and we usually assume that $n = 2$ or $n = 3$. For vector fields in $\mathbb{R}^3$, we will use the component notation

$$\mathbf{r} = \langle x, y, z \rangle \in \mathbb{R}^3 \Rightarrow \mathbf{F}(\mathbf{r}) = F_1(x, y, z)\mathbf{i} + F_2(x, y, z)\mathbf{j} + F_3(x, y, z)\mathbf{k},$$

where the functions $F_i : \mathbb{R}^3 \rightarrow \mathbb{R}$ (the component functions of $\mathbf{F}$) are scalar-valued. In $\mathbb{R}^2$, the notation is similar except that we only have 2 components and 2 variables. We will say that a vector field is **smooth**, or C^1 , if all its component functions are C^1 . We will also refer to a function $f : \mathbb{R}^n \supset D \rightarrow \mathbb{R}$ as a **scalar field**.

2.1.1 Examples of Vector Fields from Physics

In physics, vector fields arise naturally as **force fields** or **velocity fields**. We take some time to look at a few examples.

Example 2.1.1. The **gravitational field** in $\mathbb{R}^3$ describes the force exerted by a stationary object of mass $M > 0$ (typically, large), placed at the origin, on a smaller moving object of mass $m > 0$. This field is given by the expression

$$\mathbf{F}(\mathbf{r}) = -\frac{GMm\mathbf{r}}{|\mathbf{r}|^3} = \frac{GMm}{|\mathbf{r}|^2} \left(-\frac{\mathbf{r}}{|\mathbf{r}|} \right), \quad (2.1.1)$$

where $G > 0$ is the gravitational constant. Notice that, due to the “ $-$ ” coefficient, all of the vectors point towards the origin, with magnitude $|\mathbf{F}(\mathbf{r})| = \frac{GMm}{|\mathbf{r}|^2}$.

Example 2.1.2. An **electrostatic force field** in $\mathbb{R}^3$ describes the force exerted by a stationary electric charge q_0 placed at the origin on a moving object of charge q :

$$\mathbf{E}(\mathbf{r}) = \frac{cq_0q\mathbf{r}}{|\mathbf{r}|^3} = \frac{cq_0q}{|\mathbf{r}|^2} \left(\frac{\mathbf{r}}{|\mathbf{r}|} \right), \quad (2.1.2)$$

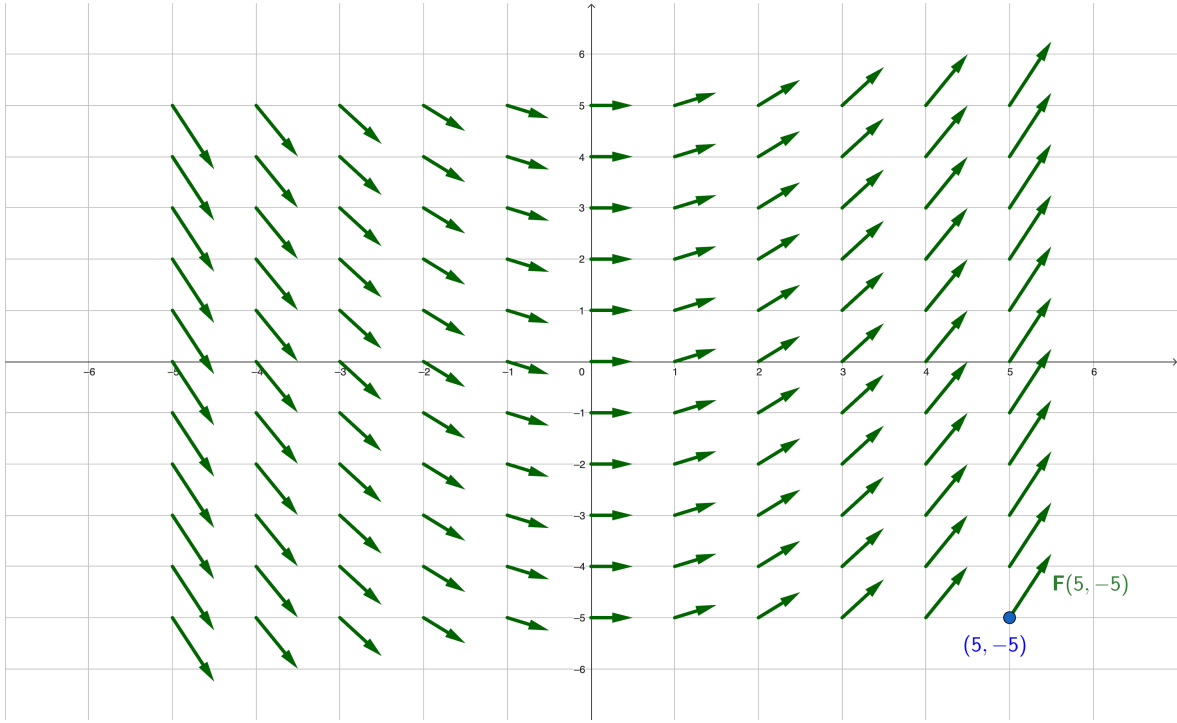


Figure 2.1: A vector field in $\mathbb{R}^2$. At each point (x, y) , we attach the vector $\mathbf{F}(x, y)$. An example is shown when $(x, y) = (5, -5)$.

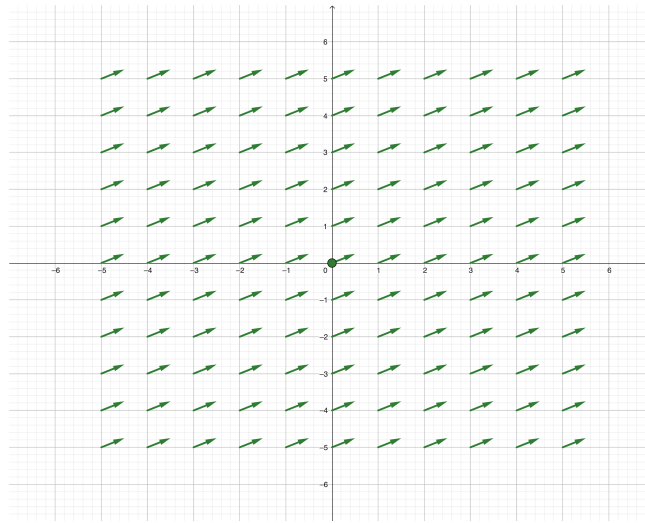
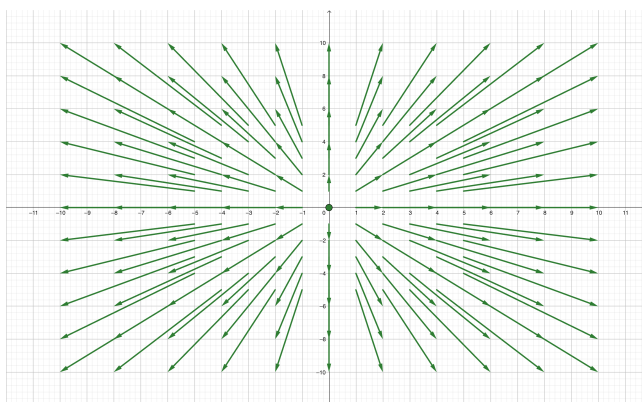
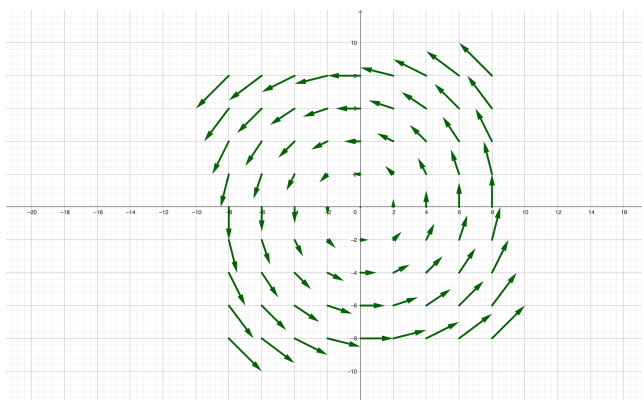


Figure 2.2: The vector field $\mathbf{F}(x, y) = \mathbf{a}$.

where $c > 0$ is a constant depending on the units. Both the gravitational field (2.1.1) and the electrostatic field (2.1.2) have the form $k\mathbf{r}/|\mathbf{r}|^3$ for some constant k , and we will often use this general form to include both examples. However, the gravitational force is always attractive (directed towards the origin), while the electrostatic force can

Figure 2.3: The vector field $\mathbf{F}(x, y) = x\mathbf{i} + y\mathbf{j}$.Figure 2.4: The vector field $\mathbf{F}(x, y) = -y\mathbf{i} + x\mathbf{j}$.

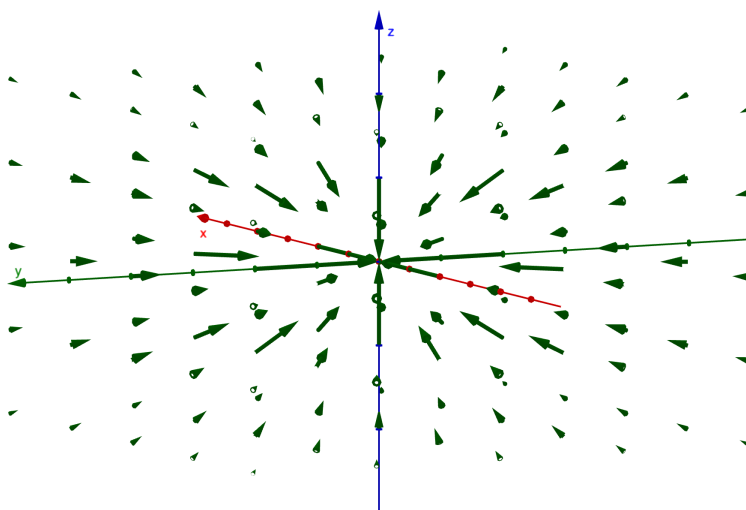
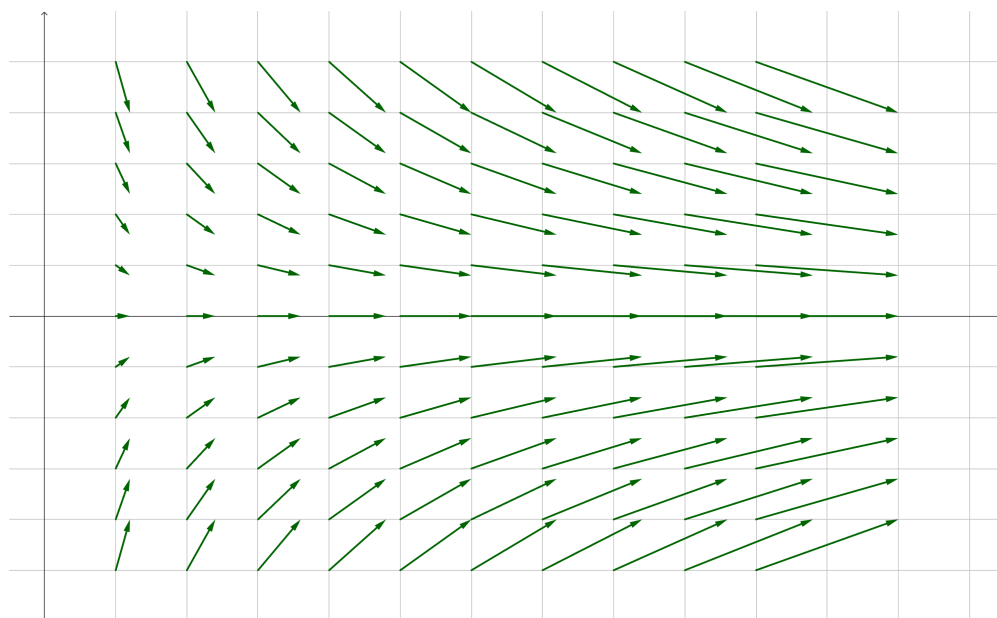
be either attractive or repulsive depending on the signs of q_0 and q .

Example 2.1.3. A **velocity field** $\mathbf{v}(x, y, z)$ describes the motion of a medium such as fluid or gas. At each point (x, y, z) , the vector $\mathbf{v}(x, y, z)$ is the velocity measured at that point. This may change from point to point (for example, when air flows through a funnel, the velocity increases when we get to the narrow part). It may also change over time, in which case the vector field may depend on the additional variable t .

Our next (and final) example appears often, both in physics and in mathematics, so we choose to give it as a definition.

Definition 2.1.4. Suppose that $f : \mathbb{R}^3 \supset D \rightarrow \mathbb{R}$ is a scalar field on some domain, and that the gradient $\nabla f := \langle f_x, f_y, f_z \rangle$ exists for every $\langle x, y, z \rangle \in D$. Then, the **gradient field** of f is defined by the equation $\mathbf{F}(\mathbf{r}) = \nabla f(\mathbf{r})$ for $\mathbf{r} \in D$. We will say that f is the **potential function** for $\mathbf{F}$.

This is a bit different from physics terminology, which is as follows: a scalar function $V : D \rightarrow \mathbb{R}$ is a potential function for the vector field $\mathbf{F}$ if $\mathbf{F} = -\nabla V$ (note the negative

Figure 2.5: A gravitational field in $\mathbb{R}^3$.Figure 2.6: A 2-dimensional velocity field (imagine a thin layer of water flowing on a surface). In this example, as $x \rightarrow \infty$ grows, the velocity becomes closer to horizontal and has larger magnitude.

sign!). For example, the Coulomb potential is given by

$$\mathbf{V}(\mathbf{r}) = \frac{cq_0q}{|\mathbf{r}|}, \text{ so that } \mathbf{E} = -\nabla\mathbf{V}.$$

In mathematics, it is easier to use the convention in Definition 2.1.4 ($\mathbf{F} = \nabla f$, without the minus sign). Most calculus textbooks follow this convention.

2.1.2 Field Lines

Field lines (also called *flow lines*, *integral curves*, or *trajectories*) are curves that are tangent to the field vectors at every point. Mathematically, we express this idea as

$$\frac{d\mathbf{r}}{dt} = \lambda(t)\mathbf{F}(\mathbf{r}(t)), \quad \lambda(t) \neq 0, \quad (2.1.3)$$

where $\mathbf{r}(t)$ is a parametrization of the curve and $\lambda(t)$ is some scalar function. We can simplify (2.1.3) by reparametrization: if $u = u(t)$ is chosen to satisfy

$$\frac{du}{dt} = \lambda(t) \Rightarrow \frac{d\mathbf{r}}{dt} = \frac{d\mathbf{r}}{du} \frac{du}{dt} = \lambda(t) \frac{d\mathbf{r}}{du}$$

then in the new parametrization $\mathbf{r} = \mathbf{r}(u)$, (2.1.3) becomes

$$\frac{d\mathbf{r}}{du} = \mathbf{F}(\mathbf{r}(u)). \quad (2.1.4)$$

To *check* that a given curve $\mathbf{r}(t)$ is a field line, it suffices to evaluate $\frac{d\mathbf{r}}{dt}$ and check that it is parallel to $\mathbf{r}$. As an example, consider again the vector field

$$\mathbf{F}(x, y) = -y\mathbf{i} + x\mathbf{j}.$$

Let us show that the curves $\mathbf{r}(t) = a \cos(t)\mathbf{i} + a \sin(t)\mathbf{j}$ (circles centered at the origin, of some radius $a > 0$) are, in fact, field lines for $\mathbf{F}(x, y)$. We have

$$\begin{aligned} \frac{d\mathbf{r}}{dt} &= -a \sin(t)\mathbf{i} + a \cos(t)\mathbf{j} \\ &= -y(t)\mathbf{i} + x(t)\mathbf{j} \\ &= \mathbf{F}(\mathbf{r}(t)). \end{aligned}$$

Hence, for each $a > 0$, the curve associated to $\mathbf{r}$ is a field line for $\mathbf{F}$.

Example 2.1.5. *On the following page, we show the field lines for the same vector fields as appeared in Figure 2.2, Figure 2.3 and Figure 2.4. When we overlay the field lines (the red curves) over the vector field (the green arrows), the orientation of the field lines are determined by the direction of the vector field.*

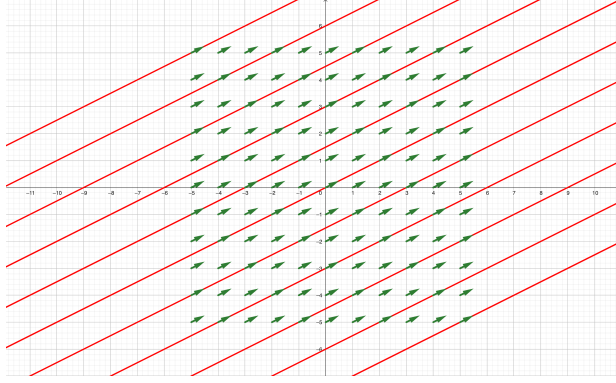


Figure 2.7: The field lines associated to the constant vector field $\mathbf{F}(x, y) = \mathbf{a}$ are parallel lines in the direction $\mathbf{a}$.

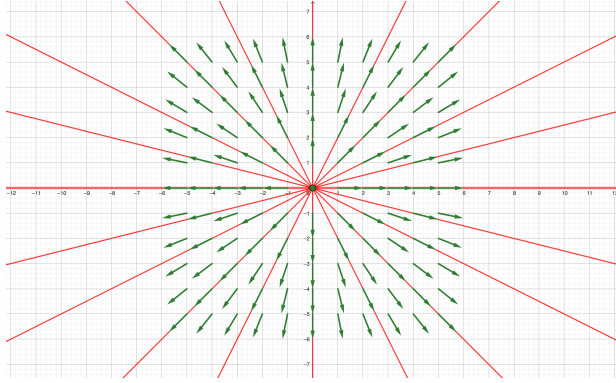


Figure 2.8: The field lines associated to the central force field $\mathbf{F}(x, y) = x\mathbf{i} + y\mathbf{j}$ are rays originating from the origin.

2.1.3 Solving for Field Lines; Existence of Field Lines

Now, in reality, we often want to *determine* an equation for the field lines of a given vector field. This corresponds to solving the equation $\frac{d\mathbf{r}}{dt} = \lambda(t)\mathbf{F}(\mathbf{r}(t))$. We can further simplify our lives by assuming that our parametrization satisfies (2.1.4). The calculation proceeds by re-writing (2.1.4) in terms of component functions:

$$\frac{dx}{dt} = F_1(\mathbf{r}(t)), \quad \frac{dy}{dt} = F_2(\mathbf{r}(t)), \quad \frac{dz}{dt} = F_3(\mathbf{r}(t)).$$

We need to solve this for the unknown functions $x(t)$, $y(t)$ and $z(t)$. We will use the *Separation of Variables* method from Partial Differential Equations: eliminate the t variable via

$$dt = \frac{dx}{F_1} = \frac{dy}{F_2} = \frac{dz}{F_3}, \quad (2.1.5)$$

and then solve the equations for x, y, z . We illustrate this in the example below.

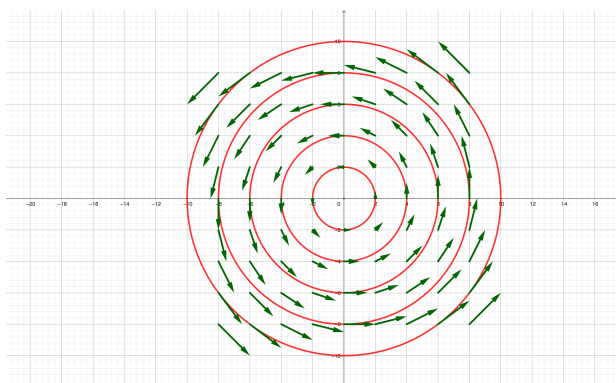


Figure 2.9: The field lines associated to the vector field $\mathbf{F}(x, y) = -y\mathbf{i} + x\mathbf{j}$ are concentric circles with the origin as a common centre. The inherited orientation of the field lines is counterclockwise.

Example 2.1.6. Determine the field lines of the vector field $\mathbf{F}(x, y) = 2x^2\mathbf{i} - \frac{y}{4}\mathbf{j}$.

For this vector field, we have $F_1(x, y) = 2x^2$ and $F_2(x, y) = -\frac{y}{4}$. Note that, so long as neither $x = 0$ or $y = 0$, (2.1.5) tells us that

$$dt = \frac{dx}{2x^2} = -\frac{4dy}{y}.$$

We integrate both sides (see the remark below for justification):

$$\int \frac{dy}{y} = -\frac{1}{8} \int \frac{dx}{x^2} \Rightarrow \ln |y| = \frac{1}{8x} + C.$$

Solving for y as a function of x , we have

$$|y| = e^{\frac{1}{8x} + C} = e^C e^{\frac{1}{8x}} \Rightarrow y = C_1 e^{\frac{1}{8x}},$$

where $C_1 = \pm e^C$ and $x \neq 0$ and $y \neq 0$.

If $x = 0$, then we have the equations

$$\frac{dx}{dt} = F_1(x, y) = 2x^2 = 0, \quad \frac{dy}{dt} = -\frac{y}{4}.$$

This corresponds to half-lines on the y -axis which point towards the origin. Similarly, if $y = 0$, we have

$$\frac{dx}{dt} = 2x^2, \quad \frac{dy}{dt} = -\frac{y}{4} = 0,$$

which corresponds to half lines on the x -axis. The full diagram looks something like the following.

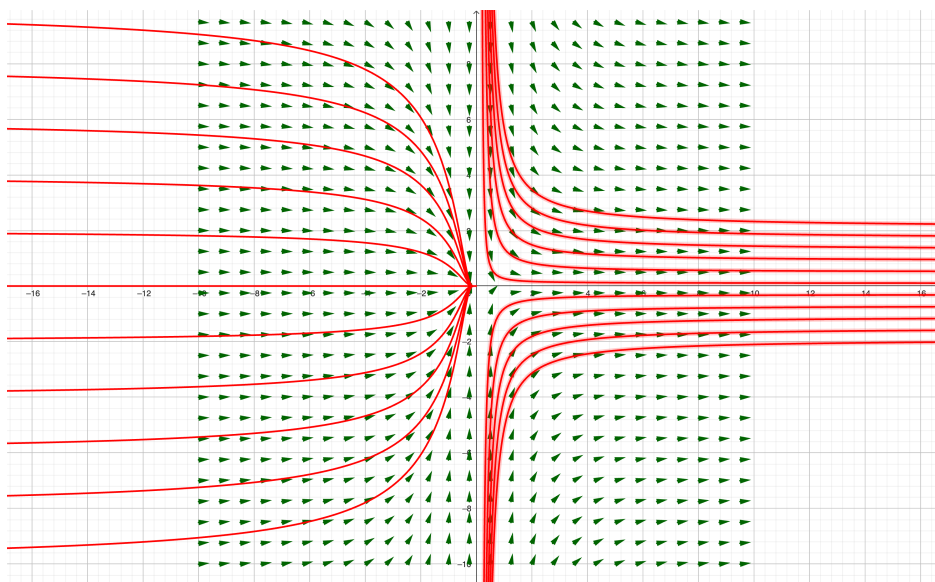


Figure 2.10: Some field lines of $\mathbf{F}(x, y) = 2x^2\mathbf{i} - \frac{y}{4}\mathbf{j}$ overlain on a portion of the vector field. In this diagram, we only include $\mathbf{F}(x, y)$ where $-10 \leq x, y \leq 10$.

Remark 2.1.1. *The integration step above is not obvious at all – in fact, it follows from the implicit function theorem. This method, leading to the equality of the anti-derivatives shown above, is known as the Method of Separation of Variables, and should be covered in Math 215/255.*

Finally, we can ask ourselves: how do we know, after following the procedure outlined above, that we have found *all* of the given field lines? The following result guarantees that our work has paid off—once we find one field line through every point $\mathbf{r}$ with $\mathbf{F}(\mathbf{r}) \neq \mathbf{0}$, we have found them all.

Theorem 2.1.7 (Existence and Uniqueness of Solutions to ODE's). *If $\mathbf{F} : \mathbb{R}^n \supset D \rightarrow \mathbb{R}^n$ is smooth on an open set $O \subset D$ containing $\mathbf{r}_0$, and if $\mathbf{F}(\mathbf{r}_0) \neq \mathbf{0}$, then there exists exactly one field line through $\mathbf{r}_0$.*

If $\mathbf{F}(\mathbf{r}_0) = \mathbf{0}$, then $\mathbf{r}_0$ is a *stationary point*. (If we interpret $\mathbf{F}$ as a velocity field, then an object placed at $\mathbf{r}_0$ should stay there indefinitely because its velocity is always zero.) In the previous example, notice that there is exactly one field line through every point in the plane, *except* at the origin—precisely where $\mathbf{F} = \mathbf{0}$! The origin, in this example, is a stationary point. By the theorem above, this gives us all field lines.

Worked Problems

1. Find all field lines of the vector field $\mathbf{F} = x(x+1)\mathbf{i} + y\mathbf{j}$, using the following procedure.

(a) Assuming that $x(x+1)$ and y are both non-zero, find a family of field lines by solving the corresponding ordinary differential equation via the Method of Separation of Variables.

We have the following differential equation when $x, x+1, y$ are all non-zero:

$$\frac{dy}{y} = \frac{dx}{x(x+1)} \quad (2.1.6)$$

Using partial fractions, $\frac{1}{x(x+1)} = \frac{1}{x} - \frac{1}{x+1}$. Then if we integrate both sides of (2.1.6), we get

$$\begin{aligned} \log |y| + C &= \log |x| - \log |x+1| \\ \log |y| + C &= \log \left| \frac{x}{x+1} \right| \\ |y| &= \left| \frac{x}{x+1} \right| e^{-C} \end{aligned}$$

The field lines are curves $y = \frac{Dx}{x+1}$, where $D = \pm e^{-C}$ is a nonzero constant.

(b) Check that the curves you found in (a) are in fact field lines.

We need to show that for an arbitrary field line $\mathbf{r}(t)$, $\mathbf{r}'(t)$ is parallel to $\mathbf{F}$ at each point on the curve. We can parametrize the field lines from (a) by $\mathbf{r}(t) = t\mathbf{i} + \frac{Dt}{t+1}\mathbf{j}$. Then $\mathbf{r}'(t) = \mathbf{i} + \frac{D}{(t+1)^2}\mathbf{j}$, and

$$\begin{aligned} \mathbf{F}(x(t), y(t)) &= x(t)(x(t)+1)\mathbf{i} + y(t)\mathbf{j} \\ &= t(t+1)\mathbf{i} + \frac{Dt}{t+1}\mathbf{j} \\ &= t(t+1) \left(\mathbf{i} + \frac{D}{(t+1)^2}\mathbf{j} \right) = t(t+1)\mathbf{r}'(t) \end{aligned}$$

(c) Consider the cases $x=0$, $x+1=0$, $y=0$ separately. Are there any field lines corresponding to these cases? Draw the picture of all the field lines you have found, with orientation.

When $x=0$ and $x=-1$, we have $\mathbf{F} = y\mathbf{j}$, with the vectors pointing vertically away from the x -axis. Along the line $y=0$, $\mathbf{F}$ is nonzero when $x \neq 0, -1$. The field vectors are $x(x+1)\mathbf{i}$, with the direction changing at each of the points $(0,0)$ and $(-1,0)$.

(d) A theorem from class says that if $\mathbf{F}$ is a smooth vector field on an open domain containing $\mathbf{r}_0$, and if $\mathbf{F}(\mathbf{r}_0) \neq 0$, then there is exactly one field line through $\mathbf{r}_0$.

Identify all points with $\mathbf{F}(\mathbf{r}) = 0$, Is it true that for every other point $\mathbf{r}_0$ in the plane, there is exactly one field line through $\mathbf{r}_0$ that you found above? If so, then this means you have found all the field lines.

We have $\mathbf{F} = 0$ at the points $(0,0), (-1,0)$. Let $\mathbf{r}_0 = (x_0, y_0)$ be a point different from $(0,0), (-1,0)$.

- If $\mathbf{r}_0$ lies on one of the lines $x = 0, x = -1$, or $y = 0$, then exactly one of the curves from part (c) passes through $\mathbf{r}_0$.
- Otherwise, let $D_0 = \frac{y_0(x_0+1)}{x_0}$ (this is well defined since $x_0 \neq 0$ and nonzero since $y_0(x_0+1) \neq 0$), Then the curve $y = \frac{D_0 x}{x+1}$ passes through $\mathbf{r}_0$. Moreover, for any other curve $y = \frac{Dx}{x+1}$ that passes through $\mathbf{r}_0$, we must also have $y_0 = \frac{Dx_0}{x_0+1}$, so that $D = D_0$.

Therefore there is exactly one field line through each point other than $(0,0), (-1,0)$.

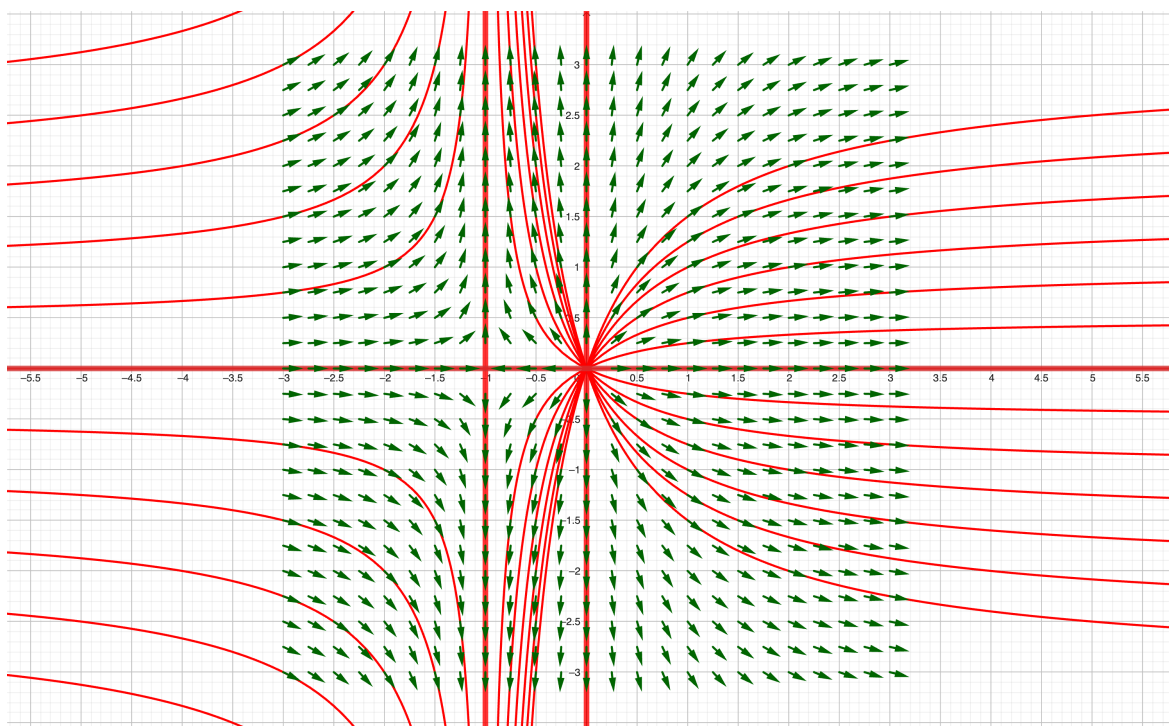


Figure 2.11: Field lines of $\mathbf{F}(x, y) = x(x+1)\mathbf{i} + y\mathbf{j}$ overlain on a portion of the vector field. The constant solutions $x = 0$, $x = -1$ and $y = 0$ correspond to the thick red lines. In this diagram, we only included the part of $\mathbf{F}(x, y)$ where $-3 \leq x, y \leq 3$.

Practice Problems

1. Find and sketch the field lines of the vector field $\mathbf{F}(x, y) = -x^3\mathbf{i} + y\mathbf{j}$ in $\mathbb{R}^2$. Indicate the flow direction in your sketch.
2. Let $f(x, y) = x^2 + 4y^4$. Find the field line of the vector field $\mathbf{F} = \nabla f$ which passes through the point $(2, 6)$.
3. Find all field lines of the vector field $\mathbf{F} = \frac{1-x^2}{x}\mathbf{i} + y^2\mathbf{j}$, using the following procedure.
 - (a) Assuming that x , $1 - x^2$, and y are all non-zero, solve the corresponding ordinary differential equation using the Method of Separation of Variables. (You may have to consider multiple cases.)
 - (b) What happens when $x = 0$, $1 - x^2 = 0$, or $y = 0$? Are there any field lines corresponding to these cases?
 - (c) Draw a picture of the field lines you found. Make sure to indicate their orientation. (You can check that by evaluating $\mathbf{F}$ at sample points.) You should have exactly one field line through every point where $\mathbf{F} \neq \mathbf{0}$.

2.2 Conservative Vector Fields

In this section, we will assume that $\mathbf{F}$ is smooth on some domain $D \subset \mathbb{R}^n$. We say that $\mathbf{F}$ is **conservative** if there is a (scalar) function $\phi : D \rightarrow \mathbb{R}$ such that

$$\mathbf{F} = \nabla\phi \text{ on } D. \quad (2.2.1)$$

As discussed in Section 2.1, we will call ϕ a *potential function* for $\mathbf{F}$. A potential function is not unique: if (2.2.1) holds for some ϕ , then it also holds with ϕ replaced by $\phi + C$ for any constant C . In questions where we need to find a potential function for a given $\mathbf{F}$, one potential function will be sufficient.

This section will focus on examples, with the goal of showing you how to prove (or disprove) that a given smooth vector field $\mathbf{F}$ is conservative.

Example 2.2.1 (Gravitational Field). *The gravitational field $\mathbf{F}(\mathbf{r}) = -\frac{k\mathbf{r}}{|\mathbf{r}|^3}$ is conservative on $\mathbb{R}^3 \setminus \{\mathbf{0}\}$. You can verify this by showing that $\mathbf{F} = \nabla\phi$ for the function $\phi(\mathbf{r}) = \frac{k}{|\mathbf{r}|}$.*

Example 2.2.2. *The function $\mathbf{F}(x, y) = y\mathbf{i} + x\mathbf{j}$ is conservative, with $\phi(x, y) = xy$. Indeed*

$$\phi_x(x, y) = y, \phi_y(x, y) = x \Rightarrow \mathbf{F}(x, y) = \nabla\phi,$$

which verifies (2.2.1).

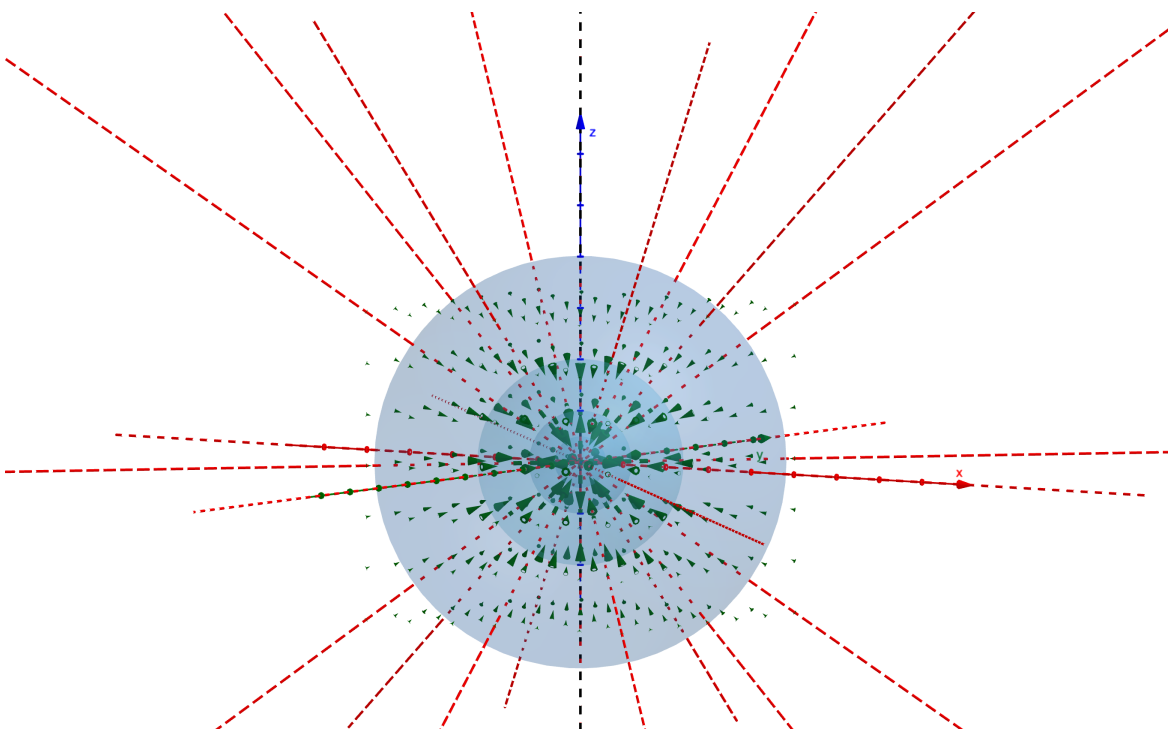


Figure 2.12: The gravitational field $\mathbf{F}(\mathbf{r}) = -k \frac{\mathbf{r}}{|\mathbf{r}|^3}$ from Example 2.2.1 is shown as green arrows. The field lines are rays terminating at the origin (which are the red lines). The equipotential surfaces are concentric spheres with common centre at the origin.

2.2.1 Equipotential curves and surfaces

Suppose that $\mathbf{F}$ is a conservative field, with an associated potential function ϕ . We will also assume that ϕ is C^1 on some domain D . An **equipotential curve** (in $\mathbb{R}^2$) or **equipotential surface** (in $\mathbb{R}^3$) is given by the equation $\phi = C$, where C is a constant.

Suppose that $\mathbf{r}(t)$ is a field line associated to $\mathbf{F}$, parametrized so that $\lambda(t) = 1$ and $\mathbf{r}'(t) = \mathbf{F}(\mathbf{r}(t))$. This implies that

$$\mathbf{r}'(t) = \mathbf{F}(\mathbf{r}(t)) = \nabla \phi(\mathbf{r}(t)). \quad (2.2.2)$$

Furthermore, by the Chain Rule and then (2.2.2) we have

$$\begin{aligned} \frac{d}{dt} \phi(\mathbf{r}(t)) &= \nabla \phi(\mathbf{r}(t)) \cdot \mathbf{r}'(t) = \mathbf{F}(\mathbf{r}(t)) \cdot \mathbf{F}(\mathbf{r}(t)) \\ &= |\mathbf{F}(\mathbf{r}(t))|^2. \end{aligned}$$

In particular, we always have

$$\frac{d}{dt} \phi(\mathbf{r}(t)) \geq 0,$$

with $\frac{d}{dt}\phi(\mathbf{r}(t)) = 0$ if and only if $\mathbf{F}(\mathbf{r}(t)) = \mathbf{0}$. Hence, if $\mathbf{F} \neq \mathbf{0}$, then ϕ is always increasing along field curves.

An additional geometric interpretation of (2.2.2) is as follows. Assume that $\mathbf{F} = \nabla\phi \neq \mathbf{0}$ on D . We know from Math 226 that $\nabla\phi$ is perpendicular to the level curves of ϕ . This shows that the field lines $\mathbf{r}' = \nabla\phi$ are perpendicular to equipotential curves, and always point in the direction of steepest increase of ϕ .

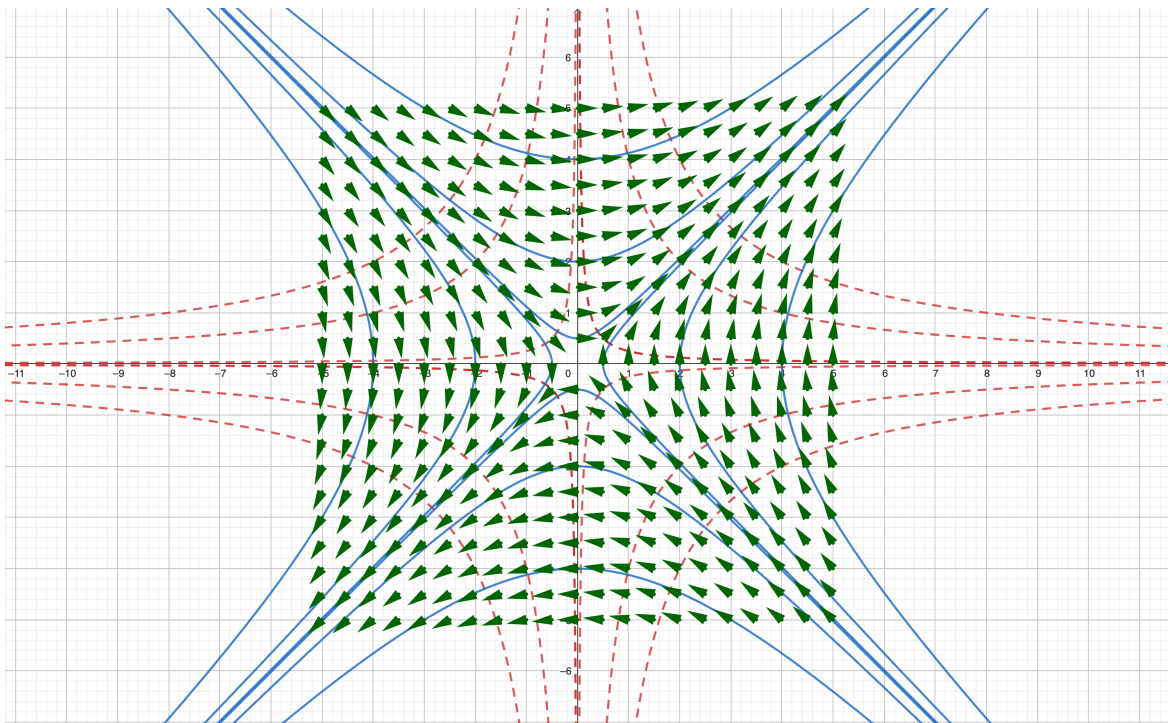


Figure 2.13: The conservative field $\mathbf{F}(x, y) = y\mathbf{i} + x\mathbf{j}$ from Example 2.2.2, is shown as green arrows. The field lines (shown here in solid blue) are given by equations of the form $y^2 - x^2 = C$ where $C \in \mathbb{R}$ is some constant. The equipotential curves (shown in dashed red) are hyperbolas and satisfy the equations $xy = D$ where $D \in \mathbb{R}$ constant. As an exercise: check that the quadratic equations $y^2 - x^2 = C$ define field lines for this vector field $\mathbf{F}(x, y)$!

2.2.2 Proving a Field is Conservative/Non-Conservative

The two examples below provide templates for determining whether a given field is conservative (Example 2.2.3) or non-conservative (Example 2.2.4).

Example 2.2.3 (A Conservative Field in $\mathbb{R}^3$). *Let*

$$\mathbf{F}(x, y) = \frac{2 \sin(x) \cos(x)}{z} \mathbf{i} - \frac{2 \sin(y) \cos(y)}{z} \mathbf{j} - \frac{\sin^2(x) + \cos^2(y)}{z^2} \mathbf{k}$$

on the domain $\mathbb{R}^3 \setminus \{z = 0\}$. Prove $\mathbf{F}$ is conservative by finding a potential function ϕ .

In order for ϕ to be a potential function, we must have that $\nabla\phi = \mathbf{F}$, so that

$$\frac{\partial\phi}{\partial x} = \frac{2\cos(x)\sin(x)}{z}, \quad \frac{\partial\phi}{\partial y} = -\frac{2\sin(y)\cos(y)}{z}, \quad \frac{\partial\phi}{\partial z} = -\frac{\sin^2(x) + \cos^2(y)}{z^2}.$$

We use these three equalities to reconstruct our function $\phi(x, y, z)$, as follows. First, note that

$$\frac{\partial\phi}{\partial x} = \frac{2\cos(x)\sin(x)}{z} \Rightarrow \phi(x, y, z) = \int \frac{2\cos(x)\sin(x)}{z} dx + g(y, z) = \frac{\sin^2(x)}{z} + g(y, z),$$

where $g(y, z)$ is some unknown function of the variables (y, z) . This unknown function $g(y, z)$ replaces the constant of integration, since we have only integrated in the x variable (and so $\frac{\partial}{\partial x}g(y, z) \equiv 0$).

Next,

$$\frac{\partial\phi}{\partial y} = \frac{\partial}{\partial y} \left[\frac{\sin^2(x)}{z} + g(y, z) \right] = \frac{\partial g}{\partial y}(y, z) = -\frac{2\sin(y)\cos(y)}{z},$$

and so our unknown function $g(y, z)$ necessarily satisfies

$$g(y, z) = \int -\frac{2\cos(y)\sin(y)}{z} dy + h(z) = \frac{\cos^2(y)}{z} + h(z),$$

where, now, $h(z)$ is an unknown function in the variable z alone! Hence, our unknown function $\phi(x, y, z)$ of three variables satisfies

$$\phi(x, y, z) = \frac{\sin^2(x)}{z} + \frac{\cos^2(y)}{z} + h(z).$$

As a final calculation:

$$\begin{aligned} \frac{\partial\phi}{\partial z} &= \frac{\partial}{\partial z} \left(\frac{\sin^2(x) + \cos^2(y)}{z} + h(z) \right) \\ &= -\frac{\sin^2(x) + \cos^2(y)}{z^2} + h'(z) \\ &= -\frac{\sin^2(x) + \cos^2(y)}{z^2}. \end{aligned}$$

This calculation shows that $h'(z) = 0$, which just means that $h(z) = C$ is some constant function. Setting $C = 0$, the function $\phi(x, y, z) = \frac{\sin^2(x) + \cos^2(y)}{z}$ satisfies all three equations. Hence, $\mathbf{F}$ is conservative, with a potential function ϕ .

Notice that, in the above example, each calculation re-wrote $\phi(x, y, z)$ as a function of “one less” variable, with the conclusion that the final function $h(z)$ was computable as a function of z alone (i.e. in this case, a constant function). This is not always possible, as the next example shows.

Example 2.2.4 (A Non-Conservative Field in $\mathbb{R}^2$). *In this example, we give three separate proofs that the smooth vector field $\mathbf{F}(x, y) = -y\mathbf{i} + x\mathbf{j}$ is not conservative on $\mathbb{R}^2$. Each proof provides a different insight into the problem.*

Proof #1 (Calculation). In the first proof, we assume that a potential function $\phi(x, y)$ satisfying

$$\frac{\partial \phi}{\partial x} = -y, \quad \frac{\partial \phi}{\partial y} = x,$$

exists, and then arrive at a contradiction. This follows the same template as Example 2.2.3. First, we know that

$$\frac{\partial \phi}{\partial x} = -y \Rightarrow \phi(x, y) = -\int y dx = -xy + g(y),$$

for some unknown function $g(y)$ in the variable y alone. Then, as a second step,

$$\frac{\partial \phi}{\partial y} = \frac{\partial}{\partial y}(-xy + g(y)) = -x + g'(y) = x.$$

So, we have that $g'(y) = 2x$. This is a contradiction, since g should be a function of the variable y alone.

Proof #2 (Geometric). From Section 2.1, we know that the field lines of $\mathbf{F}(x, y) = -y\mathbf{i} + x\mathbf{j}$ are circles centered at the origin. Now, suppose that $\mathbf{F}$ were conservative, with $\mathbf{F} = \nabla\phi$. Then, $\phi(\mathbf{r}(t))$ would necessarily increase along any field line $\mathbf{r}(t)$. But this leads to a contradiction!

Indeed, choose some starting point $\mathbf{r}_0 = \mathbf{r}(t_0)$ on the field line $\mathbf{r}$ (which is a circle). Then, there necessarily exists some $t_1 > t_0$ such that $\mathbf{r}(t_1) = \mathbf{r}_0$ (this corresponds to a full rotation of the circle). However, since ϕ is increasing, we then have

$$\phi(\mathbf{r}_0) = \phi(\mathbf{r}(t_0)) < \phi(\mathbf{r}(t_1)) = \phi(\mathbf{r}_0),$$

a contradiction.

Proof #3 (Equality of Mixed Partial). Suppose again that $\mathbf{F} = \nabla\phi$ for some C^2 function¹ ϕ . Then,

$$F_1 = \frac{\partial \phi}{\partial x}, \quad F_2 = \frac{\partial \phi}{\partial y}.$$

Since ϕ is C^2 , by Clairaux's theorem on the equality of mixed partials one obtains

$$\frac{\partial F_1}{\partial y} = \frac{\partial^2 \phi}{\partial y \partial x} = \frac{\partial^2 \phi}{\partial x \partial y} = \frac{\partial F_2}{\partial x}.$$

¹How do we know that ϕ is C^2 ? If $\mathbf{F} = \nabla\phi$, then $\phi_x = -y$ and $\phi_y = x$ have continuous derivatives of all orders, therefore so does ϕ !

But if $F_1 = -y$ and $F_2 = x$, then $\frac{\partial F_1}{\partial y} = -1$ and $\frac{\partial F_2}{\partial x} = 1$, again a contradiction.

Example 2.2.4 illustrates how the quality of “being a conservative vector field” is quite restrictive, having both geometric (Proof #2) and analytic (Proof #3) consequences for $\mathbf{F}$.

2.2.3 Conservative Fields and Equality of Mixed Partial Derivatives

We record the idea of Proof #3 in Example 2.2.4 as the following.

Proposition 2.2.5. *Suppose that $\mathbf{F} = F_1\mathbf{i} + F_2\mathbf{j} + F_3\mathbf{k}$ is a C^1 vector field on a domain $D \subset \mathbb{R}^3$. If $\mathbf{F}$ is conservative, then one necessarily has*

$$\frac{\partial F_1}{\partial y} = \frac{\partial F_2}{\partial x}, \quad \frac{\partial F_1}{\partial z} = \frac{\partial F_3}{\partial x}, \quad \frac{\partial F_2}{\partial z} = \frac{\partial F_3}{\partial y}. \quad (2.2.3)$$

Proof. Since $\mathbf{F} = \nabla\phi$ for some C^2 function $\phi : D \rightarrow \mathbb{R}$, we can use the equality of the mixed partial derivatives of ϕ to derive the equations (2.2.3). $\square$

For a smooth vector field $\mathbf{F} = F_1\mathbf{i} + F_2\mathbf{j}$ on $D \subset \mathbb{R}^3$, we have $F_3 = 0$ and there is no z dependence, so that (2.2.3) reduces to $\frac{\partial F_1}{\partial y} = \frac{\partial F_2}{\partial x}$.

Proposition 2.2.5 gives a handy tool for proving that a given vector field is *not* conservative. After all, if (2.2.3) fails for a given smooth vector field $\mathbf{F}$, then it is necessarily non-conservative. This is just the contrapositive of Proposition 2.2.5.

However, it is also entirely possible for a vector field $\mathbf{F}$ to satisfy equation (2.2.5) on its domain of definition and *still not be conservative*. This is demonstrated by the following example.

Example 2.2.6. *Let*

$$\mathbf{F}(x, y) = -\frac{y}{x^2 + y^2}\mathbf{i} + \frac{x}{x^2 + y^2}\mathbf{j},$$

so that $\mathbf{F}$ is a smooth vector field on $\mathbb{R}^2 \setminus \{(0, 0)\}$. Then

$$\frac{\partial F_1}{\partial y} = \frac{y^2 - x^2}{(x^2 + y^2)^2} = \frac{\partial F_2}{\partial x},$$

so that the corresponding two-dimensional equation (2.2.3) of Proposition 2.2.5 holds. Nevertheless, $\mathbf{F}$ fails to be conservative on its domain.

To prove that $\mathbf{F}$ is non-conservative, we show that its field lines are the circles

$$\mathbf{r}(t) = a \cos(t)\mathbf{i} + a \sin(t)\mathbf{j}.$$

Note that

$$\mathbf{r}'(t) = \underbrace{-a \sin(t)}_{\mathbf{y}(t)} \mathbf{i} + \underbrace{a \cos(t)}_{x(t)} \mathbf{j},$$

so that

$$\begin{aligned} \mathbf{F}(\mathbf{r}(t)) &= \frac{y(t)}{x^2(t) + y^2(t)} \mathbf{i} + \frac{x(t)}{x^2(t) + y^2(t)} \mathbf{j} \\ &= \frac{y(t)}{a^2} \mathbf{i} + \frac{x(t)}{a^2} \mathbf{j}. \end{aligned}$$

Therefore, we have $\mathbf{r}'(t) = a^2 \mathbf{F}(\mathbf{r}(t))$, and so $\mathbf{r}$ is indeed a field line of $\mathbf{F}$. As in Proof #2 in Example 2.2.4, this shows that $\mathbf{F}$ is not conservative on $\mathbb{R}^2$.

Worked Problems

1. Let $\mathbf{F} = F_1(x)\mathbf{i} + F_2(y)\mathbf{j} + F_3(z)\mathbf{k}$ be a smooth vector field on $\mathbb{R}^3$ —so, each component function is a single-variable function. Prove that $\mathbf{F}$ is conservative on $\mathbb{R}^3$, by finding a function ϕ such that $\mathbf{F} = \nabla\phi$.

Since F_1 is continuous on $\mathbb{R}$, the Fundamental Theorem of Calculus asserts that the function $f_1(x) = \int_0^x F_1(t)dt$ defined on $\mathbb{R}$ satisfies $\frac{d}{dx}f_1(x) = F_1(x)$ for each $x \in \mathbb{R}$. Similarly, there are functions f_2, f_3 defined on $\mathbb{R}$ with $\frac{d}{dy}f_2(y) = F_2(y)$ and $\frac{d}{dz}f_3(z) = F_3(z)$. Define $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}$ by

$$\phi(x, y, z) = f_1(x) + f_2(y) + f_3(z)$$

Then

$$\frac{\partial \phi}{\partial x} = \frac{\partial}{\partial x}(f_1(x) + f_2(y) + f_3(z)) = \frac{d}{dx}f_1(x) = F_1(x),$$

and similarly, $\frac{\partial \phi}{\partial y} = F_2(y)$, $\frac{\partial \phi}{\partial z} = F_3(z)$.

2. Find all values of the parameters a, b, c such that the vector field

$$\mathbf{F}(x, y) = (3x^2 + 2x)e^{ax+y}z^c\mathbf{i} + e^{ax+y}x^b z^c\mathbf{j} + ce^{ax+y}x^b z^2\mathbf{k}$$

is conservative on $\mathbb{R}^3$. For such a, b, c , find a potential function for $\mathbf{F}$.

Suppose that $\mathbf{F}$ is conservative, with $\mathbf{F} = \nabla f$. Then we must have

$$\frac{\partial}{\partial y}f(x, y, z) = F_2 = e^{ax+y}x^b z^c,$$

so that $f(x, y, z) = e^{ax+y}x^bz^c + g(x, z)$. We must also have (assuming, for now, that $b \neq 0$, see below)

$$\frac{\partial}{\partial x}f(x, y, z) = (ax^b + bx^{b-1})e^{ax+y}z^c + \frac{\partial g(x, z)}{\partial x} = F_1 = (3x^2 + 2x)e^{ax+y}z^c,$$

so that $a = 3$, $b = 2$, and $g = g(z)$. Note that if we had $b = 0$, then we would have

$$\frac{\partial}{\partial x}f(x, y, z) = e^{ax+y}z^c + \frac{\partial g(x, z)}{\partial x},$$

and that does not match the equation for F_1 . Finally,

$$\frac{\partial}{\partial z}f(x, y, z) = ce^{ax+y}x^bz^{c-1} + g'(z) = ce^{ax+y}x^bz^2,$$

so that $c = 3$ and we can take $g \equiv 0$. The answer is $a = c = 3$, $b = 2$, and $f(x, y, z) = e^{3x+y}x^2z^3$.

Practice Problems

1. Let $\mathbf{F}(x, y, z) = e^y\mathbf{i} + (xe^y - \cos z)\mathbf{j} + y \sin z$ in $\mathbb{R}^3$. Prove that $\mathbf{F}$ conservative, and find a potential function f .
2. Consider the vector field

$$\mathbf{G} = \frac{F_1(x)}{F_3(z)}\mathbf{i} + \frac{F_2(y)}{F_3(z)}\mathbf{j} + \mathbf{k},$$

where F_1, F_2, F_3 are smooth functions on a domain $D \subset \mathbb{R}^3$ and F_3 is never zero on D . Does $\mathbf{G}$ have to be conservative on D ? If yes, prove it. If no, give a counterexample.

3. Give an example of two smooth vector fields $\mathbf{F}$ and $\mathbf{G}$ on $\mathbb{R}^2$ such that $\mathbf{F}$ and $\mathbf{G}$ have the same field lines, $\mathbf{F}$ is conservative, and $\mathbf{G}$ is not conservative.

(Hint: Worked Problem 1 of this Section might be helpful in constructing a simple conservative field $\mathbf{F}$. Then, try modifying it to make a $\mathbf{G}$ that has the same field lines but is non-conservative.)

4. Find all functions $f(y)$ and $g(x)$, where $f(y)$ depends only on y and $g(x)$ depends only on x , such that the vector field $\mathbf{F}(x, y) = f(y)\mathbf{i} + g(x)\mathbf{j}$ is conservative on $\mathbb{R}^2$. For such f and g , find the potential function for $\mathbf{F}$.

2.3 Line Integrals

Let C be a smooth curve on $\mathbb{R}^n$ (where $n = 2$ or $n = 3$) with endpoints, and let $\mathbf{r} : [a, b] \rightarrow \mathbb{R}^n$ be a smooth parametrization of C .

2.3.1 Scalar Line Integrals as Limits of Riemann Sums

Let $f : D \rightarrow \mathbb{R}$ be a function on some domain $D \subset \mathbb{R}^n$ which contains the curve C . We define the **line integral** of f over C

$$\int_C f(x, y, z) ds, \text{ where } ds \text{ is arc-length,} \quad (2.3.1)$$

as the limit of the Riemann sums

$$S := \sum_{i=1}^n f(\mathbf{r}(t_i^*)) |\Delta \mathbf{r}_i| = \sum_{i=1}^n f(x_i^*, y_i^*, z_i^*) |\Delta \mathbf{r}_i|.$$

Let us unpack this definition. The points t_i are coming from some partition

$$a = t_0 < t_1 < \cdots < t_n = b$$

of the interval $I = [a, b]$. Having fixed this partition, we then define $\Delta \mathbf{r}_i = \mathbf{r}(t_i) - \mathbf{r}(t_{i-1})$, and let

$$\mathbf{r}(t_i^*) := (x_i^*, y_i^*, z_i^*), \text{ for some } t_i^* \in [t_{i-1}, t_i].$$

We call these points $\mathbf{r}(t_i^*)$ *evaluation points* or *sampling points*. The situation is depicted below. We then take the limit over all such partitions as $\max_i |t_i - t_{i-1}| \rightarrow 0$. For the integral (2.3.1) to exist, we need the limit to exist and to be independent of the choice of the partitions and sampling points.

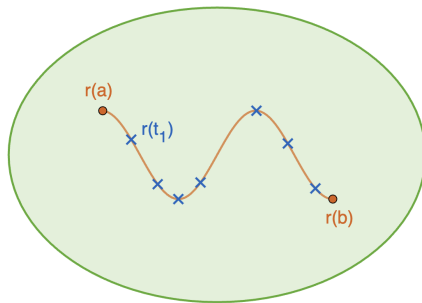


Figure 2.14: The curve C (in orange) is contained in the domain $D \subset \mathbb{R}^n$ (the green oval). The “cross-marks” are the $\mathbf{r}(t_i)$.

To derive a formula for evaluating line integrals, we first make the approximation

$$\Delta \mathbf{r}_i = \frac{\mathbf{r}(t_i) - \mathbf{r}(t_{i-1})}{t_i - t_{i-1}} (t_i - t_{i-1}) \approx \mathbf{r}'(t_i) \Delta t_i, \quad (2.3.2)$$

where $\Delta t_i = t_i - t_{i-1}$. The approximation will have less error as $\Delta t_i \rightarrow 0$. Using (2.3.2), we have

$$S_n \approx \sum_{i=1}^n f(\mathbf{r}(t_i)) |\mathbf{r}'(t_i)| \Delta t_i = S_n^*. \quad (2.3.3)$$

But, S_n^* is a Riemann sum associated to the following (one-dimensional) definite integral

$$\int_a^b f(\mathbf{r}(t)) |\mathbf{r}'(t)| dt.$$

Recalling that arc-length parametrization gives $ds = |\mathbf{r}'(t)| dt$, we arrive at the following formula.

Evaluating Scalar Line Integrals

Suppose that $C \subset D$ is a smooth arc and $f : D \rightarrow \mathbb{R}$ is continuous on C . Then, the line integral of f on C satisfies the equality

$$\int_C f ds = \int_a^b f(\mathbf{r}(t)) |\mathbf{r}'(t)| dt = \int f(\mathbf{r}(s)) ds. \quad (2.3.4)$$

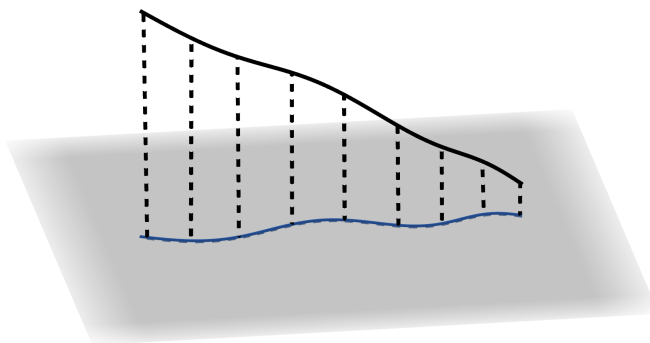
You can see several applications of this formula in the Worked Problems below. A few important remarks regarding scalar line integrals:

- If $C : [a, b] \rightarrow \mathbb{R}^n$ is a smooth curve, and if f is continuous on some domain containing C , then the limit above exists. (We do not prove it here, but this is similar to the existence of Riemann integrals of continuous functions on bounded intervals.) This does not always happen for arbitrary curves C and functions $f : D \rightarrow \mathbb{R}^n$. In cases where the assumptions above do not hold, the limit may still exist, but we need to justify it on a case-by-case basis.
- When $f(x, y, z) = 1$ for all x, y, z on the curve, these are the same Riemann sums that we used to define arc length of the curve. Hence, $\int_C 1 ds$ (or $\int_C ds$ for short) is just the length of the curve.
- We can extend arc length integrals to piecewise smooth curves by integrating on each smooth piece and then adding up the integrals. Similarly, improper line integrals (e.g. line integrals over curves C of infinite length) can be evaluated by integrating on a finite part of the curve and then taking the limit (if it exists).

2.3.2 A Physical Interpretation of Line Integrals

Line integrals often have natural interpretations in physical space. For example, if $f : D \rightarrow \mathbb{R}$ is some continuous function on a smooth arc $C \subset D$, we can interpret the line integral $\int_C f ds$ as:

- The mass of a wire with shape given by the curve C , if we make the assumption that the density at point $\mathbf{r}$ is given by $f(\mathbf{r})$. So, the line integral becomes an expression for *total mass* over the wire.
- If C lies in the xy -plane and $f(x, y) \geq 0$, we can view the line integral $\int_C f ds$ as the area of a curvilinear fence that has height $f(x, y)$ at every point (x, y) of C . One example is shown in the following figure.



- The *average* of a function $f(\mathbf{r})$ over the curve C is given by the expression:

$$\frac{\int_C f ds}{\int_C ds}.$$

Since $\int_C ds$ represents the length of C , this is similar to taking the average of a single-variable function $f : I \rightarrow \mathbb{R}$ over an interval $I \subset \mathbb{R}$.

2.3.3 Vector Line Integrals

We now define line integrals of vector fields. In this type of integrals, we are not just integrating each component of $\mathbf{F}$ separately with respect to arc length: the integral also takes into account how well $\mathbf{F}$ is aligned with C .

Again, we assume that C is a smooth curve, parametrized by $\mathbf{r}(t) : [a, b] \rightarrow \mathbb{R}^n$. We further assume that $\mathbf{F} : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ is continuous on some domain D which contains C . We then define the **line integral** of $\mathbf{F}$ over C as

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_C \mathbf{F} \cdot \mathbf{T} ds \quad (2.3.5)$$

The function $\mathbf{F} \cdot \mathbf{T}$ is equal (up to $\pm$ signs) to the length of the tangential component $\mathbf{F}_T$ of $\mathbf{F}$ along C —a scalar quantity. This is depicted in the following diagram.

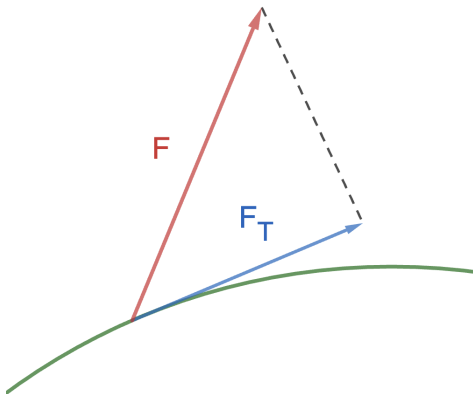


Figure 2.15: The vector field $\mathbf{F}$ and its tangential component $\mathbf{F}_T$ along some smooth arc. The dashed line reminds you that $\mathbf{F}_T$ is the orthogonal projection of $\mathbf{F}$ onto the unit tangent vector $\mathbf{T}$ of $\mathbf{r}$.

Evaluating Vector Line Integrals

Recall that $\mathbf{T} = \frac{\mathbf{r}'}{|\mathbf{r}'|}$ and that the arc-length differential satisfies $ds = |\mathbf{r}'|dt$. Using this, the line integral of a vector field $\mathbf{F}$ is then:

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_a^b \mathbf{F} \cdot \frac{\mathbf{r}'}{|\mathbf{r}'|} |\mathbf{r}'| dt = \int_a^b \mathbf{F} \cdot \mathbf{r}' dt. \quad (2.3.6)$$

In terms of components, this gives the equation

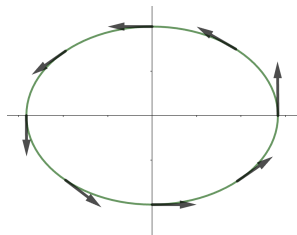
$$\begin{aligned} \int_C \mathbf{F} \cdot d\mathbf{r} &= \int_a^b F_1 x'(t) dt + \int_a^b F_2 y'(t) dt + \int_a^b F_3 z'(t) dt \\ (\text{Alternative Notation}) &= \int_C F_1 dx + F_2 dy + F_3 dz. \end{aligned}$$

So, to evaluate the line integral of a vector field $\mathbf{F} : D \rightarrow \mathbb{R}^3$, it suffices to evaluate the component definite integrals.

Line integrals $\int_C \mathbf{F} \cdot d\mathbf{r}$ of vector field $\mathbf{F}$ also have natural interpretations:

- The work done by a force $\mathbf{F}$ in moving a particle along a path $\mathbf{r}(t)$. (Note: if $\mathbf{F} \cdot \mathbf{T}$ is negative, so that the motion is against the direction of the force, this means that the force has opposed the motion and therefore made a negative contribution to it.)

- The *circulation* of $\mathbf{F}$ along C —this is especially true in the case where C is a closed curve. This is shown in the picture below, with the circulation measured in the positive (counterclockwise) direction.



Worked Problems

1. Find $\int_C z \, ds$ if C is the arc of the helix $\mathbf{r}(t) = \cos(t)\mathbf{i} + \sin(t)\mathbf{j} + t\mathbf{k}$ over the interval $t \in I = [0, 2\pi]$.

One readily calculates that $ds = \sqrt{2}dt$, whereas $z(t) = t$ along the helix. Hence, the line integral formula (2.3.4) gives

$$\int_C z \, ds = \int_0^{2\pi} t\sqrt{2} \, dt = \frac{t^2\sqrt{2}}{2} \Big|_0^{2\pi} = 2\sqrt{2}\pi^2.$$

2. Find $\int_C x^2y \, ds$ if C is the line segment from $(0, 1)$ to $(1, 3)$ in $\mathbb{R}^2$.

We parametrize C as $\mathbf{r}(t) = (t, 2t+1)$ for $t \in I = [0, 1]$. Then, $\mathbf{r}'(t) = (1, 2) \Rightarrow ds = \sqrt{5}dt$. Hence, applying the line integral formula (2.3.4) gives

$$\begin{aligned} \int_C x^2y \, ds &= \int_0^1 t^2(2t+1)\sqrt{5} \, dt = \int_0^1 \sqrt{5}(2t^3 + t^2) \, dt \\ &= \sqrt{5} \left(\frac{t^4}{2} + \frac{t^3}{3} \right) \Big|_0^1 \\ &= \frac{5\sqrt{5}}{6}. \end{aligned}$$

3. Evaluate the line integral $\int_C (x-y) \, ds$, where C is the curve parametrized by $x = t^2 + t$, $y = t^2 - t$, $z = \frac{t^2}{2} + 1$, from $t = 0$ to $t = 1$.

We have $\mathbf{r}'(t) = (2t+1)\mathbf{i} + (2t-1)\mathbf{j} + t\mathbf{k}$, so that

$$|\mathbf{r}'(t)|^2 = (4t^2 + 4t + 1) + (4t^2 - 4t + 1) + t^2 = 9t^2 + 2.$$

Therefore

$$\int_C (x - y) ds = \int_0^1 2t\sqrt{9t^2 + 2} dt.$$

Substituting $u = 9t^2 + 2$, $du = 18t$, $2 \leq u \leq 11$, we get

$$\begin{aligned} \int_C (x - y) ds &= \int_2^{11} \sqrt{u} \frac{du}{9} \\ &= \frac{2}{3} \left[\frac{u^{3/2}}{9} \right]_2^{11} = \frac{2}{27} (11^{3/2} - 2^{3/2}) \end{aligned}$$

4. Evaluate the integral $\int_C xy \, ds$, where C is the part of the curve of intersection of the surfaces $x^2 + y^2 = 1$ and $z = 2y$ where $x \geq 0$ and $y \geq 0$.

We use the parametrization $x = \cos t$, $y = \sin t$, $z = 2 \sin t$, $0 \leq t \leq \pi/2$. Then $\mathbf{r}'(t) = -\sin t \mathbf{i} + \cos t \mathbf{j} + 2 \cos t \mathbf{k}$, so that

$$|\mathbf{r}'| = \sqrt{\sin^2 t + \cos^2 t + 4 \cos^2 t} = \sqrt{1 + 4 \cos^2 t},$$

$$\int_C xy \, ds = \int_0^{\pi/2} \cos t \sin t \sqrt{1 + 4 \cos^2 t} \, dt.$$

Letting $u = 1 + 4 \cos^2 t$, $du = -8 \cos t \sin t \, dt$, we get

$$\int_C xy \, ds = -\frac{1}{8} \int_5^1 \sqrt{u} \, du = \frac{1}{8} \cdot \frac{2}{3} u^{3/2} \Big|_1^5 = \frac{1}{12} (5^{3/2} - 1).$$

5. (**Warning: Long Calculation!**) The mass of a wire in the shape of a parametrized curve $\mathbf{r}(t)$ with density $f(x, y, z)$ is given by

$$M = \int_C f \, ds.$$

The center of mass of the same wire has coordinates $(\frac{M_x}{M}, \frac{M_y}{M}, \frac{M_z}{M})$, where

$$M_x = \int_C x f(x, y, z) \, ds, \quad M_y = \int_C y f(x, y, z) \, ds, \quad M_z = \int_C z f(x, y, z) \, ds.$$

Find the mass and the center of mass of a wire in the shape of the helix $x = \cos t$, $y = \sin t$, $z = at$, $0 \leq t \leq \pi$, if the density of the wire at a point (x, y, z) is $f(x, y, z) = x^2 + y^2 + z^2$.

We have $f(\mathbf{r}(t)) = \cos^2 t + \sin^2 t + a^2 t^2 = 1 + a^2 t^2$, $\mathbf{r}'(t) = (-\sin t, \cos t, a)$, $|\mathbf{r}'(t)| = \sqrt{1 + a^2}$. The mass of the wire is

$$M = \int_0^\pi f(\mathbf{r}(t)) |\mathbf{r}'(t)| \, dt = \int_0^\pi (1 + a^2 t^2) \sqrt{1 + a^2} \, dt$$

$$= \sqrt{1+a^2} \left(t + \frac{a^2 t^3}{3} \right) \Big|_0^\pi = \sqrt{1+a^2} \left(\pi + \frac{a^2 \pi^3}{3} \right).$$

To find the center of mass, we evaluate M_x, M_y, M_z as follows:

$$\begin{aligned} M_x &= \int_0^\pi x(t) f(\mathbf{r}(t)) |\mathbf{r}'(t)| dt = \int_0^\pi (\cos t) (1 + a^2 t^2) \sqrt{1 + a^2} dt \\ &= \sqrt{1 + a^2} \left(\sin t \Big|_0^\pi + \int_0^\pi a^2 t^2 \cos t dt \right) \\ &= \sqrt{1 + a^2} \left(a^2 t^2 \sin t \Big|_0^\pi - \int_0^\pi 2a^2 t \sin t dt \right) \\ &= -2a^2 \sqrt{1 + a^2} \int_0^\pi t \sin t dt \\ &= 2a^2 \sqrt{1 + a^2} \left(t \cos t \Big|_0^\pi - \int_0^\pi \cos t dt \right) \\ &= 2a^2 \sqrt{1 + a^2} \left(-\pi - \sin t \Big|_0^\pi \right) = -2a^2 \pi \sqrt{1 + a^2}, \end{aligned}$$

$$\begin{aligned} M_y &= \int_0^\pi y(t) f(\mathbf{r}(t)) |\mathbf{r}'(t)| dt = \int_0^\pi (\sin t) (1 + a^2 t^2) \sqrt{1 + a^2} dt \\ &= \sqrt{1 + a^2} \left(-\cos t \Big|_0^\pi + \int_0^\pi a^2 t^2 \sin t dt \right) \\ &= \sqrt{1 + a^2} \left(2 - a^2 t^2 \cos t \Big|_0^\pi + \int_0^\pi 2a^2 t \cos t dt \right) \\ &= \sqrt{1 + a^2} \left(2 + a^2 \pi^2 + 2a^2 t \sin t \Big|_0^\pi - \int_0^\pi 2a^2 \sin t dt \right) \\ &= \sqrt{1 + a^2} \left(2 + a^2 \pi^2 + 2a^2 \cos t \Big|_0^\pi \right) \\ &= \sqrt{1 + a^2} (2 + a^2 \pi^2 - 4a^2), \end{aligned}$$

$$\begin{aligned} M_z &= \int_0^\pi z(t) f(\mathbf{r}(t)) |\mathbf{r}'(t)| dt = \int_0^\pi at(1 + a^2 t^2) \sqrt{1 + a^2} dt \\ &= \sqrt{1 + a^2} \int_0^\pi (at + a^3 t^3) dt = \sqrt{1 + a^2} \left(\frac{at^2}{2} + \frac{a^3 t^4}{4} \right) \Big|_0^\pi \\ &= \sqrt{1 + a^2} \left(\frac{a\pi^2}{2} + \frac{a^3 \pi^4}{4} \right). \end{aligned}$$

Hence the center of mass is $(\bar{x}, \bar{y}, \bar{z})$, where

$$\bar{x} = \frac{M_x}{M} = \frac{-2a^2}{1 + \frac{a^2 \pi^2}{3}}, \quad \bar{y} = \frac{M_y}{M} = \frac{2 + a^2 \pi^2 - 4a^2}{\pi + \frac{a^2 \pi^3}{3}},$$

$$\bar{z} = \frac{M_z}{M} = \frac{\frac{a\pi^2}{2} + \frac{a^3\pi^4}{4}}{1 + \frac{a^2\pi^2}{3}} = \frac{6a\pi + 3a^3\pi^3}{12 + 4a^2\pi^2}.$$

6. Find $\int_C \mathbf{F} \cdot d\mathbf{r}$ if $\mathbf{F} = z\mathbf{i} + 2y\mathbf{j}$ and $\mathbf{r}(t) = t\mathbf{i} + t^2\mathbf{j} + t^3\mathbf{k}$, for $t \in [0, 2]$.

Along the curve parametrized by $\mathbf{r}(t)$, we have

$$\mathbf{F}(\mathbf{r}(t)) = z(t)\mathbf{i} + 2y(t)\mathbf{j} = t^3\mathbf{i} + 2t^2\mathbf{j}.$$

Moreover,

$$\mathbf{r}'(t) = \mathbf{i} + 2t\mathbf{j} + 3t^2\mathbf{k}.$$

Hence, the line integral is given by

$$\begin{aligned} \int_C \mathbf{F} \cdot d\mathbf{r} &= \int_0^2 (t^3\mathbf{i} + 2t^2\mathbf{j}) \cdot (\mathbf{i} + 2t\mathbf{j} + 3t^2\mathbf{k}) dt = \int_0^2 (t^3 + 4t^3) dt \\ &= \int_0^2 5t^3 dt = \frac{5t^4}{4} \Big|_0^2 \\ &= 20. \end{aligned}$$

7. Find $\int_C \mathbf{F} \cdot d\mathbf{r}$ if $\mathbf{F} = xy\mathbf{i} + 2\mathbf{k}$ and $\mathbf{r}(t) = \cos t\mathbf{i} + \sin t\mathbf{j} + t\mathbf{k}$ for $t \in I = [0, 2\pi]$.

We have

$$\mathbf{F}(\mathbf{r}(t)) = \cos t \sin t \mathbf{i} + 2\mathbf{k}$$

$$\mathbf{r}'(t) = -\sin t \mathbf{i} + \cos t \mathbf{j} + \mathbf{k},$$

so that

$$\begin{aligned} \int_C \mathbf{F} \cdot d\mathbf{r} &= \int_0^{2\pi} [(\cos t \sin t) \cdot (-\sin t) + 2 \cdot 1] dt = \int_0^{2\pi} (-\sin^2 t \cos t + 2) dt \\ &= -\frac{\sin^3 t}{3} + 2t \Big|_0^{2\pi} \\ &= 4\pi. \end{aligned}$$

8. Evaluate the line integral $\int_C -ydx + xdy + x^2dz$, if C is the part of the intersection curve of the elliptic cylinder $x^2 + \frac{y^2}{4} = 1$ and the surface $z = x^3$ that lies in the region $x \geq 0, y \geq 0$, between the points $(1, 0, 1)$ and $(0, 2, 0)$, oriented in the direction of increasing y .

Parametrize the curve by $x = \cos t$, $y = 2 \sin t$, $z = \cos^3 t$, $0 \leq t \leq \pi/2$. (Note that $x = \cos t$, $y = 2 \sin t$ parametrizes the ellipse $x^2 + \frac{y^2}{4} = 1$ in the xy -plane, since $(\cos t)^2 + \frac{1}{4}(2 \sin t)^2 = \cos^2 t + \sin^2 t = 1$.) Then

$$\begin{aligned} \int_C -y dx + x dy + x^2 dz &= \int_0^{\pi/2} \left[(-2 \sin t)(-\sin t) + (\cos t)(2 \cos t) + (\cos^2 t)(-3 \cos^2 t \sin t) \right] dt \\ &= \int_0^{\pi/2} [2 + 3 \cos^4 t(-\sin t)] dt \\ &= 2t + \frac{3}{5} \cos^5 t \Big|_0^{\pi/2} = \pi - \frac{3}{5}. \end{aligned}$$

9. Let $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ be a smooth function with $\nabla f(\mathbf{r}) \neq \mathbf{0}$ for all $\mathbf{r}$. Let $S = \{\mathbf{r} : f(\mathbf{r}) = a\}$ be a level surface of f . Prove that

$$\int_C \nabla f \cdot d\mathbf{r} = 0$$

for any oriented smooth curve C which is contained in S .

We have $\int_C \mathbf{F} \cdot d\mathbf{r} = \int_C \mathbf{F} \cdot \mathbf{T} ds$. If C lies on a level surface of f , then the vector $\mathbf{T}$ tangent to the curve C is also tangent to the surface S . Therefore $\mathbf{T}$ is perpendicular to the vector ∇f which is normal to S . But this means that $\nabla f \cdot \mathbf{T} = 0$ on C , so that $\int_C \nabla f \cdot d\mathbf{r} = \int_C \nabla f \cdot \mathbf{T} ds = 0$.

10. Let $\mathbf{F} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be a smooth vector field. Let C be a smooth curve in $\mathbb{R}^3$. Prove that

$$\left| \int_C \mathbf{F} \cdot d\mathbf{r} \right| \leq \int_C |\mathbf{F}| ds.$$

Since $|\mathbf{F} \cdot \mathbf{T}| = |\mathbf{F}||\mathbf{T}|\cos\theta \leq |\mathbf{F}||\mathbf{T}| = |\mathbf{F}|$, where θ is the angle between $\mathbf{F}$ and $\mathbf{T}$, we get

$$\left| \int_C \mathbf{F} \cdot d\mathbf{r} \right| = \left| \int_C \mathbf{F} \cdot \mathbf{T} ds \right| \leq \int_C |\mathbf{F} \cdot \mathbf{T}| ds \leq \int_C |\mathbf{F}| ds.$$

(The absolute value of an integral is always $\leq$ than the integral of the absolute value of the integrand!)

Practice Problems

1. Evaluate the integral $\int_C \sqrt{x^2 + y^2} ds$, where C is the parametrized curve $\mathbf{r}(t) = (t \cos t, t \sin t)$, $0 \leq t \leq 2\pi$.

2. Evaluate the line integral $\int_C f \, ds$, where $f(x, y, z) = z - x$ and $\mathbf{r}(t) = e^{-t}\mathbf{i} + \sqrt{2}t\mathbf{j} + e^t\mathbf{k}$, $0 \leq t \leq 1$.
3. Evaluate the line integral $\int_C 2yz \, ds$, if C is parametrized by $\mathbf{r}(t) = t\mathbf{i} + 2e^t\mathbf{j} + e^{2t}\mathbf{k}$ for $0 \leq t \leq 1$.
4. Evaluate the line integral $\int_C xy \, dx - xy^2 \, dy$, where C is the curve with equation $x^2 = y^3$, from $(-1, 1)$ to $(1, 1)$.
5. Evaluate the line integral $\int_C x \, dy - y \, dz$, where C is the oriented line segment from $(1, 0, 0)$ to $(3, 1, 6)$.
6. Evaluate the integral $\int_C y \, dx + x \, dy$, where C is the arc of the circle $(x-1)^2 + y^2 = 1$ from $(0, 0)$ to $(1, 1)$ in the first quadrant of the xy -plane.
7. Evaluate the integral $\int_C \mathbf{F} \cdot d\mathbf{r}$, where $\mathbf{F} = y\mathbf{i} - z\mathbf{j}$ and C is the part of the curve of intersection of the surfaces $x + y + z = 2$ and $y = z^2$ from $(0, 1, 1)$ to $(2, 0, 0)$.
8. Evaluate the integral $\int_C \mathbf{F} \cdot d\mathbf{r}$, where $\mathbf{F} = y\mathbf{i} + z\mathbf{j} + x^2\mathbf{k}$ and C is the intersection curve of the surfaces $z = y^2$ and $x^2 + y^2 = 4$, oriented counterclockwise when viewed from the direction of the positive z -axis.
9. Consider the integral $\int_C f \, ds$, where $f(x, y, z) = e^{-(x^2+y^2+z^2)}$ and C is the line where the planes $x + y - z = 0$ and $x + 2y + z = 0$ intersect. Prove that this integral is finite.

(Hint: if you set up this integral using the “usual” parametrization of a line where the velocity is constant, you will get an improper integral. If you do not know how to evaluate it, it will suffice if you use inequalities to prove that the integral is finite.

2.4 Path Independence

We now introduce a fundamental concept for the line integrals of conservative vector fields: *path independence*. We say that a vector field $\mathbf{F}$ has *path-independent line integrals* on a domain $D \subset \mathbb{R}^3$ if for any $P_1, P_2 \in D$, and for any two piecewise smooth curves C_1, C_2 from P_1 to P_2 , we have $\int_{C_1} \mathbf{F} \cdot d\mathbf{r} = \int_{C_2} \mathbf{F} \cdot d\mathbf{r}$.

Proposition 2.4.1 (Path Independence of Line Integrals). *Suppose that C is a smooth curve parametrized by $\mathbf{r}(t)$ for $t \in [a, b]$, and $D \subset \mathbb{R}^n$ is some domain containing C . Further suppose that $\mathbf{F} : D \rightarrow \mathbb{R}^n$ is a continuous vector field which satisfies $\mathbf{F} = \nabla\phi$ for some scalar function $\phi : D \rightarrow \mathbb{R}$.*

Then,

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \phi(\mathbf{r}(b)) - \phi(\mathbf{r}(a)) \quad (2.4.1)$$

In particular, the line integral of $\mathbf{F}$ over $\mathbf{r}$ depends only upon the value of ϕ at the endpoints of the curve (which are $\mathbf{r}(a)$ and $\mathbf{r}(b)$).

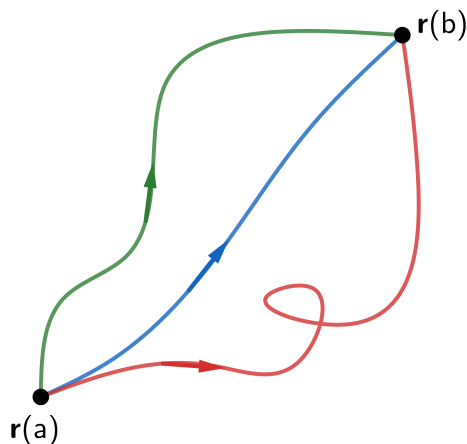


Figure 2.16: We refer to the conclusion (2.4.1) as *path independence*. This is because, regardless of which curve we take from $\mathbf{r}(a)$ to $\mathbf{r}(b)$, we will arrive at the same value for the line integral.

Proof of Proposition 2.4.1. This conclusion follows from the Chain Rule:

$$\begin{aligned} \int_C \mathbf{F} \cdot d\mathbf{r} &= \int_C (\nabla \phi) \cdot d\mathbf{r} = \int_a^b \left[\phi_x(t)x'(t) + \phi_y(t)y'(t) + \phi_z(t)z'(t) \right] dt \\ &= \int_a^b \frac{d}{dt} \phi(\mathbf{r}(t)) dt \\ &= \phi(\mathbf{r}(b)) - \phi(\mathbf{r}(a)) \end{aligned}$$

□

Remark 2.4.1. The conclusion of Proposition 2.4.1 continues to hold if C is a piecewise smooth curve (the union of several smooth arcs joined at their endpoints to form a single curve). To prove this, apply the proposition to each smooth arc, and sum up what you get!

In general, there is no equivalent of Proposition 2.4.1 for non-conservative vector fields, and line integrals may well depend on the shape of the curve rather than just on its endpoints.

Example 2.4.2. Suppose that $\mathbf{F}(x, y) = -y\mathbf{i} + x\mathbf{j}$ and let C_1 and C_2 be as in Figure 2.18. Along C_1 , we are travelling in the opposite direction as $\mathbf{F}$; hence, we have:

$$\int_{C_1} \mathbf{F} \cdot d\mathbf{r} < 0.$$

However, along C_2 , we travel in the same direction as $\mathbf{F}$; hence, we have:

$$\int_{C_2} \mathbf{F} \cdot d\mathbf{r} > 0.$$

And yet, C_1 and C_2 have the same initial and terminal points. This shows that $\mathbf{F}$ does not have the path independence property.

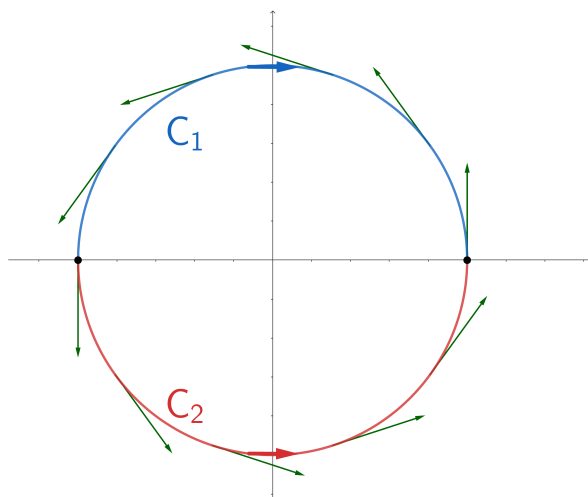


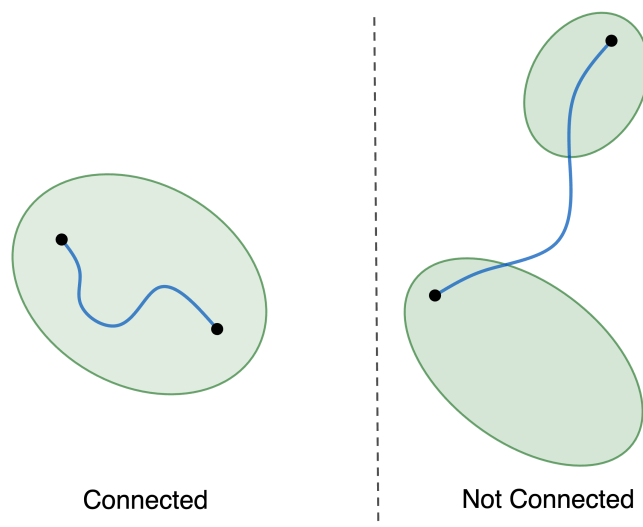
Figure 2.17: The vector field $\mathbf{F}$ (represented by green vectors) along with the two curves C_1 (in blue) and C_2 (in red). The large arrows represent the direction of travel along the curves C_1 and C_2 .

The previous example leads us to ask the following foundational question for line integrals of vector fields on $\mathbb{R}^n$.

Question 2.4.3. Suppose that $\mathbf{F}$ has path-independent line integrals, in the sense that given any two points P_1, P_2 in the domain of $\mathbf{F}$, the integral $\int_C \mathbf{F} \cdot d\mathbf{r}$ has the same value for all paths C from P_1 to P_2 . Does this necessarily imply that $\mathbf{F}$ is conservative?

As we shall prove below, this is indeed true under reasonable assumptions on the domain. We assume (for now) that $D \subset \mathbb{R}^n$ is open. We also need our domains to admit enough paths along which we can integrate our vector field. This is the property below.

Definition 2.4.4. We say that $D \subset \mathbb{R}^n$ is **connected** if every pair of points in D can be joined by a piecewise smooth curve contained in D .



It is very easy to think up/draw connected and not connected sets. The following picture gives you an idea of what they should intuitively look like.

Caution: the definition above is a little bit different from the topological concept of connectedness that you may see elsewhere. For example, a fractal curve like Koch's snowflake is connected in the topological sense, but is not connected in the sense of Definition 2.4.4 because fractal curves are not piecewise smooth. For our purposes, the condition that we will use is the one in Definition 2.4.4 and not the topological one.

We can now give an answer to the question above.

Theorem 2.4.5. *Let $D \subset \mathbb{R}^n$ be an open, connected domain, and let $\mathbf{F}$ be a smooth vector field on D . Then, the following statements are equivalent:*

- (a) $\mathbf{F}$ is conservative on D ;
- (b) $\oint_C \mathbf{F} \cdot d\mathbf{r} = 0$ for all piecewise smooth, closed curves in D ;
- (c) Given any two points $P_1, P_2 \in D$, the line integral $\int_C \mathbf{F} \cdot d\mathbf{r}$ has the same value for all piecewise smooth curves C from P_1 to P_2 .

We use the symbol “ $\oint$ ” to denote a line integral along a closed curve (that is, the initial and terminal points are the same).

In the remainder of this section, we prove Theorem 2.4.5.

Proof of Theorem 2.4.5. We prove the following implications: (a) $\Rightarrow$ (b), (b) $\Rightarrow$ (c), then (c) $\Rightarrow$ (a).

(a) $\Rightarrow$ (b): This is a special case of Proposition 2.4.1: for closed paths, we have $\mathbf{r}(a) = \mathbf{r}(b)$ so that $\oint_C \mathbf{F} \cdot d\mathbf{r} = \phi(\mathbf{r}(b)) - \phi(\mathbf{r}(a)) = 0$.

(b) $\Rightarrow$ (c): Let C_1 and C_2 be two curves from connecting P_1 to P_2 , both entirely contained in D . We then form a closed curve, by joining C_1 to C_2 , and reversing the orientation of C_1 . This is shown in the two figures below.

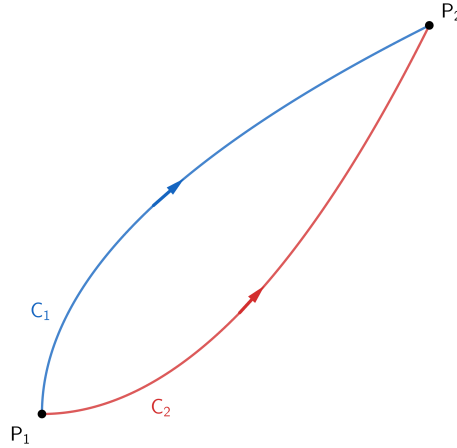


Figure 2.18: The two curves C_1 and C_2 .

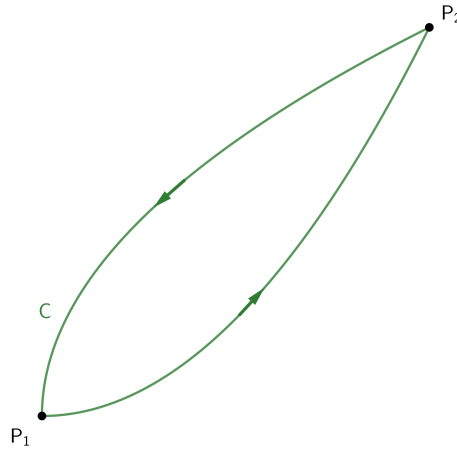


Figure 2.19: The closed curve obtained by reversing the orientation of C_1 .

Call this new curve C ; notice that C is a closed curve. Recall that we are assuming (b) holds, so that

$$0 = \oint_C \mathbf{F} \cdot d\mathbf{r} = \int_{C_2} \mathbf{F} \cdot d\mathbf{r} - \int_{C_1} \mathbf{F} \cdot d\mathbf{r}.$$

The minus sign on the second integral comes from reversing the orientation of C_1 to form C . This proves (c).

(c) $\Rightarrow$ (a): We now assume that $\mathbf{F}$ has the path independence property on D . Fixing some $P_0 = (x_0, y_0, z_0)$, define a scalar valued function $\phi(x, y, z)$ as

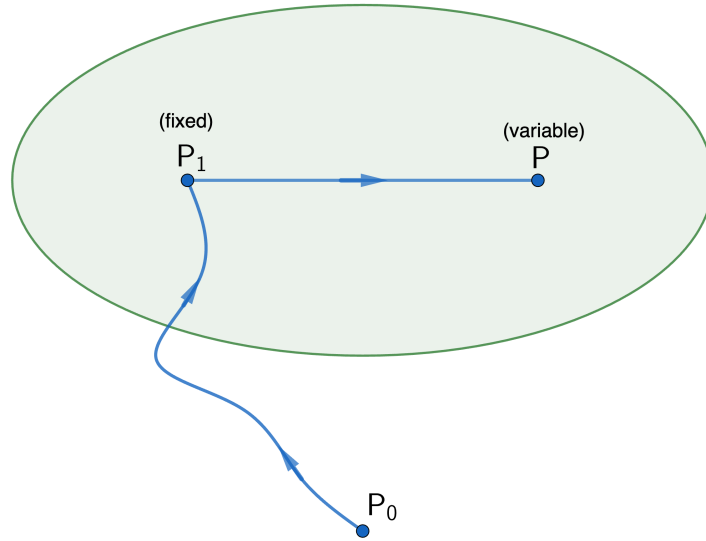
$$\phi(x, y, z) = \int_C \mathbf{F} \cdot d\mathbf{r}, \quad (2.4.2)$$

where C is any path connecting P_0 to $P = (x, y, z)$. Notice that the value of ϕ does not depend upon the choice of path – this follows since we assume the path independence property for $\mathbf{F}$.

Our goal is now to show that $\nabla\phi = \mathbf{F}$; in other words, we show that

$$\underbrace{\frac{\partial\phi}{\partial x}}_{(*)} = F_1, \quad \frac{\partial\phi}{\partial y} = F_2, \quad \frac{\partial\phi}{\partial z} = F_3. \quad (2.4.3)$$

We will only prove (*). The other two identities are proved in a similar fashion, with the x, y, z coordinates interchanged as needed.



Let's say we want to prove (*) at some $P_1 = (x_1, y_1, z_1) \in D$. This means that we want to evaluate

$$\frac{\partial\phi}{\partial x}(P_1) = \lim_{x \rightarrow x_1} \frac{\phi(x, y_1, z_1) - \phi(x_1, y_1, z_1)}{x - x_1}. \quad (2.4.4)$$

Let $P = (x, y_1, z_1)$, where x is close enough to x_1 so that $P \in D$. (This is possible, because we assumed our domain D was an open set so that it contains a neighbourhood

of P !). Recalling the definition of ϕ in (2.4.2), we now choose the paths we will use. We first fix a piecewise smooth path C_1 from P_0 to P_1 , so that

$$\phi(P_1) = \int_{C_1} \mathbf{F} \cdot d\mathbf{r}.$$

For the variable point P in (2.4.4), we let C be the union of C_1 from P_0 to P_1 and the line segment C_2 from P_1 to P , so that

$$\phi(P) := \int_C \mathbf{F} \cdot d\mathbf{r} = \int_{C_1} \mathbf{F} \cdot d\mathbf{r} + \int_{C_2} \mathbf{F} \cdot d\mathbf{r}.$$

This gets us

$$\phi(P) - \phi(P_1) = \int_{C_2} \mathbf{F} \cdot d\mathbf{r}.$$

To evaluate the last integral, we parametrize C_2 using

$$\mathbf{r}(t) := \langle x_1 + t, y_1, z_1 \rangle, \quad (2.4.5)$$

so that $\mathbf{r}(0) = P_1$, $\mathbf{r}(x - x_1) = P$, and $\mathbf{r}'(t) = \langle 1, 0, 0 \rangle$. Using our line integral formula from Section 2.3, we see that

$$\int_{C_2} \mathbf{F} \cdot d\mathbf{r} = \int_0^{x-x_1} F_1(x_1 + t, y_1, z_1) dt.$$

(Recall: if $a < 0$, then $\int_0^a = -\int_a^0$!) Putting everything together, and using the Fundamental Theorem of Calculus, we get

$$\lim_{x \rightarrow x_1} \frac{\phi(P) - \phi(P_1)}{x - x_1} = \frac{\partial}{\partial x} \int_{C_2} \mathbf{F} \cdot d\mathbf{r} \Big|_{x=x_1} = F_1(x_1, y_1, z_1).$$

Hence $\frac{\partial \phi}{\partial x}(P_1) = F_1(P_1)$, as claimed.

This proves the first identity in (2.4.3). The proofs of the remaining two identities are similar, but with the y and z variables playing the role of the x variable in the above proof. The path independence assumption is used here in an important way: to prove the second and third identity in (2.4.3), we need to use paths that end in line segments parallel, respectively, to the y -axis and z -axis. Path independence guarantees that the integrals on these very different paths give us the same function ϕ . $\square$

The theorem can be extended in various ways. In many of our applications, $D \subset \mathbb{R}^n$ will be a set of the form $D = S \cup B$, where S is an open set and B is a set of some or all of the boundary points of S . The theorem extends to this case: define ϕ as above but on all of D , prove that ϕ is continuous on D and that $\mathbf{F} = \nabla \phi$ on S . Additionally, (b) could be replaced by the (formally weaker but actually equivalent) condition that $\oint_C \mathbf{F} \cdot d\mathbf{r} = 0$ for all piecewise smooth, closed, and simple curves in D . If the domain is not connected but instead consists of several connected parts, we may apply the theorem on each part separately.

Worked Problems

1. Find a function $F_3(x, y, z)$ such that the vector field

$$\mathbf{F}(x, y, z) = (y \cos z - yz \cos x)\mathbf{i} + (x \cos z - z \sin x)\mathbf{j} + F_3(x, y, z)\mathbf{k}$$

is conservative on $\mathbb{R}^3$. With that choice of F_3 , find $\int_C \mathbf{F} \cdot d\mathbf{r}$, where C is a smooth curve from $(0, 1, \pi)$ to $(\pi/2, 4, 0)$.

Suppose $\mathbf{F} = \nabla f$. Then $f_x = F_1 = y \cos z - yz \cos x$, so that

$$f(x, y, z) = xy \cos z - yz \sin x + g(y, z).$$

Differentiating this in y , we get

$$f_y = x \cos z - z \sin x + g_y.$$

This matches F_3 when $g_y = 0$, so we can take $g(y, z) = h(z)$ for any function $h(z)$ independent of y . The simplest example is when $g \equiv 0$, in which case $f(x, y, z) = xy \cos z - yz \sin x$ and

$$F_3 = f_z = -xy \sin z - y \sin x.$$

Finally,

$$\begin{aligned} \int_C \mathbf{F} \cdot d\mathbf{r} &= f(\pi/2, 4, 0) - f(0, 1, \pi) \\ &= \frac{\pi}{2} \cdot 4 \cos 0 - 4 \cdot 0 \cdot \sin \frac{\pi}{2} - 0 - 1 \cdot \pi \cdot \sin 0 = 2\pi. \end{aligned}$$

Practice Problems

- Let $\mathbf{F}(x, y, z) = e^{2y}\mathbf{i} + (axe^{2y} + be^z)\mathbf{j} - 3ye^z\mathbf{k}$. Find the values of a, b for which $\mathbf{F}$ is conservative, and the corresponding potential function. For these values of a and b , evaluate $\int_C \mathbf{F} \cdot d\mathbf{r}$, where C is a piecewise smooth curve from $(1, 1, 0)$ to $(0, 1, 3)$.
- Find all values of the parameters a, b, c such that the vector field

$$\mathbf{F}(x, y, z) = 4xy^3\mathbf{i} + (ax^2y^2 + b \cos z)\mathbf{j} + (y \sin z + c \cos z)\mathbf{k}$$

is conservative on $\mathbb{R}^3$. For all such a, b, c :

- find a potential function f such that $\mathbf{F} = \nabla f$,
- evaluate the integral $\int_C \mathbf{F} \cdot d\mathbf{r}$, if C is the line segment from $(0, 1, 0)$ to $(1, 1, \pi/2)$.

3. Let $\mathbf{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be a C^1 vector field. Assume that the following holds: for any $P, Q \in \mathbb{R}^2$, and for any two curves C_1, C_2 from P to Q **that are unions of at most 5 line segments each**, we have $\int_{C_1} \mathbf{F} \cdot d\mathbf{r} = \int_{C_2} \mathbf{F} \cdot d\mathbf{r}$. Does it follow that $\mathbf{F}$ is conservative? Prove your answer.
4. (8 marks) Let f and g be two C^1 functions on D . Let $\mathbf{F} = f\nabla g$ and $\mathbf{G} = g\nabla f$. Prove that $\mathbf{F}$ has path-independent line integrals if and only if so does $\mathbf{G}$.

Chapter 3

Parametric Surfaces and Surface Integration

3.1 Surfaces and Area Integrals

The next selection of topics focuses on **parametric surfaces**. In $\mathbb{R}^3$, these will be given by continuous, vector-valued functions $\mathbf{r} : D \rightarrow \mathbb{R}^3$, where $D \subset \mathbb{R}^2$ is a subset of *two-dimensional* space. Usually, we will assume that D is the closure of an open set. In terms of components, we write

$$\mathbf{r}(u, v) = x(u, v)\mathbf{i} + y(u, v)\mathbf{j} + z(u, v)\mathbf{k}, \quad (u, v) \in D.$$

The **surface** $\mathcal{S} \subset \mathbb{R}^3$ is the *range* of the vector-valued function $\mathbf{r}$. We say that $\mathbf{r}$ is a parametrization of $\mathcal{S}$.

Example 3.1.1. Suppose that $f : D \rightarrow \mathbb{R}$ is a scalar function on some set $D \subset \mathbb{R}^2$. Letting $z = f(x, y)$, the graph of f can be written as a parametric surface, with the (x, y) variables as the parameters:

$$\mathbf{r}(x, y) = x\mathbf{i} + y\mathbf{j} + f(x, y)\mathbf{k}. \quad (3.1.1)$$

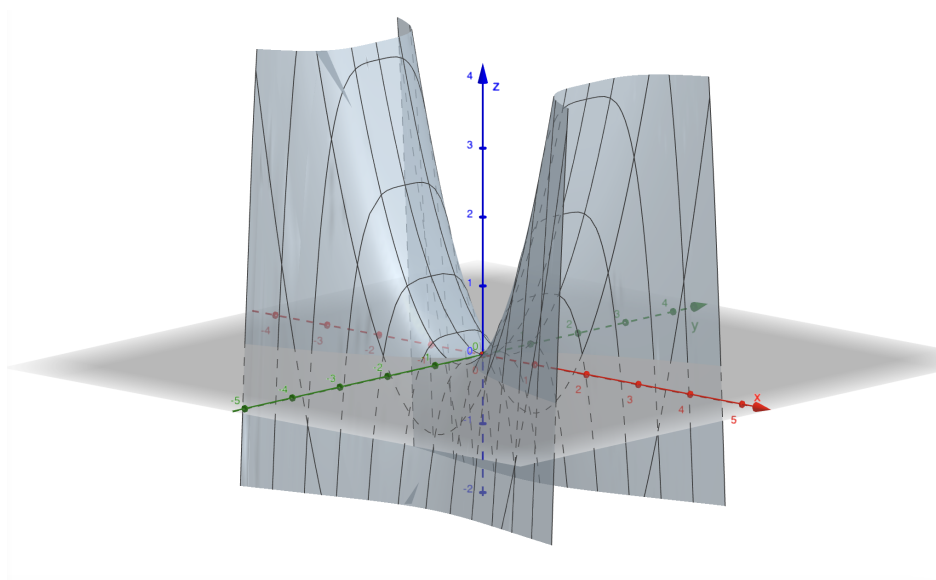
Example 3.1.2. The sphere $x^2 + y^2 + z^2 = a^2$ (for $a > 0$) can be parametrized using spherical coordinates (θ, ϕ) as

$$x = a \cos \theta \sin \phi, \quad y = a \sin \theta \sin \phi, \quad z = a \cos \phi,$$

where $(\theta, \phi) \in D$ live in the domain

$$D := \{(\theta, \phi) : 0 \leq \theta \leq 2\pi, 0 \leq \phi \leq \pi\}.$$

We could also use combinations of spherical or cylindrical coordinates to parametrize cylinders, cones, etc.

Figure 3.1: An example of a surface in $\mathbb{R}^3$.

We will usually assume that the function $\mathbf{r}$ and all its derivatives are continuous as needed. We will also need a further condition, as follows. Let $P = \mathbf{r}(u_0, v_0)$ be a point on $\mathcal{S}$. Notice that the equation

$$\mathbf{r}(u, v_0) = x(u, v_0)\mathbf{i} + y(u, v_0)\mathbf{j} + z(u, v_0)\mathbf{k},$$

and $\mathbf{r}(u_0, v)$ (defined similarly), are parametric equations for two curves $\mathcal{C}_u$ and $\mathcal{C}_v$ (respectively) that are contained in $\mathcal{S}$ and intersect at P . At the point P , the curves $\mathcal{C}_u$ and $\mathcal{C}_v$ have their associated tangent vectors

$$\mathbf{r}_u = \frac{\partial \mathbf{r}}{\partial u}(u_0, v_0) \quad \text{and} \quad \mathbf{r}_v = \frac{\partial \mathbf{r}}{\partial v}(u_0, v_0).$$

Our additional assumption is that $\mathbf{r}_u$ and $\mathbf{r}_v$ are not parallel, except possibly when (u_0, v_0) belongs to the boundary of D . We write this as

$$\mathbf{r}_u \times \mathbf{r}_v \neq \mathbf{0}. \quad (3.1.2)$$

To see what could go wrong if (3.1.2) fails, consider the vector-valued function $\mathbf{r}(u, v) = (u + 2v)\mathbf{i} + (u + 2v)\mathbf{j} + (u + 2v)\mathbf{k}$, with $\mathbf{r}_v = 2\mathbf{r}_u$. Then $x(u, v) = y(u, v) = z(u, v)$, so that $\mathbf{r}$ parametrizes the line $x = y = z$ rather than a surface. On the other hand, if (3.1.2) does hold, then the curves on $\mathcal{S}$ corresponding to fixed values of u and v form a nice grid on $\mathcal{S}$ as shown in the illustrations in this section.

Definition 3.1.3. Assume that $D \subseteq \mathbb{R}^2$ is the closure of an open set U , and that its boundary consists of finitely many piecewise smooth curves.

(a) We say that $\mathbf{r} : D \rightarrow \mathbb{R}^3$ is a **smooth parametrization** of its image $\mathcal{S}$ if $\mathbf{r}$ is C^1 on D , one-to-one on U , and (3.1.2) holds on U .

- (b) A surface $\mathcal{S}$ is **smooth** if it admits a smooth parametrization.
- (c) A surface is **piecewise smooth** if it consists of a finite number of smooth surfaces joined together along their common boundaries.
- (d) A surface is **closed** if it bounds a closed and bounded region in $\mathbb{R}^3$.

If $\mathbf{r}$ is defined only on D (and not on some neighbourhood of it), then the partial derivatives of $\mathbf{r}$ may not be well defined on its boundary. In that case, we understand the C^1 condition to mean that the first order partials of D (initially defined on the interior of D) extend to continuous functions on all of D including the boundary.

We will see below that a sphere, and the graph of a C^1 function $z = g(x, y)$, are smooth surfaces. Additionally, a sphere in $\mathbb{R}^3$ is a closed surface since it bounds a ball in 3 dimensions.

3.1.1 Normal Vector for Parametric Surfaces

We start with the formulas to normal vectors and tangent planes for parametric surfaces. Suppose that $\mathbf{r} : D \rightarrow \mathbb{R}^3$ is a smooth parametrization of $\mathcal{S}$. Let $P = (x_0, y_0, z_0) \in \mathcal{S}$, so that $P = \mathbf{r}(u_0, v_0)$ for some fixed values $(u_0, v_0) \in D$. Define the curves $\mathcal{C}_u$ and $\mathcal{C}_v$, and the tangent vectors $\mathbf{r}_u(u_0, v_0)$ and $\mathbf{r}_v(u_0, v_0)$, as above. The normal vector at P should then be perpendicular to these two vectors. Assuming (3.1.2), we may use

$$\boxed{\mathbf{n} = \mathbf{r}_u \times \mathbf{r}_v} \quad (3.1.3)$$

as the equation for our normal vector at P . Since $\mathbf{r}_u$ and $\mathbf{r}_v$ are functions of $(u, v) \in D$, the vector $\mathbf{n} = \mathbf{n}(u, v)$ is a vector-valued function on D .

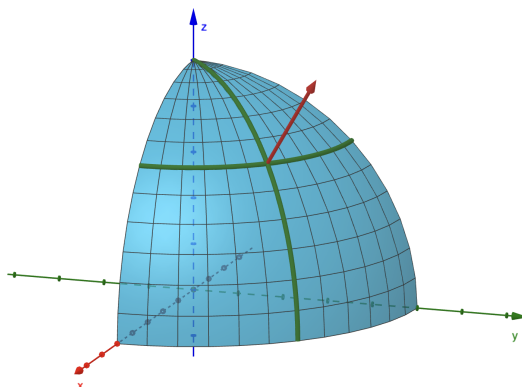


Figure 3.2: A surface (in blue with guide-lines) parametrized by $\mathbf{r}(u, v)$ with two curves $\mathbf{r}(u, v_0)$ and $\mathbf{r}(u_0, v)$ given in green. The normal vector $\mathbf{n}$ at the point of intersection is perpendicular to both of the tangent vectors to these green curves.

Once we have found the normal vector $\mathbf{n}$, we can proceed with the equation of the tangent plane. We just write the equation of the plane that passes through $P = (x_0, y_0, z_0)$ and is perpendicular to $\mathbf{n} = \langle n_1, n_2, n_3 \rangle$:

$$n_1(x - x_0) + n_2(y - y_0) + n_3(z - z_0) = 0.$$

3.1.2 The Area Element

We also want to calculate the surface area element of parametric surfaces—again, in terms of the parametrizing function $\mathbf{r}$. We will use it to evaluate the area of surfaces and to set up surface integrals.

Assume that $\mathcal{S}$ is parametrized by some function $\mathbf{r}(u, v)$ on $D \subset \mathbb{R}^2$. We start by subdividing $\mathcal{S}$ using the (not necessarily rectangular) grid which consists of the curves $u = \text{constant}$ and $v = \text{constant}$. This has the effect of dividing $\mathcal{S}$ into approximate parallelograms, spanned by vectors $\mathbf{r}_u \Delta u$ and $\mathbf{r}_v \Delta v$. (We used a very similar argument in deriving the change-of-variables formula in Math 226.) This situation is illustrated in Figure 3.3 below.

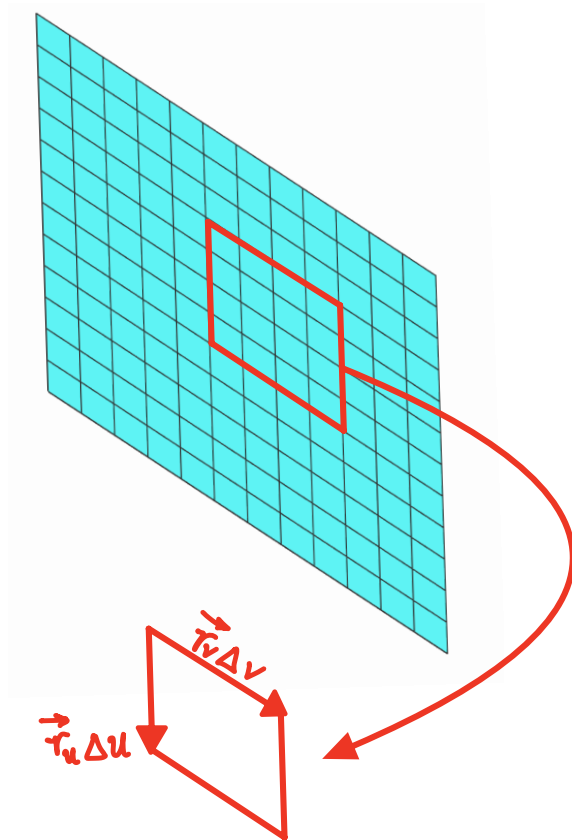


Figure 3.3: The area of the indicated parallelogram is $|\mathbf{n}| \Delta u \Delta v$.

The area of such a parallelogram can be calculated as

$$\begin{aligned} dS &= |(\mathbf{r}_u \Delta u) \times (\mathbf{r}_v \Delta v)| = |\mathbf{r}_u \times \mathbf{r}_v| \Delta u \Delta v \\ &= |\mathbf{n}| \Delta u \Delta v, \end{aligned}$$

where $\mathbf{n}$ is the normal vector from the previous subsection. This delivers the following **integration formula**:

$$\boxed{\iint_{\mathcal{S}} f(x, y, z) dS = \iint_D f(\mathbf{r}(u, v)) |\mathbf{r}_u \times \mathbf{r}_v| du dv} \quad (3.1.4)$$

where dS denotes integration with respect to surface area.

The integration formula (3.1.4) shows us how to calculate some common quantities, such as:

- The area of $\mathcal{S}$ is given by

$$(\text{Area of } \mathcal{S}) = \iint_{\mathcal{S}} 1 dS = \iint_D |\mathbf{r}_u \times \mathbf{r}_v| du dv$$

- If $\mathcal{S}$ is a surface made of a material of variable density $f(x, y, z)$, then the integral $\iint_{\mathcal{S}} f(x, y, z) dS$ is the *total mass* of $\mathcal{S}$.
- The **average value** of a function f over the surface $\mathcal{S}$ is given by

$$\frac{1}{(\text{Area of } \mathcal{S})} \iint_{\mathcal{S}} f(x, y, z) dS = \frac{\iint_{\mathcal{S}} f(x, y, z) dS}{\iint_{\mathcal{S}} 1 dS}$$

- The **centroid** of $\mathcal{S}$ has coordinates $(\bar{x}, \bar{y}, \bar{z})$ where

$$\bar{x} = \frac{1}{(\text{Area of } \mathcal{S})} \iint_{\mathcal{S}} x dS, \text{ and similarly for } \bar{y}, \bar{z}.$$

Finally, it is very important that we realize: if $\mathbf{n} = \mathbf{r}_u \times \mathbf{r}_v$ is the normal vector found above, then the vectors $\lambda \mathbf{n}$ are also normal vectors, for any $\lambda \in \mathbb{R} \setminus \{0\}$. But, for the purpose of *integration*, we must use $|\mathbf{n}| = |\mathbf{r}_u \times \mathbf{r}_v|$ with *no rescaling*.

We now give several more specialized formulas for important classes of surfaces.

3.1.3 Area of Graphs of Functions

Suppose that $\mathcal{S}$ is the graph of a function $z = g(x, y)$, where $(x, y) \in D$. We parametrize it by

$$\mathbf{r}(x, y) = x\mathbf{i} + y\mathbf{j} + g(x, y)\mathbf{k}.$$

The curves $\mathcal{C}_x$ and $\mathcal{C}_y$ have the tangent vectors

$$\mathbf{r}_x = \mathbf{i} + g_x\mathbf{k}, \quad \mathbf{r}_y = \mathbf{j} + g_y\mathbf{k},$$

which produce the normal vector

$$\mathbf{n} = \mathbf{r}_x \times \mathbf{r}_y = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & g_x \\ 0 & 1 & g_y \end{vmatrix} = -g_x\mathbf{i} - g_y\mathbf{j} + \mathbf{k}.$$

Hence

$$dS = |\mathbf{n}| dx dy = \sqrt{1 + g_x^2 + g_y^2} dx dy \quad (3.1.5)$$

Therefore the area of the part of the graph of $z = g(x, y)$ with $(x, y) \in D$ is

$$\iint_{\mathcal{S}} 1 dS = \iint_D \sqrt{1 + g_x^2 + g_y^2} dx dy. \quad (3.1.6)$$

3.1.4 Geometric Interpretation of the Area Element

We continue to work with the same set-up as in Section 3.1.3, and we would like to understand the geometric meaning of the formula (3.1.5). Let γ denote the angle between $\mathbf{n}$ (the normal vector defined above) and $\mathbf{k}$ (the unit vector in the vertical direction). Then, using the definition of the dot product, we have

$$\mathbf{n} \cdot \mathbf{k} = |\mathbf{n}| |\mathbf{k}| \cos \gamma.$$

We have $|\mathbf{k}| = 1$ and, from the computation of $\mathbf{n}$ in Section 3.1.3,

$$\mathbf{n} := -g_x\mathbf{i} - g_y\mathbf{j} + \mathbf{k} \Rightarrow \mathbf{n} \cdot \mathbf{k} = 1.$$

Therefore $|\mathbf{n}| = \frac{1}{\cos \gamma} = \sec \gamma$, and

$$dS = |\mathbf{n}| dx dy = \frac{dx dy}{\cos \gamma} = \frac{dA}{\cos \gamma},$$

where dS and dA denotes the respective area differentials on the surface and on its projection to the xy -plane. This whole situation is illustrated in Figure 3.3. Intuitively, if $\mathbf{n}$ is close to vertical, then $\mathcal{S}$ is close to being horizontal and dS is close to dA ; if on the other hand the angle between $\mathbf{n}$ and $\mathbf{k}$ is close to $\pi/2$ (so that $\cos \gamma$ is close to 0),

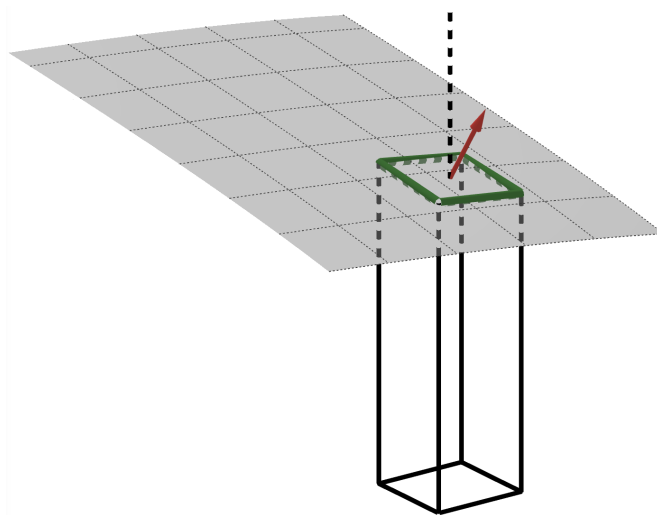


Figure 3.4: The normal vector (in red) makes angle γ with the $\mathbf{k}$ direction. The area element on the surface is represented by the green parallelogram.

then the surface is almost vertical, so that a large part of $\mathcal{S}$ can project to a very small area in the xy -plane.

We can carry this one step further by considering a surface $\mathcal{S}$ given by the equation $G(x, y, z) = 0$. In this scenario, ∇G (if non-zero) is normal to the surface. It does not necessarily have the “right” magnitude for the area element. However, a clever calculation fixes this issue. Notice that

$$|\cos \gamma| = \frac{|\nabla G \cdot \mathbf{k}|}{|\nabla G||\mathbf{k}|} = \frac{|G_z|}{|\nabla G|} = \frac{|G_z|}{\sqrt{G_x^2 + G_y^2 + G_z^2}}.$$

This, in turn, implies that

$$dS = \frac{|\nabla G|}{|G_z|} dxdy = \frac{\sqrt{G_x^2 + G_y^2 + G_z^2}}{|G_z|} dxdy, \text{ if } G_z \neq 0. \quad (3.1.7)$$

If $G_z = 0$ at some point of the domain, it may still be possible to set up and calculate the surface area as an improper integral. Other times, it may be preferable to use some other parametrization of the surface near the troublesome point. For example, if $G_z = 0$ but $G_y \neq 0$ on some part of the surface, we could parametrize that part of it by the x and z variables instead of x and y .

3.1.5 Surface Area for Sphere and Cone

We conclude this section with calculating the surface elements for two “canonical surfaces”.

Example 3.1.4. Let $\mathcal{S}$ be the sphere of radius $a > 0$ centered at the origin, parametrized by (θ, ϕ) as in Example 3.1.2. Then, a calculation (which you should do!) shows that

$$|\mathbf{n}| = |\mathbf{r}_\theta \times \mathbf{r}_\phi| = a^2 \sin \phi. \quad (3.1.8)$$

Note the similarity to the Jacobian $r^2 \sin \phi$ which appears when we change from Cartesian to Spherical Coordinates.

The identity (3.1.8) is useful when calculating integrals over the sphere. For example,

$$\begin{aligned} \iint_{\{x^2+y^2+z^2=1\}} z^2 dS &= \int_0^{2\pi} \int_0^\pi \underbrace{\cos^2 \phi}_{z=\cos \phi} \underbrace{\sin \phi d\phi d\theta}_{(*)} \\ &= \int_0^{2\pi} \left[-\frac{\cos^3 \phi}{3} \right]_0^\pi d\theta = \int_0^{2\pi} \frac{2}{3} d\theta = \frac{4\pi}{3}. \end{aligned}$$

The expression $(*)$ is the area element with $a = 1$.

Example 3.1.5. Consider the positive half-cone symmetric about the z -axis,

$$z = b\sqrt{x^2 + y^2}, \quad b > 0.$$

We use the parametrization based on cylindrical coordinates,

$$x(\theta, t) = t \cos \theta, \quad y(\theta, t) = t \sin \theta, \quad z(\theta, t) = bt,$$

for $0 \leq \theta \leq 2\pi$ and $t \geq 0$. Then

$$\mathbf{r}_t = \langle \cos \theta, \sin \theta, 1 \rangle, \quad \mathbf{r}_\theta = \langle -t \sin \theta, t \cos \theta, 0 \rangle,$$

so that

$$\mathbf{n} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \cos \theta & \sin \theta & 1 \\ -t \sin \theta & t \cos \theta & 0 \end{vmatrix} = -bt \cos \theta \mathbf{i} - bt \sin \theta \mathbf{j} + t \mathbf{k}.$$

Hence,

$$|\mathbf{n}| = \left(b^2 t^2 (\cos^2 \theta + \sin^2 \theta) + t^2 \right)^{1/2} = t\sqrt{b^2 + 1}.$$

We use this to calculate integrals over the cone. For instance, if $\mathcal{S}$ is the conical segment

$$z = b\sqrt{x^2 + y^2}, \quad x^2 + y^2 \leq a^2,$$

for some $a, b > 0$, we have

$$\begin{aligned}\iint_S z^3 dS &= \int_0^a \int_0^{2\pi} \underbrace{(bt)^3}_z \underbrace{(t\sqrt{b^2+1})}_{dS} d\theta dt = 2\pi b^3 \sqrt{b^2+1} \int_0^a t^4 dt \\ &= 2\pi b^3 \sqrt{b^2+1} \left. \frac{t^5}{5} \right|_0^a \\ &= 2\pi b^3 \sqrt{b^2+1} \frac{a^5}{5}.\end{aligned}$$

Worked Problems

1. Find the area of the part of the plane $2x - y + z = 10$ that lies above the disc $(x-3)^2 + y^2 \leq 1$.

We consider our plane as a parametric surface, parametrized by

$$\mathbf{r}(x, y) = x\mathbf{i} + y\mathbf{j} + (10 - 2x + y)\mathbf{k}.$$

Then

$$\mathbf{r}_x = \mathbf{i} - 2\mathbf{k}, \quad \mathbf{r}_y = \mathbf{j} + \mathbf{k},$$

so that

$$\mathbf{n} = \mathbf{r}_x \times \mathbf{r}_y = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & -2 \\ 0 & 1 & 1 \end{vmatrix} = 2\mathbf{i} - \mathbf{j} + \mathbf{k} \Rightarrow |\mathbf{n}| = \sqrt{6}.$$

Hence, we have

$$(\text{Area of } \mathcal{S}) = \sqrt{6} \cdot (\text{Area of disc}) = \sqrt{6}\pi.$$

(For a more general plane $Ax + By + Cz = D$ with $C \neq 0$, a similar calculation shows that

$$dS = \frac{\sqrt{A^2 + B^2 + C^2}}{|C|} dxdy.$$

The area element is a constant multiple of $dxdy$ because the normal vector to the plane makes a constant angle with the z -axis. A good exercise is to try to work out the details.)

2. Find the area of the part of the paraboloid $z = x^2 + y^2$ that lies below the plane $z = 9$.

Call the surface described above $\mathcal{P}$. In particular, if $(x, y, z) \in \mathcal{P}$, then we must have $z \leq 9$, and also $x^2 + y^2 \leq 9$. Set $g(x, y) = x^2 + y^2$ with associated domain $D := \{(x, y) : x^2 + y^2 \leq 9\}$. Then $\mathcal{P}$ is the graph of g . Using that $g_x = 2x$ and $g_y = 2y$, formula (3.1.6) gives

$$(\text{Area of } \mathcal{P}) = \iint_D |\mathbf{n}| dx dy = \iint_{\{x^2+y^2 \leq 9\}} \sqrt{1+4x^2+4y^2} dx dy$$

We evaluate this integral in polar coordinates

$$\int_0^{2\pi} \int_0^3 \sqrt{1+4r^2} r dr d\theta = \int_0^{2\pi} \frac{1}{12} (1+4r^2)^{3/2} \Big|_0^3 d\theta = \underbrace{\frac{\pi}{6} (37^{3/2} - 1)}_{\text{Area of } \mathcal{P}}$$

3. Assume that $\mathcal{P}$ is the paraboloid defined in the previous problem. Calculate $\iint_{\mathcal{P}} z dS$.

This is a modification of the previous example. We have

$$\begin{aligned} \iint_{\mathcal{P}} z dS &= \iint_D \underbrace{(x^2 + y^2)}_{z(x,y)} \cdot \underbrace{\sqrt{1+4x^2+4y^2} dx dy}_{dS} = \int_0^{2\pi} \int_0^3 r^2 \cdot \sqrt{1+4r^2} r dr d\theta \\ &= \frac{\pi}{8} \left[\frac{37^{5/2}}{5} - \frac{37^{3/2}}{3} - \frac{1}{5} + \frac{1}{3} \right]. \end{aligned}$$

The last calculation can be done using the substitution $u = 1 + 4r^2$. (Exercise: complete the details!)

4. A parametric surface is given by the equations $x = u^2, y = uv, z = \frac{1}{2}v^2$, with $u, v \geq 0$.

(a) Find the equation of the tangent plane to the surface at the point P where $x = 4, y = 4, z = 2$.

We have $\mathbf{r}_u = \langle 2u, v, 0 \rangle$ and $\mathbf{r}_v = \langle 0, u, v \rangle$, so that the normal vector is

$$\mathbf{n} = \mathbf{r}_u \times \mathbf{r}_v = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2u & v & 0 \\ 0 & u & v \end{vmatrix} = v^2 \mathbf{i} - 2uv \mathbf{j} + 2u^2 \mathbf{k},$$

The point P corresponds to $u^2 = 4, uv = 4, v^2 = 4$, so that the normal vector at P is $\mathbf{n} = 4\mathbf{i} - 8\mathbf{j} + 8\mathbf{k}$. (Optional: we can also solve these equations to get $u = v = 2$.) Therefore the tangent plane is

$$4(x - 4) - 8(y - 4) + 8(z - 2) = 0,$$

which can be simplified to $(x - 4) - 2(y - 4) + 2(z - 2) = 0$, $x - 2y + 2z = 0$.

(b) Find the area of the part of the surface where $0 \leq u \leq 1$ and $0 \leq v \leq 2$.

Using $\mathbf{n}$ from part (a), we have

$$|\mathbf{n}| = \sqrt{v^4 + 4u^2v^2 + 4u^4} = v^2 + 2u^2.$$

The area is

$$\begin{aligned} \int_0^2 \int_0^1 (v^2 + 2u^2) du dv &= \int_0^2 \left[uv^2 + \frac{2u^3}{3} \right]_{u=0}^1 dv \\ &= \int_0^2 \left(v^2 + \frac{2}{3} \right) dv = \left[\frac{v^3}{3} + \frac{2v}{3} \right]_{v=0}^2 = \frac{8}{3} + \frac{4}{3} = 4. \end{aligned}$$

Practice Problems

- Evaluate the surface integral $\iint_{\mathcal{S}} \sqrt{x} dS$, if $\mathcal{S}$ is the parametrized surface $\mathbf{r}(u, v) = (u^2, v, u + v)$ with $0 \leq u \leq 1$, $0 \leq v \leq 1$.
- A parametric surface is given by the equations $x = uv$, $y = u^2 - v^2$, $z = u^2 + v^2$, with $u, v \geq 0$. Find the equation of the plane tangent to the surface at the point P where $(u, v) = (2, 1)$.
- Let $\mathcal{S}$ be the surface $\mathbf{r}(u, v) = \langle u \cos v, u \sin v, v \rangle$ with $0 \leq u \leq 1$, $0 \leq v \leq \pi$.
 - Find the tangent plane to the surface at the point where $u = 1/2$, $v = \pi/6$.
 - Evaluate the surface integral $\iint_{\mathcal{S}} (x^2 + y^2)^{1/2} dS$.
- Let $\mathcal{S}$ be the surface parametrized by $\mathbf{r}(u, v) = (u + \sin u, v + \sin v, 1 - \cos(u + v))$, $0 < u < \pi$, $0 < v < \pi$ (the boundary of the rectangle is not included).
 - Find all (u, v) such that the tangent plane to the surface at $\mathbf{r}(u, v)$ is parallel to the xy -plane.
 - Prove that the area of $\mathcal{S}$ is less than $5\pi^2$.
- Evaluate the integral $\iint_{\mathcal{S}} (x^2 + 4y^2 + 1) dS$, if $\mathcal{S}$ is the part of the paraboloid $z = \frac{x^2}{2} + y^2$ that lies above the region $D = \{(x, y) : x^2 + 4y^2 \leq 1 \text{ and } y \geq 0\}$ in the xy -plane.

6. Let $\mathcal{C}$ be a smooth simple curve in the plane $z = 0$, parametrized by $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$ with $a \leq t \leq b$. Construct a surface $\mathcal{S}$ in three dimensions by taking the union of all line segments that connect points $\mathbf{r}(t)$ on the curve to the fixed point $P = (0, 0, 1)$. Prove that

$$\frac{1}{2}(\text{length of } \mathcal{C}) \leq (\text{area of } \mathcal{S}).$$

Give an example of $\mathcal{C}$ such that equality holds in the above formula.

(Hint: How would you parametrize the surface? You may want to start by parametrizing each line segment separately.)

7. Let D be a planar domain such that

$$D \subset \{(x, y) : x^2 + y^2 \leq 64 \text{ and } \frac{x^2}{36} + \frac{y^2}{49} \geq 1\}.$$

(This is just an inclusion and D does not have to be equal to the set on the right.)

Let $\mathcal{S}$ be the part of the sphere $x^2 + y^2 + z^2 = 64$ that lies above D . Prove that

$$(\text{area of } D) \leq \frac{3}{4}(\text{area of } \mathcal{S}).$$

8. Let $\mathcal{S}$ be the part of the sphere $x^2 + y^2 + z^2 = 1$ that lies above the cone $z = \sqrt{x^2 + y^2}$. Consider the integral

$$\iint_{\mathcal{S}} \sqrt{x^2 + y^2} dS.$$

A student has evaluated this integral as follows.

The unit normal to $\mathcal{S}$ is the vector $\mathbf{N} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$. The angle γ between $\mathbf{N}$ and $\mathbf{k}$ is $\gamma = \phi$ in spherical coordinates, so that the area element is $\frac{1}{\cos \gamma} = \frac{1}{\cos \phi}$. Also in spherical coordinates, $\sqrt{x^2 + y^2} = \sin \phi$ on the sphere. Therefore

$$\begin{aligned} \iint_{\mathcal{S}} \sqrt{x^2 + y^2} dS &= \int_0^{2\pi} \int_0^{\pi/4} \frac{\sqrt{x^2 + y^2}}{\cos \gamma} d\phi d\theta \\ &= \int_0^{2\pi} \int_0^{\pi/4} \frac{\sin \phi}{\cos \phi} d\phi d\theta \\ &= 2\pi \left[-\ln(\cos \phi) \right]_0^{\pi/4} \\ &= 2\pi \left(-\ln(\cos(\pi/4)) + \ln(\cos 0) \right) \\ &= 2\pi(-\ln(\sqrt{2}/2) - \ln(1)) = -2\pi \ln(\sqrt{2}/2). \end{aligned}$$

Is this solution correct? If not, explain where the student has made a mistake (if there is more than one error, indicate all of them), and provide the correct solution.

3.2 Flux Integrals

3.2.1 Oriented Surfaces

A smooth surface $\mathcal{S} \subset \mathbb{R}^3$ is **orientable** if we can define a *continuous* vector field $\mathbf{N} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ on $\mathcal{S}$ so that $|\mathbf{N}| = 1$ and $\mathbf{N}$ is normal to $\mathcal{S}$ at each point. If such a field exists, it is unique up to the choice of sign (if $\mathbf{N}$ has the above property, then so does $-\mathbf{N}$). The **orientation** of $\mathcal{S}$ is the choice of one of $\mathbf{N}$ and $-\mathbf{N}$. Assuming that we choose $\mathbf{N}$, the **positive side** of $\mathcal{S}$ is the side out of which $\mathbf{N}$ is pointing. An **oriented surface** is an orientable surface $\mathcal{S}$ together with a choice of $\mathbf{N}$.

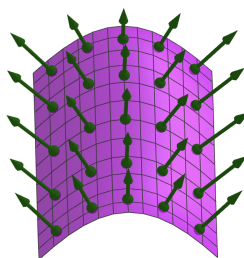


Figure 3.5: An example of an oriented surface. We have demarcated a uniformly-spaced collection of points $P \in \mathcal{S}$ as well as the associated normal vector $\mathbf{N}(P)$ at each point.

Some common types of orientable surfaces are:

- All *closed surfaces* are orientable. When $\mathcal{S}$ is closed, there is a natural notion of “inside” and “outside”. Unless stated otherwise, the default orientation will always be with $\mathbf{N}$ pointing outward (to the outside of $\mathcal{S}$).
- If $\mathcal{S} = \mathcal{S}_g$ is the *graph of a C^1 function* $g : \mathbb{R}^2 \rightarrow \mathbb{R}$, then $\mathcal{S}_g$ is also orientable. This follows because a formula for the (non-normalized!) normal vector is given by

$$\mathbf{n}(x, y) := -g_x(x, y)\mathbf{i} - g_y(x, y)\mathbf{j} + \mathbf{k},$$

which is continuous (since g is smooth) and nonzero because the last component is always 1. In this situation, the default orientation has $\mathbf{N}$ pointing in the same direction as the vector $\mathbf{n}(x, y)$ defined above, “upward” in the sense that the $\mathbf{k}$ -component of $\mathbf{N}$ is positive.

3.2.2 Orientation and Boundary Curves

If $\mathcal{S}$ is oriented *and also* has a boundary (i.e. a curve $\mathcal{C}$), then the orientation on $\mathcal{S}$ induces an orientation on $\mathcal{C}$. The induced orientation is as follows: if we imagine a

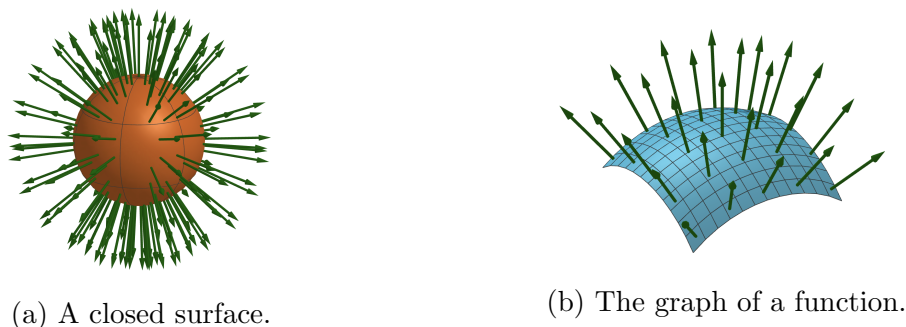


Figure 3.6: Two pictures of the examples above.

person standing on the curve $\mathcal{C}$ with their head pointing in the direction of the normal vector $\mathbf{N}$, looking forward in the direction indicated by the orientation of $\mathcal{C}$, then their right hand should point away from $\mathcal{S}$. For example, if $\mathcal{S}$ is the unit disc in the xy -plane with the orientation provided by $\mathbf{N} = \mathbf{k}$, then $\mathcal{C}$ is the unit circle oriented counterclockwise.

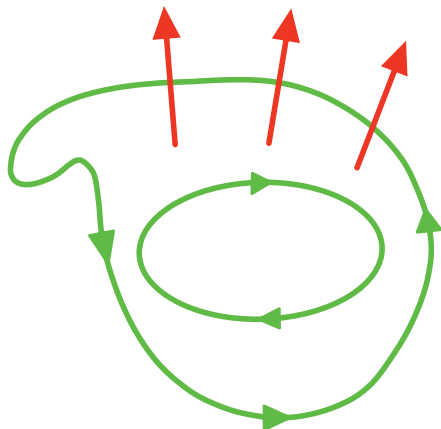


Figure 3.7: In the above picture, $\mathcal{S}$ is the part of the plane that lies between two curves $\mathcal{C}_1$ (“on the outside”) and $\mathcal{C}_2$ (“on the inside”). If $\mathcal{S}$ is oriented so that the normal vectors $\mathbf{N}$ (in red) point towards the viewer, then the matching boundary orientation is counterclockwise on $\mathcal{C}_1$ (the outside) and clockwise on $\mathcal{C}_2$ (the inside).

A **piecewise smooth** surface $\mathcal{S}$ can be broken up into finitely many smooth pieces $\mathcal{S}_1, \dots, \mathcal{S}_J$, joined along shared boundary curves. Assume that each piece is orientable, with the corresponding normal vectors $\mathbf{N}_1, \dots, \mathbf{N}_J$. If the choices of normal vectors on each of the smooth parts are compatible, in the sense that they point out of the same side of the surface, then they define a “global” orientation on $\mathcal{S}$. The compatibility condition can be checked using the induced orientation on joint boundary curves: the induced curve orientations should be *opposite* to each other.

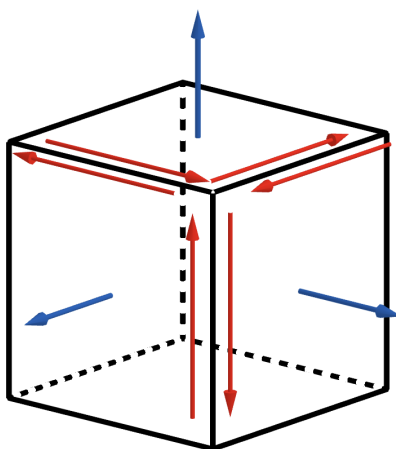
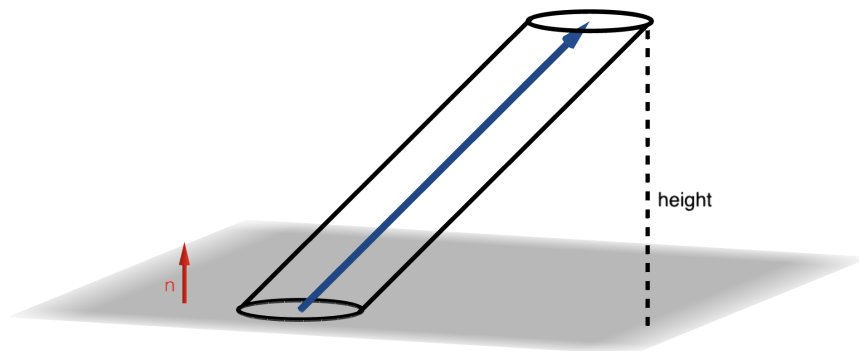


Figure 3.8: The cube is a piecewise smooth surface. If we choose the normal vectors $\mathbf{N}_1$, $\mathbf{N}_2$ and $\mathbf{N}_3$ (in blue) to point outward, the associated boundary curves have compatible (i.e. opposite) orientations!

3.2.3 Flux Integrals

Now suppose that $\mathbf{F}$ is the velocity of a gas or fluid flowing through a porous membrane $\mathcal{S}$. Our question is: how much gas or fluid passes from one side of $\mathcal{S}$ to the other side in a given unit of time? Notice that, in order for this question to make sense, we need to assume that $\mathcal{S}$ is *orientable*. This is so that the notion of “one side” and “the other side” have a well-defined meaning.

For a small patch of $\mathcal{S}$, the fluid flowing through that patch in unit time forms a cylinder (usually non-rectangular); the base area of this cylinder is dS and its height is given by $h = \mathbf{F} \cdot \mathbf{N}$, so that the volume is $\mathbf{F} \cdot \mathbf{N} dS$.



Summing up over all such cylinders and passing to the limit as $dS \rightarrow 0$, we obtain

the flux integral:

$$(\text{Flux Integral}) \quad \iint_{\mathcal{S}} \mathbf{F} \cdot \mathbf{N} dS = \iint_{\mathcal{S}} \mathbf{F} \cdot d\mathbf{S}, \quad (3.2.1)$$

where $d\mathbf{S} = \mathbf{N}dS$ is simply a notation we use to suppress the orientation vector $\mathbf{N}$. Often, if $\mathcal{S}$ is a closed surface with orientation $\mathbf{N}$ pointing outward, we may use the notation $\oiint_{\mathcal{S}}$ instead of $\iint_{\mathcal{S}}$.

3.2.4 Calculating Flux Integrals

We will now spend some time learning how to calculate flux integrals. This can be done using geometric considerations, a suitable parametrization of the surface, or some combination of both.

1. Using geometric considerations. This often works if $\mathbf{F} \cdot \mathbf{N}$ is particularly simple or easy to compute. Other times, it can be combined with our second approach below (using a parametrization) to simplify the calculations. Let's see an example.

Example 3.2.1. Evaluate $\oiint \mathbf{F} d\mathbf{S}$ if $\mathbf{F}(\mathbf{r}) = \frac{c\mathbf{r}}{|\mathbf{r}|^3}$ and $\mathcal{S}$ is the sphere $x^2 + y^2 + z^2 = a^2$, for $a > 0$.

The unit normal vector pointing outward is $\mathbf{N}(\mathbf{r}) = \frac{\mathbf{r}}{|\mathbf{r}|} = \frac{\mathbf{r}}{a}$. We also notice that $\mathbf{F}$ is radial, therefore parallel to $\mathbf{N}$. So,

$$\mathbf{F} \cdot \mathbf{N} = \frac{c\mathbf{r}}{a^3} \cdot \frac{\mathbf{r}}{a} = \frac{c|\mathbf{r}|^2}{a^4} = \frac{c}{a^2}.$$

At the end, we used once more that $|\mathbf{r}| = a$ on $\mathcal{S}$. In particular, we see that $\mathbf{F} \cdot \mathbf{N}$ is constant on the sphere $\mathcal{S}$. Hence,

$$\begin{aligned} \oiint_{\mathcal{S}} \mathbf{F} \cdot d\mathbf{S} &= \iint_{\mathcal{S}} \mathbf{F} \cdot \mathbf{N} dS = \iint_{\mathcal{S}} \frac{c}{a^2} dS \\ &= \frac{c}{a^2} (\text{Area of } \mathcal{S}) = 4\pi c. \end{aligned}$$

Note that the flux of $\mathbf{F}$ over $\mathcal{S}$ does not depend upon the radius a !

2. Using parametrization. If $\mathcal{S}$ is a parametric surface given by $\mathbf{r}(u, v)$ for $(u, v) \in D$, we know that a normal vector is given by $\mathbf{n} = \mathbf{r}_u \times \mathbf{r}_v$, so that $dS = |\mathbf{n}| du dv$. As long as $\mathbf{n} \neq \mathbf{0}$, the unit normal is given by $\mathbf{N} = \pm \frac{\mathbf{n}}{|\mathbf{n}|}$ (where the $\pm$ is chosen to match the orientation). So,

$$\begin{aligned} \iint_{\mathcal{S}} \mathbf{F} \cdot d\mathbf{S} &= \iint_{\mathcal{S}} \mathbf{F} \cdot \mathbf{N} dS = \iint_D \mathbf{F} \cdot \left(\pm \frac{\mathbf{n}}{|\mathbf{n}|} \right) |\mathbf{n}| du dv \\ &= \pm \iint_D \mathbf{F} \cdot \mathbf{n} du dv \end{aligned}$$

3. Special Case Parametrizations. Assume that $\mathcal{S}$ is the graph of a function $z = g(x, y)$, with $\mathbf{N}$ pointing up. Previously, we found the normal vector $\mathbf{n} = -g_x\mathbf{i} - g_y\mathbf{j} + \mathbf{k}$. Then $\mathbf{N} = \frac{\mathbf{n}}{|\mathbf{n}|}$ (why does this give the correct orientation?), so that

$$d\mathbf{S} = \mathbf{N} dS = \frac{\mathbf{n}}{|\mathbf{n}|} |\mathbf{n}| dx dy = (-g_x\mathbf{i} - g_y\mathbf{j} + \mathbf{k}) dx dy.$$

More generally, if $\mathcal{S}$ is given by the equation $G(x, y, z) = 0$ and $G_z = \frac{\partial G}{\partial z} \neq 0$ on $\mathcal{S}$, then we proved that

$$\mathbf{n} = \frac{\nabla G}{|G_z|}, \quad dS = \frac{|\nabla G|}{|G_z|} dx dy.$$

For the unit normal, we use $\mathbf{N} = \pm \frac{\nabla G}{|\nabla G|}$. This is just a unit vector in the direction of the vector ∇G perpendicular to the level surface of G , with the $\pm$ sign again depending on the choice of orientation. Hence

$$d\mathbf{S} = \mathbf{N} dS = \pm \frac{\nabla G}{|\nabla G|} \cdot \frac{|\nabla G|}{|G_z|} dx dy = \pm \frac{\nabla G}{|G_z|} dx dy,$$

Worked Problems

1. Evaluate $\iint_{\mathcal{S}} \mathbf{F} \cdot d\mathbf{S}$, if $\mathbf{F}(x, y) = x\mathbf{i} + z\mathbf{k}$ and $\mathcal{S}$ is parametrized by $\langle x, y, z \rangle = \langle u^2, uv, \frac{v^2}{2} \rangle$ and $0 \leq u, v \leq 1$, with $\mathbf{N}$ pointing up.

We have

$$\mathbf{r}_u = \langle 2u, v, 0 \rangle, \quad \mathbf{r}_v = \langle 0, u, v \rangle \Rightarrow \mathbf{n} = \mathbf{r}_u \times \mathbf{r}_v = \langle v^2, -2uv, 2u^2 \rangle.$$

Notice that, for all $0 < u < 1$, we have $2u^2 > 0$, so that $\mathbf{n}$ points up everywhere except for part of the boundary curve. Now, on $\mathcal{S}$, we know that $x = u^2$ and $z = \frac{v^2}{2}$, so that

$$\mathbf{F} \cdot \mathbf{n} = (u^2 v^2 + \frac{1}{2} v^2 \cdot 2u^2) = 2u^2 v^2.$$

Hence,

$$\begin{aligned} \iint_{\mathcal{S}} \mathbf{F} \cdot d\mathbf{S} &= \int_0^1 \int_0^1 \mathbf{F} \cdot \mathbf{n} du dv = \int_0^1 \int_0^1 (2u^2 v^2) du dv = 2 \frac{u^3}{3} \Big|_0^1 \cdot \frac{v^3}{3} \Big|_0^1 \\ &= \frac{2}{9}. \end{aligned}$$

2. Let $\mathbf{F} = x\mathbf{j} + 2y\mathbf{k}$, and let $\mathcal{S}$ be the part of the plane $x + y + z = 1$ that lies in the 1st octant ($x, y, z \geq 0$), with $\mathbf{N}$ pointing up. Evaluate $\iint_{\mathcal{S}} \mathbf{F} \cdot d\mathbf{S}$.

Our surface can be written as

$$z = g(x, y), \text{ where } g(x, y) = 1 - x - y.$$

This gives

$$d\mathbf{S} = (-g_x\mathbf{i} - g_y\mathbf{j} + \mathbf{k}) dx dy = (\mathbf{i} + \mathbf{j} + \mathbf{k}) dx dy,$$

so that

$$\mathbf{F} \cdot d\mathbf{S} = (0 \cdot 1 + x \cdot 1 + 2y \cdot 1) dx dy = (x + 2y) dx dy.$$

We also have $D = \{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq 1 - x\}$. The calculation then goes as follows:

$$\begin{aligned} \iint_{\mathcal{S}} \mathbf{F} \cdot d\mathbf{S} &= \int_0^1 \int_0^{1-x} (x + 2y) dy dx = \int_0^1 (xy + y^2) \Big|_{y=0}^{y=1-x} dx \\ &= \int_0^1 (1 - x) dx = \left[x - \frac{x^2}{2} \right]_{x=0}^1 \\ &= \frac{1}{2}. \end{aligned}$$

3. Find $\iint_{\mathcal{S}} \mathbf{F} \cdot d\mathbf{S}$, if $\mathbf{F} = y\mathbf{i} - x\mathbf{j}$ and $\mathcal{S}$ is the part of the graph of $z = x + y^2$ that lies above the rectangle $D := \{(x, y) : 0 \leq x \leq 3, 0 \leq y \leq 2\}$, oriented with $\mathbf{N}$ pointing up.

Our surface is given by the equation $z = g(x, y)$ with $g(x, y) = x + y^2$. Therefore

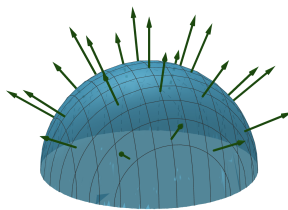
$$d\mathbf{S} = (-g_x\mathbf{i} - g_y\mathbf{j} + \mathbf{k}) dx dy = (-\mathbf{i} - 2y\mathbf{j} + \mathbf{k}) dx dy,$$

which gives

$$\mathbf{F} \cdot d\mathbf{S} = (y \cdot (-1) - x(-2y) + 0) dx dy = (2xy - y) dx dy.$$

The calculation is then

$$\begin{aligned} \iint_{\mathcal{S}} \mathbf{F} \cdot d\mathbf{S} &= \int_0^3 \int_0^2 (2xy - y) dy dx = \int_0^3 \left(xy^2 - \frac{y^2}{2} \right) \Big|_{y=0}^2 dx \\ &= \int_0^3 (4x - 2) dx = (2x^2 - 2x) \Big|_0^3 \\ &= 12. \end{aligned}$$



4. Evaluate the integral $\iint_{\mathcal{S}} \mathbf{F} \cdot d\mathbf{S}$, where $\mathbf{F} = x\mathbf{i} + y\mathbf{j}$ and $\mathcal{S}$ is the half-sphere $x^2 + y^2 + z^2 = 1$, $z \geq 0$, with the normal vector pointing up.

The unit normal vector to the sphere is $\mathbf{N} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, so that $\mathbf{F} \cdot \mathbf{N} = x^2 + y^2$. In spherical coordinates (θ, ϕ) ,

$$\mathbf{F} \cdot \mathbf{N} = (\cos^2 \theta + \sin^2 \theta) \sin^2 \phi = \sin^2 \phi,$$

so that

$$\begin{aligned} \int_{\mathcal{S}} \mathbf{F} \cdot d\mathbf{S} &= \int_0^{2\pi} \int_0^{\pi/2} \sin^2 \phi \sin \phi \, d\phi d\theta \\ &= \int_0^{2\pi} \int_0^{\pi/2} (1 - \cos^2 \phi) \sin \phi \, d\phi d\theta \\ &= 2\pi \left[-\cos \phi + \frac{\cos^3 \phi}{3} \right]_0^{\pi/2} \\ &= 2\pi \left(1 - \frac{1}{3} \right) = \frac{4\pi}{3}. \end{aligned}$$

Practice Problems

- Let $\mathcal{S}$ be the surface $\mathbf{r}(u, v) = (u \cos v, u \sin v, 2v)$ with $0 \leq u \leq 2$, $0 \leq v \leq 2\pi$, oriented with $\mathbf{N}$ pointing up. Evaluate the integral $\iint_{\mathcal{S}} \mathbf{F} \cdot d\mathbf{S}$, where $\mathbf{F} = -y\mathbf{i} + (x + y^2)\mathbf{j} + z\mathbf{k}$.
- Evaluate the integral $\iint_{\mathcal{S}} \mathbf{F} \cdot d\mathbf{S}$, where $\mathbf{F} = z\mathbf{i} + y^2\mathbf{j}$ and $\mathcal{S}$ is the part of the cylinder $x^2 + y^2 = 4$ where $x \geq 0$, $y \geq 0$, $0 \leq z \leq 1$, with the normal pointing away from the z -axis.
- Evaluate the surface integral $\iint_{\mathcal{S}} \mathbf{F} \cdot d\mathbf{S}$, if $\mathbf{F} = z\mathbf{k}$ and $\mathcal{S}$ is the rectangle with vertices $(0, 2, 0)$, $(0, 0, 4)$, $(5, 2, 0)$, $(5, 0, 4)$, oriented so that the normal vector points upward.

4. Evaluate the surface integral $\iint_{\mathcal{S}} \mathbf{F} \cdot d\mathbf{S}$, where $\mathbf{F} = x^2\mathbf{i}$ and $\mathcal{S}$ is the part of the plane $x + \frac{y}{2} + \frac{z}{3} = 1$ that lies in the first octant ($x, y, z \geq 0$), with the normal vector pointing up.
5. Evaluate the integral $\iint_{\mathcal{S}} \mathbf{F} \cdot d\mathbf{S}$, where $\mathbf{F} = \mathbf{i} + 2\mathbf{j} + 2\mathbf{k}$ and $\mathcal{S}$ is the surface parametrized by $\mathbf{r}(u, v) = \langle uv, u^2 - v^2, u^2 + v^2 \rangle$ with $0 \leq u \leq 1$, $0 \leq v \leq u$, with the normal vector pointing up.
6. Evaluate the integral $\iint_{\mathcal{S}} \mathbf{F} \cdot d\mathbf{S}$, where $\mathbf{F} = x^2\mathbf{i} + y^4\mathbf{j}$ and $\mathcal{S}$ is the part of the surface $xy^2z = 3$ that lies above the rectangle $1 \leq x \leq 2$, $1 \leq y \leq 3$, with the normal vector pointing up.

Chapter 4

Main Theorems of Vector Calculus

4.1 Divergence and Curl

Consider a vector field $\mathbf{F} = F_1\mathbf{i} + F_2\mathbf{j} + F_3\mathbf{k}$, defined on a domain $D \subset \mathbb{R}^3$, and assume that all component functions are C^1 on D . We will sometimes refer to C^1 vector fields as **smooth**. We focus on the following two functions related to our vector field.

Definition 4.1.1 (Divergence and Curl). *The **divergence** of $\mathbf{F}$ is*

$$\operatorname{div} \mathbf{F} := \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z}, \quad (4.1.1)$$

*and the **curl** of $\mathbf{F}$ is*

$$\operatorname{curl} \mathbf{F} := \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z} \right) \mathbf{i} - \left(\frac{\partial F_3}{\partial x} - \frac{\partial F_1}{\partial z} \right) \mathbf{j} + \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) \mathbf{k}. \quad (4.1.2)$$

The formulas (4.1.1) and (4.1.2) are easier to remember in vector notation. If we define the “vector”

$$\nabla := \mathbf{i} \frac{\partial}{\partial x} + \mathbf{j} \frac{\partial}{\partial y} + \mathbf{k} \frac{\partial}{\partial z},$$

then the divergence and curl of $\mathbf{F}$ are, formally,

$$\boxed{\operatorname{div} \mathbf{F} = \nabla \cdot \mathbf{F}, \quad \operatorname{curl} \mathbf{F} = \nabla \times \mathbf{F}.}$$

The calculation for curl goes as follows:

$$\nabla \times \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_1 & F_2 & F_3 \end{vmatrix} = \mathbf{i} \begin{vmatrix} \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_2 & F_3 \end{vmatrix} - \mathbf{j} \begin{vmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial z} \\ F_1 & F_3 \end{vmatrix} + \mathbf{k} \begin{vmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \\ F_1 & F_2 \end{vmatrix},$$

matching (4.1.2) if the 2×2 determinants are “evaluated” by taking the indicated derivatives. We recommend memorizing the cross product formula (or equivalently, the determinant expression) for curl, as (4.1.2) can be difficult to remember any other way. We will spend most of this section determining what information about $\mathbf{F}$ the divergence and curl encode.

4.1.1 Some Examples of Divergence and Curl

Let us calculate the divergence and curl of some vector fields. We first remark that, whenever $\mathbf{F} = \mathbf{a}$ for some constant vector $\mathbf{a}$, then

$$\operatorname{div} \mathbf{F} = \nabla \cdot \mathbf{F} = 0, \quad \operatorname{curl} \mathbf{F} = \nabla \times \mathbf{F} = \mathbf{0}.$$

So, for constant vector fields, we expect no divergence or curl (a helpful example to keep in mind when we start discussing the geometric interpretation of these functions). Let's see some more interesting examples.

Example 4.1.2. Suppose that $\mathbf{F}$ is a planar vector field, so that $\mathbf{F} = F_1(x, y)\mathbf{i} + F_2(x, y)\mathbf{j}$. Then

$$\operatorname{curl} \mathbf{F} = \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) \mathbf{k}$$

since $F_3 = 0$ and all derivatives in z are 0. Whenever we talk about the curl of a 2-dimensional vector field in the xy -plane, we understand it to be the above expression, with the third direction added.

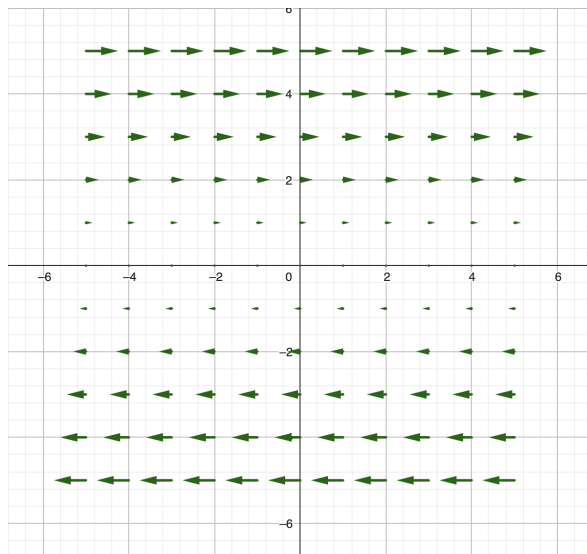


Figure 4.1: In Example 4.1.3, the vector field “circulates” around the vector $-\mathbf{k}$. Imagine that the vector field represents the flow of water on a surface, and that you put a floating stick parallel to the y -axis on top of it. What do you think will happen to the stick?

Example 4.1.3. Let $\mathbf{F} = y\mathbf{i}$. Then, $\nabla \cdot \mathbf{F} = 0$ and

$$\operatorname{curl} \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ y & 0 & 0 \end{vmatrix} = -\mathbf{k}.$$

We interpret the curl of $\mathbf{F}$ as the twisting motion of our vector field (see Figure 4.1).

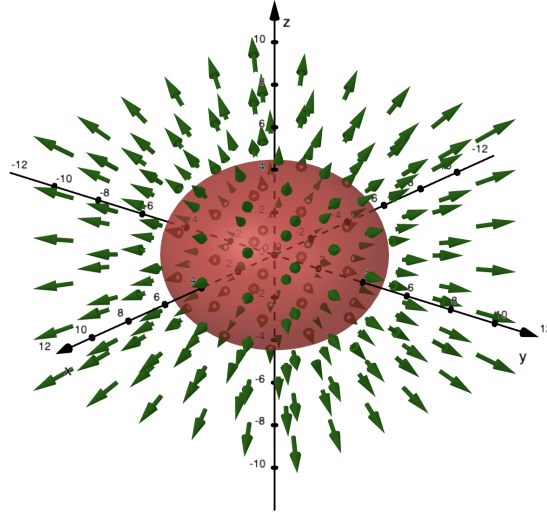


Figure 4.2: The field $\mathbf{F}$ from Example 4.1.4 with $c > 0$. The sphere is included to show you the symmetry of the vector field about the origin.

Example 4.1.4. Let $\mathbf{F} = c\mathbf{r} = c(x\mathbf{i} + y\mathbf{j} + z\mathbf{k})$ for some constant $c \neq 0$. Then

$$\nabla \cdot \mathbf{F} = c(1 + 1 + 1) = 3c,$$

and

$$\nabla \times \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ x & y & z \end{vmatrix} = \mathbf{0}.$$

Divergence is associated with expansion (if $c > 0$) or contraction (if $c < 0$).

The next example is quite general; so, we list it as a result (Proposition 4.1.5 below) for future reference.

Proposition 4.1.5. Suppose that $\mathbf{F} : D \rightarrow \mathbb{R}^3$ is C^1 and conservative, with $\nabla\phi = \mathbf{F}$. Then $\nabla \times \mathbf{F} = \mathbf{0}$ on D .

Proof. We know that $\mathbf{F} := \phi_x\mathbf{i} + \phi_y\mathbf{j} + \phi_z\mathbf{k}$. Calculating the curl then gives:

$$\nabla \times \mathbf{F} = (\phi_{zy} - \phi_{yz})\mathbf{i} - (\phi_{zx} - \phi_{xz})\mathbf{j} + (\phi_{yx} - \phi_{xy})\mathbf{k}.$$

If $\mathbf{F}$ is C^1 on its domain, then ϕ is C^2 (why?). Hence, equality of mixed partials holds (e.g. $\phi_{zy} = \phi_{yz}$), and the above calculation shows that the curl of $\mathbf{F}$ is zero on the domain of $\mathbf{F}$. $\square$

The special properties of having zero divergence and zero curl have a name.

Definition 4.1.6. A vector field $\mathbf{F} : D \rightarrow \mathbb{R}^3$ is said to be:

- *incompressible* if $\operatorname{div} \mathbf{F} \equiv 0$;
- *irrotational* if $\operatorname{curl} \mathbf{F} \equiv \mathbf{0}$.

An important example is provided by the *gravitational field* $\mathbf{G} = c\mathbf{r}/|\mathbf{r}|^3$. We have already seen that $\mathbf{G}$ is conservative; by Proposition 4.1.5, $\mathbf{G}$ is irrotational. Moreover, $\mathbf{G}$ is also incompressible, so that $\nabla \cdot \mathbf{G} = 0$. We prove this in our Worked Problems Section.

Proposition 4.1.7. *If $\mathbf{F}$ is C^2 , then $\nabla \times \mathbf{F}$ is incompressible. That is: $\nabla \cdot (\nabla \times \mathbf{F}) = 0$.*

Proof. This is a calculation similar to that in Proposition 4.1.5 (and also based on the equality of mixed partials), but longer. The details are left as an exercise. $\square$

Example 4.1.8. *We return to another important example from Chapter 2. Let*

$$\mathbf{F}(x, y) := \underbrace{-\frac{y}{x^2 + y^2}}_{F_1(x, y)} \mathbf{i} + \underbrace{\frac{x}{x^2 + y^2}}_{F_2(x, y)} \mathbf{j}, \text{ for } (x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}.$$

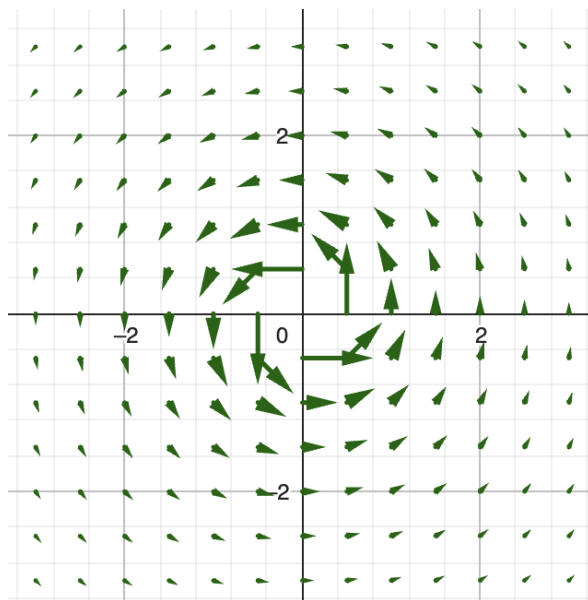


Figure 4.3: The vector field $\mathbf{F}$ in Example 4.1.8 is irrotational, but not conservative. Imagine that the vector field represents the flow of water on a surface, and that you put a floating stick on top of it, away from the origin. Then the stick will travel in circles around the origin, but it will not rotate!

Since $\mathbf{F}$ is a planar field, we know that

$$\operatorname{curl} \mathbf{F} = \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) \mathbf{k}.$$

Direct calculation gives

$$\frac{\partial F_2}{\partial x} = \frac{x^2 - y^2}{y^2 + x^2} = \frac{\partial F_1}{\partial y},$$

so that $\mathbf{F}$ is irrotational. Nevertheless, we have already seen (with several different proofs) that $\mathbf{F}$ fails to be conservative.

4.1.2 Flux and Circulation Density

We now come to some deeper results regarding the divergence and curl of $\mathbf{F}$. They relate these functions to certain integrals of $\mathbf{F}$ on curves and surfaces, providing several rigorous versions of the general statement that divergence is associated with expansion/contraction, and curl with twisting motion. We start with a local result.

Theorem 4.1.9 (Divergence as Flux Density). *Let $\mathcal{S}_\epsilon \subset \mathbb{R}^3$ be a sphere of radius $\epsilon > 0$ centered at $P \in \mathbb{R}^3$, with $\mathbf{N}$ pointing outward. Assume that $\mathbf{F}$ is C^2 on some open set containing B_ϵ , the solid ball of radius ϵ centered at P . Then the divergence of $\mathbf{F}$ at P satisfies*

$$\operatorname{div} \mathbf{F}(P) = \lim_{\epsilon \rightarrow 0} \frac{1}{\operatorname{Vol}(\mathcal{S}_\epsilon)} \iint_{B_\epsilon} \mathbf{F} \cdot \mathbf{N} dS = \lim_{\epsilon \rightarrow 0} \frac{1}{\frac{4}{3}\pi\epsilon^3} \iint_{\mathcal{S}_\epsilon} \mathbf{F} \cdot \mathbf{N} dS. \quad (4.1.3)$$

In words, the divergence can be understood (in the limit) as the flux density of $\mathbf{F}$ over $\mathcal{S}_\epsilon$. Moreover, one has the approximation

$$\iint_{\mathcal{S}_\epsilon} \mathbf{F} \cdot \mathbf{N} dS = \frac{4}{3}\pi\epsilon^3 \cdot \operatorname{div} \mathbf{F}(P) + O(\epsilon^4) \quad (4.1.4)$$

as $\epsilon \rightarrow 0$.

Here, $O(\epsilon^4)$ is the “Big- O Notation”, meaning that our error has magnitude at most $K\epsilon^4$ for some constant $K > 0$.

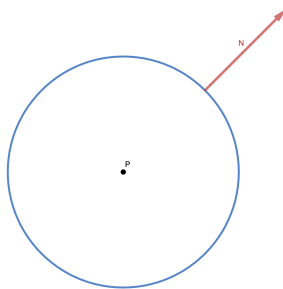


Figure 4.4: The sphere $\mathcal{S}_\epsilon$ about P with associated normal vector $\mathbf{N}$.

Proof. Assume that $P = (0, 0, 0)$ and write $\mathbf{F} = F_1\mathbf{i} + F_2\mathbf{j} + F_3\mathbf{k}$. Our goal is to show (4.1.4). To do that, we will start with the flux integral, apply linear approximation to $\mathbf{F}$, evaluate the leading term, and keep track of the error estimates.

We proceed as follows. On the sphere $\mathcal{S}_\epsilon = \{\mathbf{r} : |\mathbf{r}| = \epsilon\}$, our normal vector $\mathbf{N}$ takes the form

$$\mathbf{N} = \frac{1}{\epsilon}(x\mathbf{i} + y\mathbf{j} + z\mathbf{k}).$$

To approximate the behaviour of $\mathbf{F}$ near the origin, we expand each F_j (for $j = 1, 2, 3$) in a Taylor series near $\mathbf{0}$.¹ We then evaluate the contribution of each component function F_j to the expression (4.1.4). Let us do the calculation for F_1 in full detail.

The Taylor approximation about $\mathbf{0}$ gives, for $(x, y, z) \in \mathcal{S}_\epsilon$,

$$F_1(\mathbf{r}) = \underbrace{F_1(\mathbf{0}) + F_{1,x}(\mathbf{0})x + F_{1,y}(\mathbf{0})y + F_{1,z}(\mathbf{0})z}_{\text{linear approximation}} + \underbrace{O(\epsilon^2)}_{\text{error from higher order terms}}.$$

Now, using our formulas for $\mathbf{N}$ and F_1 , we have

$$F_1\mathbf{i} \cdot \mathbf{N} = F_1 \frac{x}{\epsilon} = \frac{1}{\epsilon} [F_1(\mathbf{0})x + F_{1,x}(\mathbf{0})x^2 + F_{1,y}(\mathbf{0})xy + F_{1,z}(\mathbf{0})xz] + \underbrace{\frac{x}{\epsilon}O(\epsilon^2)}_{(*)}.$$

The new error $(*)$ still has order $O(\epsilon^2)$ because $|x| \leq \epsilon$ (so that $|\frac{x}{\epsilon}| \leq 1$) on $\mathcal{S}_\epsilon$.

We now evaluate the contribution of $F_1\mathbf{i} \cdot \mathbf{N}$ to the integral (4.1.4). By symmetry, we have

$$\iint_{\mathcal{S}_\epsilon} x dS = \iint_{\mathcal{S}_\epsilon} xy dS = \iint_{\mathcal{S}_\epsilon} xz dS = 0.$$

because for each of the functions x , xy , and xz , the integrals of their positive and negative parts are equal in size and have opposite signs. Also by symmetry, this time with respect to interchanging the x, y, z variables,

$$\begin{aligned} \iint_{\mathcal{S}_\epsilon} x^2 dS &= \frac{1}{3} \iint_{\mathcal{S}_\epsilon} (x^2 + y^2 + z^2) dS = \frac{1}{2} \iint_{\mathcal{S}_\epsilon} \epsilon^2 dS = \frac{\epsilon^2}{3} (\text{Area of } \mathcal{S}_\epsilon) \\ &= \frac{4\pi}{3} \epsilon^4. \end{aligned}$$

Putting these calculations together, we see that:

$$\begin{aligned} \iint_{\mathcal{S}_\epsilon} (F_1\mathbf{i}) \cdot \mathbf{N} dS &= \left(\frac{1}{\epsilon} F_{1,x}(\mathbf{0}) \right) \cdot \left(\frac{4\pi\epsilon^4}{3} \right) + \underbrace{\iint_{\mathcal{S}_\epsilon} O(\epsilon^2) dS}_{\text{Area} = O(\epsilon^2)} \\ &= \frac{4\pi\epsilon^3}{3} F_{1,x}(\mathbf{0}) + O(\epsilon^4). \end{aligned}$$

Given that similar calculations hold for $F_2\mathbf{j}$ and $F_3\mathbf{k}$, we obtain (4.1.4). $\square$

¹The error term estimate is the part that requires the C^2 assumption. See the chapter on Taylor approximation in Math 226.

We also have a similar formula for the curl of $\mathbf{F}$.

Theorem 4.1.10 (Curl as Circulation Density). *If $\mathbf{F}$ is C^2 on a neighbourhood of P , and C_ϵ is a circle of radius ϵ centred at P in the plane normal to $\mathbf{N}$, and oriented counterclockwise when viewed from the direction of $\mathbf{N}$, then*

$$\lim_{\epsilon \rightarrow 0} \frac{1}{\pi \epsilon^2} \oint_{C_\epsilon} \mathbf{F} \cdot d\mathbf{r} = \mathbf{N} \cdot \text{curl } \mathbf{F}(P). \quad (4.1.5)$$

The proof uses similar ideas as the theorem on divergence and flux density. The starting point is, again, using linear approximation with error estimates for each component of $\mathbf{F}$. We use this approximation to evaluate the line integrals, then use symmetry and cancellation to simplify the integrals. The details are left for the interested student to work out.

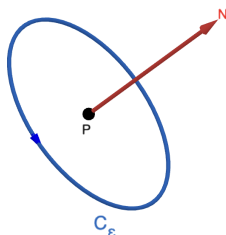


Figure 4.5: Curl as circulation density: the situation described in the theorem.

Worked Problems

1. Let $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ and let $\mathbf{G}(\mathbf{r}) = c \frac{\mathbf{r}}{|\mathbf{r}|^3}$ (for a gravitational field, choose $c < 0$). Show, via direct calculation, that $\nabla \cdot \mathbf{G} = 0$.

In terms of components, we have

$$\mathbf{G}(x, y, z) = c \left(\frac{x}{(x^2 + y^2 + z^2)^{3/2}} \mathbf{i} + \frac{y}{(x^2 + y^2 + z^2)^{3/2}} \mathbf{j} + \frac{z}{(x^2 + y^2 + z^2)^{3/2}} \mathbf{k} \right).$$

By the quotient rule,

$$\frac{\partial G_1}{\partial x} = c \left[\frac{x^2 + y^2 + z^2}{(x^2 + y^2 + z^2)^{5/2}} - \frac{3x^2}{(x^2 + y^2 + z^2)^{5/2}} \right]$$

and similarly for G_2 and G_3 . Hence, we have

$$\nabla \cdot \mathbf{G} = c \left[\frac{3(x^2 + y^2 + z^2)}{(x^2 + y^2 + z^2)^{5/2}} - \frac{3x^2 + 3y^2 + 3z^2}{(x^2 + y^2 + z^2)^{5/2}} \right] = 0.$$

Practice Problems

1. Let $\mathbf{F} = (y + z)^2\mathbf{i} + y\mathbf{j} + z\mathbf{k}$. Find the divergence and curl of $\mathbf{F}$.
2. (a) Let $\mathbf{F} = x \cos y\mathbf{i} + e^z\mathbf{j} + \cos y\mathbf{k}$. Find the curl and divergence of $\mathbf{F}$.
 (b) Is it true that $\nabla \times \mathbf{F}$ is always perpendicular to $\mathbf{F}$ for all C^1 vector fields $\mathbf{F}$ in $\mathbb{R}^3$? Explain why or why not.
3. Prove that if f and g are two C^2 functions on an open set $X \subset \mathbb{R}^3$, then $\nabla \cdot (\nabla f \times \nabla g) = 0$.
4. Let $\mathbf{F}$ be a 2-dimensional C^2 vector field on an open set U in the xy -plane. Let $P \in U$, and let $\mathbf{r}(t)$ be the field line of $\mathbf{F}$ that passes through P . Assume that:
 - $\mathbf{F}$ has constant magnitude 1 on a neighbourhood of P ,
 - the field line $\mathbf{r}(t)$ has curvature a at P .

Find the curl of $\mathbf{F}$ at P . Prove your answer.

4.2 Green's Theorem in Two Dimensions

In this section, we focus on **Green's Theorem**, which relates circulation integrals on a closed curve to curl integrals on the region bounded by said curve.

Theorem 4.2.1. *Let $\mathcal{C}$ be a simple, closed, piecewise smooth curve in $\mathbb{R}^2$, oriented counterclockwise, and let R denote the region bounded by $\mathcal{C}$. If $\mathbf{F} = F_1\mathbf{i} + F_2\mathbf{j}$ is C^1 on some open set containing R , then*

$$\oint_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{r} = \iint_R \text{curl } \mathbf{F} \cdot \mathbf{k} \, dA. \quad (4.2.1)$$

In terms of the component functions F_1 and F_2 , this means

$$\oint_{\mathcal{C}} F_1 dx + F_2 dy = \iint_R \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) dx dy. \quad (4.2.2)$$

We postpone the proof for now, focusing first on the geometric interpretation and applications of the formula above.

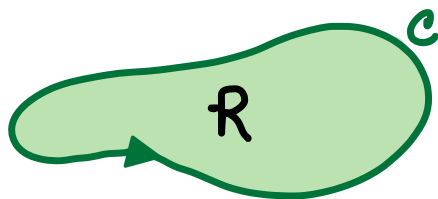


Figure 4.6: An example of a curve $\mathcal{C}$ and region R satisfying the first part of the hypothesis of Green's Theorem.

4.2.1 Green's Theorem: Some Special Cases

Let us take some time to consider the following situations where Green's Theorem may be applied.

- Let's first show that Green's Theorem is consistent with the "local version" from the last section. Suppose that $\mathcal{C} = \mathcal{C}_\epsilon$ is a circle of radius $\epsilon > 0$ centred at some point $P \in \mathbb{R}^2$, and let R_ϵ denote its interior. Applying formula (4.2.1) to a vector field $\mathbf{F}$ on R_ϵ , we have

$$\oint_{\mathcal{C}_\epsilon} \mathbf{F} \cdot d\mathbf{r} = \iint_{R_\epsilon} \text{curl } \mathbf{F} \cdot \mathbf{k} \, dA \approx (\text{Area of } R_\epsilon) \cdot (\text{curl } \mathbf{F}(P) \cdot \mathbf{k}),$$

and the approximation should work well so long as $\epsilon > 0$ is assumed to be small enough (so that $\text{curl } \mathbf{F}$ is close to constant on R_ϵ).

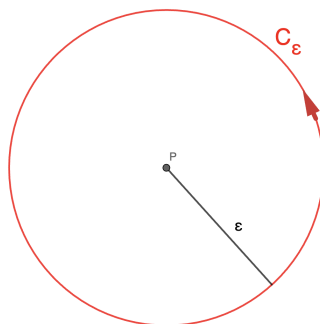
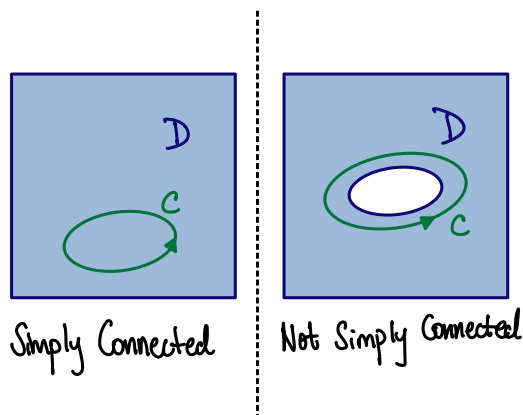


Figure 4.7: The circle $\mathcal{C}_\epsilon$ about P .

- If $\mathbf{F}$ is *irrotational* on the region R bounded by $\mathcal{C}$ (so that $\nabla \times \mathbf{F} = \mathbf{0}$ at every point in R), then formula (4.2.1) shows that $\oint_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{r} = 0$. In particular, if $\mathbf{F}$ is irrotational *everywhere* in $\mathbb{R}^2$, then $\oint_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{r} = 0$ for all $\mathcal{C} \subset \mathbb{R}^2$ which satisfy the assumption of Green's Theorem. (This is true, for example, whenever $\mathbf{F}$ is conservative!)

- A domain $D \subset \mathbb{R}^2$ is **simply connected** if every simple, closed curve in D encloses a region *contained in* D .



If $\nabla \times \mathbf{F} = \mathbf{0}$ everywhere on a *simply connected* domain D , then $\oint_C \mathbf{F} \cdot d\mathbf{r} = 0$ for every simple, closed, piecewise smooth curve C contained in D . With an additional argument, we can also drop the assumption that C is simple. (How would you go about proving this?) Hence, we have the following:

$$\nabla \times \mathbf{F} = \mathbf{0} \text{ on } D \text{ and } D \text{ is simply connected} \Rightarrow \mathbf{F} \text{ is conservative on } D.$$

The last implication can fail if D is not simply connected. The following example illustrates this failure for a specific function $\mathbf{F}$ and region D .

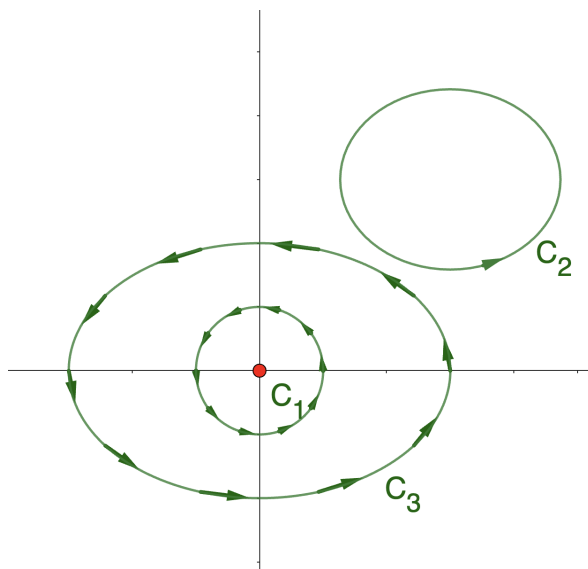
Example 4.2.2. Let $\mathbf{F}(x, y) = -\frac{y}{x^2+y^2}\mathbf{i} + \frac{x}{x^2+y^2}\mathbf{j}$ on $D = \mathbb{R}^2 \setminus \{\mathbf{0}\}$. From Example 4.1.8, we know that $\nabla \times \mathbf{F} = \mathbf{0}$ everywhere on its domain D . And yet, if C_1 is a circle around the origin, then

$$\oint_{C_1} \mathbf{F} \cdot d\mathbf{r} = 2\pi \neq 0!$$

(Exercise: parametrize the circle by $\mathbf{r}(t) = a \cos(t)\mathbf{i} + a \sin(t)\mathbf{j}$ with $0 \leq t \leq 2\pi$, and complete the calculation.) This can only happen because the domain $D = \mathbb{R}^2 \setminus \{\mathbf{0}\}$ is not simply connected.

In the last example, if we had instead taken a curve $C_2 \subset \mathbb{R}^2 \setminus \{\mathbf{0}\}$ satisfying the hypothesis of Green's Theorem and bounding a region that does not contain the origin, then we would indeed have that $\oint_{C_2} \mathbf{F} \cdot d\mathbf{r} = 0$. This is because $\mathbf{F}$ is *well defined* and *irrotational* for *every point* of the region bounded by C_2 —a key difference from C_1 .

On the other hand, let C_3 be **any** simple, closed, piecewise smooth curve that circles the origin once counterclockwise. We can show, with an extension of Green's Theorem

Figure 4.8: The three curves $\mathcal{C}_1$, $\mathcal{C}_2$ and $\mathcal{C}_3$ discussed in Example 4.2.2

to be discussed shortly, that

$$\oint_{\mathcal{C}_3} \mathbf{F} \cdot d\mathbf{r} = \oint_{\mathcal{C}_1} \mathbf{F} \cdot d\mathbf{r} = 2\pi.$$

A useful way to write this is as follows. Define the “delta function” $\delta_{\mathbf{0}}$ in the plane by saying that

$$\iint_R \delta_{\mathbf{0}} dA = \begin{cases} 1, & \text{if } \mathbf{0} \in R, \\ 0, & \text{if } \mathbf{0} \notin R. \end{cases}$$

(The “delta function” is not actually a function, and has to be defined rigorously as either a distribution or a measure. We will nonetheless use this terminology, and defer the rigorous definition to more advanced courses such as Math 420.) With that notation, the vector field $\mathbf{F}$ defined above satisfies

$$\text{curl } \mathbf{F} \cdot \mathbf{k} = 2\pi\delta_{\mathbf{0}},$$

since

$$\iint_R \text{curl } \mathbf{F} \cdot \mathbf{k} dA = \oint_{\partial R} \mathbf{F} \cdot d\mathbf{r} = \begin{cases} 2\pi, & \text{if } \mathbf{0} \in R, \\ 0, & \text{if } \mathbf{0} \notin R \end{cases}$$

and $\partial R = \mathcal{C}$ denotes the boundary curve of R .

4.2.2 Using Green's Theorem to simplify integrals

In some cases, Green's Theorem allows us to convert a complicated line integral to an area integral that is much easier to evaluate. For instance, let us evaluate the integral

$$\int_{\mathcal{C}} \left(\sqrt{x^2 - x + 16} + xy \right) dx + (2x + x^2 + \sin(y^3)) dy,$$

where $\mathcal{C}$ is the circle $x^2 + (y - 2)^2 = 4$, oriented counterclockwise. This would be unpleasant to integrate directly, but Green's Theorem can help. Let D be the disc $x^2 + (y - 2)^2 \leq 4$. We write

$$\begin{aligned} & \int_{\mathcal{C}} \left(\sqrt{x^2 - x + 16} + xy \right) dx + (2x + x^2 + \sin(y^3)) dy \\ &= \iint_D \left(\frac{\partial}{\partial x} (2x + x^2 + \sin(y^3)) - \frac{\partial}{\partial y} \left(\sqrt{x^2 - x + 16} + xy \right) \right) dA \\ &= \iint_D (2 + 2x - x) dA \\ &= 2(\text{area of } D) = 2 \cdot 4\pi = 8\pi, \end{aligned}$$

where we used that $\iint_D x dA = 0$ by symmetry.

4.2.3 Using Green's Theorem to compute areas

Simplifications in the other direction (where the line integral is simpler than the area integral) are less common, but here is one important case. Suppose that $\mathbf{F}$ is some vector field satisfying $\nabla \times \mathbf{F} = \mathbf{k}$ (so that $\text{curl } \mathbf{F} \cdot \mathbf{k} = \mathbf{k} \cdot \mathbf{k} = 1$). If you prefer a concrete example, you can use $\mathbf{F}(x, y) = x\mathbf{j}$ or $\mathbf{F}(x, y) = -y\mathbf{i}$. Then, with R and $\mathcal{C}$ as in Green's Theorem,

$$(\text{Area of } R) = \iint_R 1 dA = \oint_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{r} = \oint_{\mathcal{C}} x dy = - \oint_{\mathcal{C}} y dx. \quad (4.2.3)$$

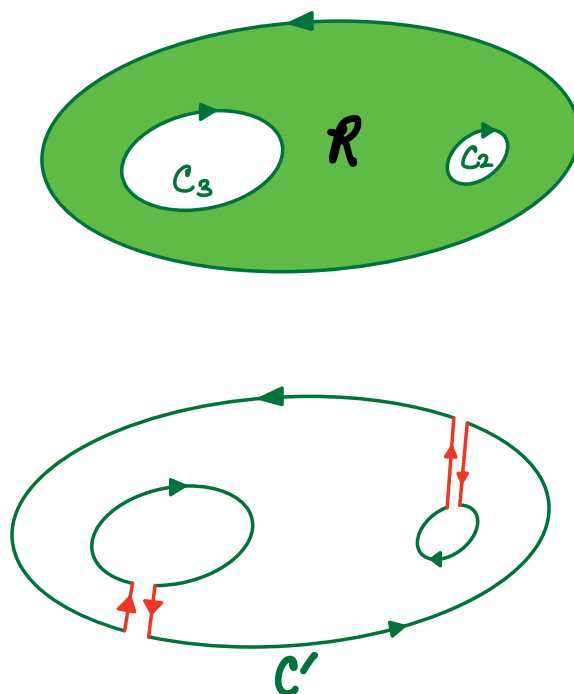
This means that we can use the above circulation integrals to compute the area of R ! We will show some specific examples of this idea in the Worked Problems below.

4.2.4 An Extension of Green's Theorem

There is a version of Green's Theorem for more complicated regions bounded by several curves, as follows.

Let $R \subset \mathbb{R}^2$ be bounded by *several* simple, closed, piecewise smooth curves $\mathcal{C}_1, \dots, \mathcal{C}_k$, *oriented positively with respect to* R . If $\mathcal{C}$ is the *total boundary* of R (that is, all of the curves $\mathcal{C}_i$ considered together) then we still have

$$\oint_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{r} = \iint_R \text{curl } \mathbf{F} \cdot \mathbf{k} dA.$$



A Sketch of the Proof of the Extension. Essentially, we want to connect each of the component pieces $\mathcal{C}_i$ of the boundary $\mathcal{C}$ of R by adding curves (such as the line segments in the figure above) to “convert” $\mathcal{C}$ into a single curve $\mathcal{C}'$, before applying our standard Green’s Theorem to $\mathcal{C}'$. One can check that each of these added-in line segments are traversed twice, once in each direction. Hence, the total contribution to the line integral over the added segments is zero! Also, if there are any curves “cutting out holes” in D , then they are counted in $\mathcal{C}'$ with the same orientation as in $\mathcal{C}$. This proves the extension. $\square$

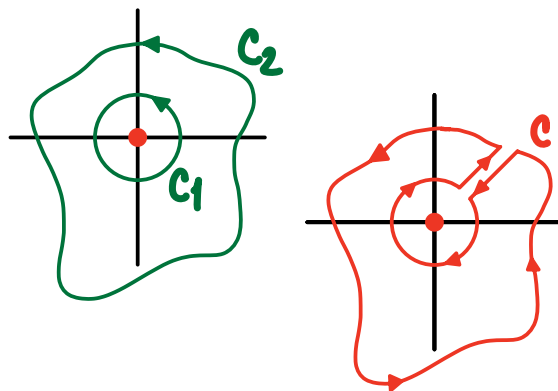
As a special case, let us return to the discussion after Example 4.2.2.

Proposition 4.2.3. *Suppose that $\mathbf{F}$ is smooth and has curl equal to $\mathbf{0}$ on $D = \mathbb{R}^2 \setminus \{\mathbf{0}\}$. Then, for any two curves $\mathcal{C}_1$ and $\mathcal{C}_2$ encircling the origin which satisfy the hypothesis of Green’s Theorem, one has*

$$\oint_{\mathcal{C}_1} \mathbf{F} \cdot d\mathbf{r} = \oint_{\mathcal{C}_2} \mathbf{F} \cdot d\mathbf{r}.$$

Proof. This follows from a similar idea as the proof of the extension of Green’s Theorem. It suffices to prove this when $\mathcal{C}_1$ is a small circle contained entirely in the region bounded by $\mathcal{C}_2$ (why?). We join the curves $\mathcal{C}_1$ and $\mathcal{C}_2$ together as in the diagram below.

We then see that $\mathcal{C}_2$ is traversed in the positive direction along $\mathcal{C}$, whereas $\mathcal{C}_1$ is traversed in the negative (i.e. opposite) direction as in the extended version of Green’s



Theorem. Hence, we have

$$\oint_{C_2} \mathbf{F} \cdot d\mathbf{r} - \oint_{C_1} \mathbf{F} \cdot d\mathbf{r} = \oint_C \mathbf{F} \cdot d\mathbf{r} = \iint_R (\nabla \times \mathbf{F}) \cdot \mathbf{k} dA = 0,$$

where at the end we used that $\mathbf{F}$ has zero curl on all of $\mathbb{R}^2 \setminus \{\mathbf{0}\}$. $\square$

4.2.5 Sketching a Proof of Green's Theorem

We conclude this section by giving some insight into the proof of Green's Theorem. The argument in Step 1 will be given in full. For Step 2, however, we will only sketch the argument without providing all details.

Step 1: Suppose that $R \subset \mathbb{R}^2$ is both x -simple and y -simple. This means that

$$R := \{(x, y) : a \leq x \leq b, f(x) \leq y \leq g(x)\},$$

for some real numbers $a < b$ and continuous functions $f, g : [a, b] \rightarrow \mathbb{R}$, and that R also has a similar representation with the roles of x and y reversed.

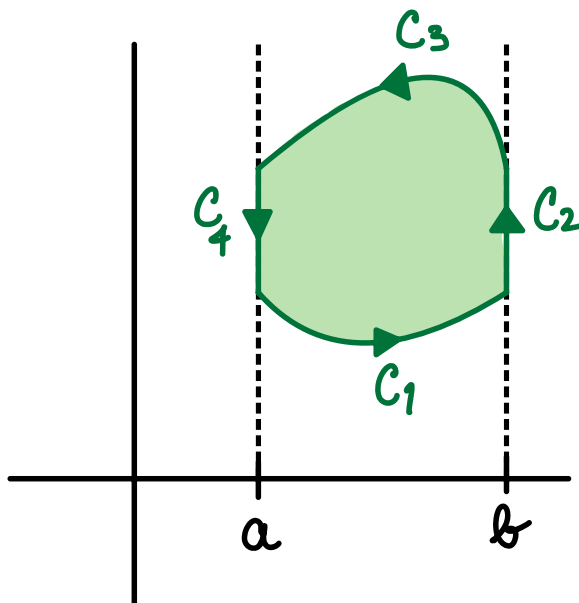
This means that $\mathcal{C}$, the boundary of R , consists of four curves $\mathcal{C}_i$ as indicated (note that $\mathcal{C}_2$ and $\mathcal{C}_4$ could be just single points). To prove Green's Theorem for the region R , then, we must show that

$$\oint_{\mathcal{C}} F_1 dx + F_2 dy = \iint_R \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) dx dy. \quad (4.2.4)$$

We will prove that

$$\oint_{\mathcal{C}} F_1 dx = \iint_R -\frac{\partial F_1}{\partial y} dx dy; \quad (4.2.5)$$

the calculation for the F_2 parts is similar but with the x and y variables interchanged.



We first remark that, since $\mathcal{C} = \mathcal{C}_1 + \mathcal{C}_2 + \mathcal{C}_3 + \mathcal{C}_4$, we can break the integral on the left-hand side of (4.2.5) into the sum of component integrals

$$\int_{\mathcal{C}_i} F_1 dx, \text{ for } i = 1, 2, 3, 4.$$

and evaluate each piece one at a time. First, the easy pieces, which are

$$\int_{\mathcal{C}_2} F_1 dx = \int_{\mathcal{C}_4} F_1 dx = 0,$$

which follows since $dx = 0$ uniformly on the line segments $\mathcal{C}_2$ and $\mathcal{C}_4$. Next, for the curves $\mathcal{C}_1$ and $\mathcal{C}_3$, we use the parametrizations $\mathbf{r}_1(t) = (t, f(t))$ and $\mathbf{r}_3(x) = (t, g(t))$ (respectively) to write

$$\int_{\mathcal{C}_1} F_1 dx = \int_a^b F_1(t, f(t)) dt \quad \text{and} \quad \int_{\mathcal{C}_3} F_1 dx = - \int_a^b F_1(t, g(t)) dt,$$

with the minus sign on $\mathcal{C}_3$ to match its orientation—that is, we start at $(b, g(b))$ and move to $(a, g(a))$. Adding up these integrals, we see that

$$\oint_{\mathcal{C}} F_1 dx = - \int_a^b [F_1(t, g(t)) - F_1(t, f(t))] dt$$

Now—and this is the key step—we apply the Fundamental Theorem of Calculus in the y -variable, to rewrite the difference as the integral of the derivative:

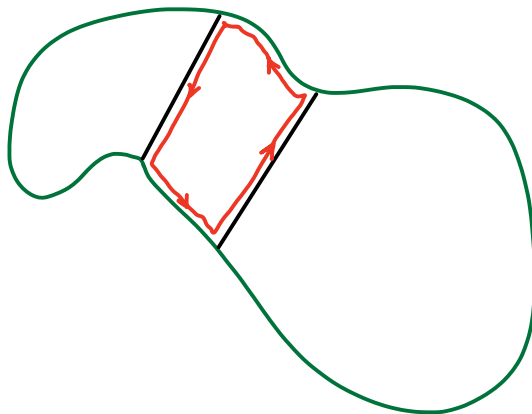
$$F_1(t, g(t)) - F_1(t, f(t)) = \int_{f(t)}^{g(t)} \frac{\partial F_1(t, y)}{\partial y} dy,$$

so that

$$\oint_C F_1 dx = - \int_a^b \int_{f(t)}^{g(t)} \frac{\partial F_1(t, y)}{\partial y} dy dt = - \iint_R \frac{\partial F_1}{\partial y} dA,$$

which matches (4.2.5).

Step 2: In general, our region R need not be either x -simple or y -simple. In most of the specific cases we will see in this course, we can write R as a union of finitely many pieces $R_1, \dots, R_m \subset \mathbb{R}^2$, pairwise disjoint except for their boundary curves, such that $R_1, \dots, R_m$ are all both x -simple and y -simple. One such piece is shown in the illustration below. The goal is to cover R by such regions so that the 2-dimensional curl integrals add up and the line integrals along the added boundary curves cancel out.



However, there are regions R that satisfy the assumptions of Green's Theorem but cannot be written as a union of finitely many pieces as above. In those cases, there are two ways we can proceed. One is to approximate R by regions R' that are finite unions of x -simple and y -simple pieces as above, so that the boundary curve of R is also approximated by the boundary curve of R' . We then apply the theorem to R' , and take the limit on both sides as $R' \rightarrow R$. (See Practice Problems for an example.) Another method involves using a change of variables to transform R , or parts of it, to regions that are x -simple and y -simple. This part of the proof is omitted here.

4.2.6 A First Glance at the Divergence Theorem

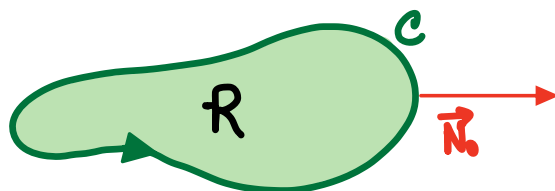
We are, in fact, within reach of our next major theorem. The Divergence Theorem works in both 2 and 3 dimensions; the 3-dimensional version will be covered shortly, but the 2-dimensional version can be stated now.

Theorem 4.2.4 (Two-Dimensional Divergence Theorem). *Let $\mathcal{C}$ be a closed, simple, counterclockwise oriented curve which bounds a region $R \subset \mathbb{R}^2$, and let $\mathbf{N}_0$ denote the*

unit normal vector to $\mathcal{C}$, oriented outward (so, not necessarily the principal normal vector, which could point inward).

Further, assume that $\mathcal{C}$, R , and a vector field $\mathbf{F} = F_1\mathbf{i} + F_2\mathbf{j}$ satisfy the hypotheses of Green's Theorem. Then

$$\iint_R \operatorname{div} \mathbf{F} \, dA = \oint_{\mathcal{C}} \mathbf{F} \cdot \mathbf{N}_0 \, ds = \oint_{\mathcal{C}} F_1 dy - F_2 dx. \quad (4.2.6)$$



The Divergence Theorem is the first equality in (4.2.6). Before giving the proof of it, we prove the second equality, which interprets the last line integral at the end of (4.2.6) as a 2-dimensional version of a flux integral.

Let us first examine what it means for $\mathbf{N}_0$ to point outward from $\mathcal{C}$. Suppose that $\mathcal{C}$ is parametrized counterclockwise as

$$\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}, \quad a \leq t \leq b,$$

then $\mathbf{r}'(t) = x'(t)\mathbf{i} + y'(t)\mathbf{j}$. We claim that the unit normal vector $\mathbf{N}_0$ satisfies

$$\mathbf{N}_0(t) = \frac{1}{|\mathbf{r}'(t)|} (y'(t)\mathbf{i} - x'(t)\mathbf{j}).$$

Indeed, we have $|\mathbf{N}_0| = 1$ and $\mathbf{N}_0 \cdot \mathbf{r}' = 0$. It remains to check that $\mathbf{N}_0$ (and not $-\mathbf{N}_0$) points outward: we determine this by considering sample points on the curve. For example, consider a point $P = \mathbf{r}(t_0)$ where $\mathbf{r}'$ points in the positive y -direction, so that $\mathbf{r}' = y'\mathbf{j}$ with $y' > 0$. Counterclockwise orientation means that D should lie to the left of P , so that $\mathbf{N}_0$ should point away from it in the direction of $\mathbf{i}$. This is consistent with the above formula.

With the aforementioned parametrization, we have

$$\begin{aligned} \oint_{\mathcal{C}} \mathbf{F} \cdot \mathbf{N}_0 \, ds &= \int_a^b \frac{1}{|\mathbf{r}'(t)|} (F_1 y'(t) + F_2 (-x'(t))) \underbrace{|\mathbf{r}'(t)|}_{ds} dt \\ &= \int_a^b [F_1 y'(t) - F_2 x'(t)] dt \\ &= \oint_{\mathcal{C}} F_1 dy - F_2 dx, \end{aligned}$$

proving the second equality in (4.2.6).

Proof of the Divergence Theorem. We apply Green's Theorem to the vector field $\mathbf{G} = -F_2\mathbf{i} + F_1\mathbf{j}$. First, note that

$$(\text{curl } \mathbf{G}) \cdot \mathbf{k} = \frac{\partial G_2}{\partial x} - \frac{\partial G_1}{\partial y} = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} = \text{div } \mathbf{F},$$

and

$$\oint_C \mathbf{G} \cdot d\mathbf{r} = \oint_C G_1 dx + G_2 dy = \oint_C F_1 dy - F_2 dx.$$

Hence, Green's Theorem applied to $\mathbf{G}$ tells us that

$$\iint_D \text{div } \mathbf{F} \, dA = \iint_D (\text{curl } \mathbf{G}) \cdot \mathbf{k} \, dA = \oint_C \mathbf{G} \cdot d\mathbf{r} = \oint_C \mathbf{F} \cdot \mathbf{N}_0 \, ds.$$

□

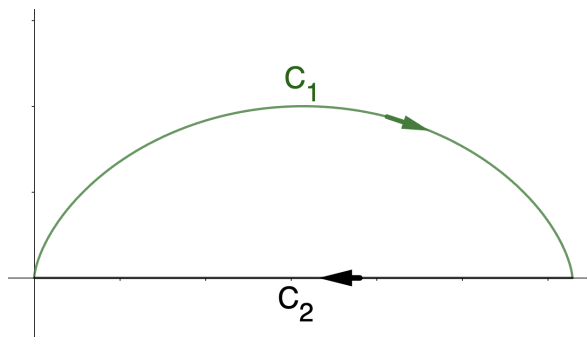
We have proven the Divergence Theorem in two dimensions. We will soon show that a similar result continues to hold in three dimensions. This, for now, concludes our discussion of Green's Theorem.

Worked Problems

1. Use Green's Theorem to calculate the area under the cycloid

$$x(t) = a(t - \sin t), \quad y(t) = a(1 - \cos t) \quad 0 \leq t \leq 2\pi,$$

and above the x axis.



Let $\mathcal{C}_1$ be the cycloid and $\mathcal{C}_2$ be the portion of the x -axis below the cycloid. The direction of *increasing* t -values corresponds to *clockwise* motion along $\mathcal{C}_1$. In order to apply Green's Theorem, we need to traverse $\mathcal{C}_1 + \mathcal{C}_2$ in the counterclockwise

direction. Let $\mathcal{C}$, then, denote $\mathcal{C}_1 + \mathcal{C}_2$, but traversed in the counterclockwise (i.e. backward) direction. Using the area formula (4.2.3) gives:

$$\begin{aligned}
 \text{Area} &= -\oint_{\mathcal{C}} y dx = \int_{\mathcal{C}_1} y dx + \int_{\mathcal{C}_2} y dx \\
 &= \int_0^{2\pi} \underbrace{a(1 - \cos t)}_{y(t)} \cdot \underbrace{a(1 - \cos t) dt}_{dx} \\
 &= \int_0^{2\pi} a^2(1 - 2\cos t + \cos^2 t) dt \\
 &= \int_0^{2\pi} a^2 \left(1 - 2\cos t + \frac{\cos 2t + 1}{2} \right) dt \\
 &= a^2 \left[t - 2\sin t + \frac{\sin 2t}{4} + \frac{t}{2} \right]_{t=0}^{2\pi} \\
 &= a^2(2\pi - 0 + \pi) \\
 &= 3\pi a^2.
 \end{aligned}$$

On the second line, we used that $\int_{\mathcal{C}_2} y dx = 0$ because $y = 0$ on $\mathcal{C}_2$.

2. (Using Green's Theorem to simplify integrals.) Evaluate the integral

$$\oint_{\mathcal{C}} (y + e^{\sqrt{x}}) dx + (2x + \cos(y^2)) dy,$$

if $\mathcal{C}$ is the union of the parabolas $y = x^2$ and $x = y^2$ (for $0 \leq x, y \leq 1$) oriented counter-clockwise.

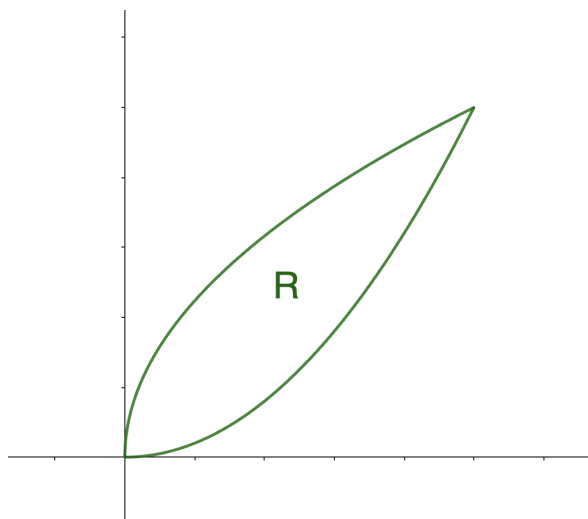
We can utilize Green's Theorem to convert the above line integral into a (much simpler) area integral on the region shown below

First, we calculate that

$$(\nabla \times \mathbf{F}) \cdot \mathbf{k} = \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} = 2 - 1 = 1.$$

Then, by Green's Theorem,

$$\begin{aligned}
 \oint_{\mathcal{C}} (y + e^{\sqrt{x}}) dx + (2x + \cos(y^2)) dy &= \int_R 1 dA \\
 &= \int_0^1 \int_{x^2}^{\sqrt{x}} dy dx = \int_0^1 (x^{1/2} - x^2) dx \\
 &= \left[\frac{x^{3/2}}{3/2} - \frac{x^3}{3} \right]_{x=0}^1 \\
 &= \frac{2}{3} - \frac{1}{3} = \frac{1}{3}.
 \end{aligned}$$



Practice Problems

1. The curve $x = \sin t$, $y = \pi^2 - t^2$, $-\pi \leq t \leq \pi$, is simple and closed. Use Green's Theorem to find the area bounded by this curve.
2. Use Green's Theorem to evaluate the line integral $\int_C (xy^2 + y^2) dx + (4xy + x^2y + \cos(y^2)) dy$, where C is the boundary of the region enclosed by the parabola $y = x^2$ and the line $y = x$, oriented counterclockwise.
3. Use Green's Theorem to evaluate the integral $\int_C y dx + x dy$, where C is the arc of the circle $(x-1)^2 + y^2 = 1$ from $(0,0)$ to $(1,1)$ in the first quadrant of the xy -plane.
4. Use Green's Theorem to evaluate the integral $\int_C y^3 dx - x^3 dy$, where C is the circle $x^2 + y^2 = 4$ oriented counterclockwise.
5. Let $\mathbf{F} = (e^{x^2} + 4x^2y)\mathbf{i} + (e^{-4y^2} - y^2x + 4x)\mathbf{j}$. Use Green's Theorem to evaluate the integral $\int_C \mathbf{F} \cdot d\mathbf{r}$, where C is the boundary of the rectangle $0 \leq x \leq 1$, $0 \leq y \leq 3$, oriented counter-clockwise.
6. Let $\mathbf{F} = (y^3 + 4x)\mathbf{i} + (3x - 4x^3)\mathbf{j}$. Find the simple, closed, oriented counterclockwise curve C for which the integral $\int_C \mathbf{F} \cdot d\mathbf{r}$ is maximized (in the sense that $\int_C \mathbf{F} \cdot d\mathbf{r} \geq \int_{C'} \mathbf{F} \cdot d\mathbf{r}$ for any other simple, closed, oriented counterclockwise curve C'). If no such curve exists, explain why.
7. Let

$$g(x) = \begin{cases} 0 & \text{if } x < 0 \text{ or } x > 1, \\ x^3 \sin(\frac{\pi}{x}) & \text{if } 0 < x \leq 1. \end{cases} \quad (4.2.7)$$

Prove that the graph of $y = g(x)$, $x \in \mathbb{R}$, is a smooth curve, with smooth parametrization $\mathbf{r}(x) = x\mathbf{i} + g(x)\mathbf{j}$. (This means that $\mathbf{r}$ is C^1 and $\mathbf{r}'$ is never zero. The main part is to prove that $\mathbf{r}'(x)$ exists, is continuous, and is nonzero at $x = 0$.) Prove further that the graph of $y = g(x)$ for $-1 \leq x \leq 1$ has finite length.

8. Let $g(x)$ be the function in (4.2.7). Let C be the closed curve, oriented counterclockwise, consisting of:

- the graph of $y = g(x)$ for $-1 \leq x \leq 1$, oriented in the direction of decreasing x ,
- the line segment from $(-1, 0)$ to $(-1, -2)$,
- the line segment from $(-1, -2)$ to $(1, -2)$,
- the line segment from $(1, -2)$ to $(1, 0)$.

Let also $\mathbf{F}$ be a C^1 vector field on $\mathbb{R}^2$, and let D be the planar region enclosed by C . (This region cannot be split into finitely many regions that are both x -simple and y -simple.) Prove that Green's Theorem remains valid in this setting, as follows.

(a) For $n = 1, 2, \dots$, let C_n be the closed curve, oriented counterclockwise, consisting of:

- the graph of $y = g(x)$ for $n^{-1} \leq x \leq 1$, oriented in the direction of decreasing x ,
- the line segment from $(n^{-1}, 0)$ to $(-1, 0)$,
- the line segment from $(-1, 0)$ to $(-1, -2)$,
- the line segment from $(-1, -2)$ to $(1, -2)$,
- the line segment from $(1, -2)$ to $(1, 0)$.

Prove that $\oint_{C_n} \mathbf{F} \cdot d\mathbf{r} \rightarrow \oint_C \mathbf{F} \cdot d\mathbf{r}$ as $n \rightarrow \infty$.

(b) For $n = 1, 2, \dots$, let D_n be the planar region enclosed by C_n . Prove that

$$\iint_{D_n} (\nabla \times \mathbf{F}) \cdot \mathbf{k} dA \rightarrow \iint_D (\nabla \times \mathbf{F}) \cdot \mathbf{k} dA \text{ as } n \rightarrow \infty.$$

(c) Assuming that Green's Theorem holds for each C_n and D_n , conclude that it also holds for C and D .

4.3 The Divergence Theorem

In Section 4.2, we used Green's Theorem to give a special case of the Divergence Theorem in two dimensions. This section develops the corresponding three-dimensional Divergence Theorem, which is the following result.

Theorem 4.3.1. *Let $D \subset \mathbb{R}^3$ be a domain bounded by a piecewise smooth closed surface $\mathcal{S}$, with unit normal vector $\mathbf{N}$ pointing out of D at all points on $\mathcal{S}$. If the vector field $\mathbf{F}$ is C^1 on some open set containing D , then*

$$\iiint_D \operatorname{div} \mathbf{F} \, dV = \oiint_{\mathcal{S}} \mathbf{F} \cdot \mathbf{N} \, dS. \quad (4.3.1)$$

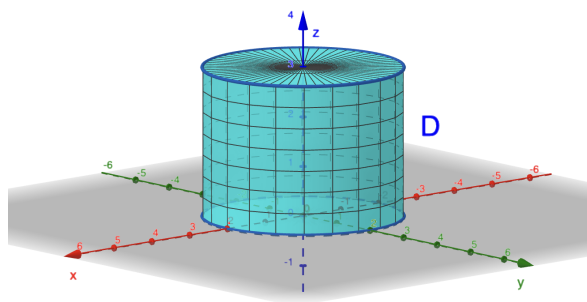
Similar to the loop notation “ $\oint_C$ ” for integrals taken over closed curves, the notation “ $\oiint_{\mathcal{S}}$ ” tells us that we are integrating over a *closed* surface in $\mathbb{R}^3$.

4.3.1 Typical Usage of the Divergence Theorem

Let us first see some typical applications of the Divergence Theorem, with concrete examples.

1. To evaluate and/or simplify complicated surface integrals.

Example 4.3.2. Suppose that $\mathbf{F} = (2x + 3e^y + 4e^{z^2})\mathbf{i} + (x - e^z)\mathbf{j} + (7\cos(y) + 2z)\mathbf{k}$. Let $\mathcal{S}$ be the surface of the circular cylinder $x^2 + y^2 \leq 4$, $0 \leq z \leq 3$, with top and bottom included. Evaluate $\oiint_{\mathcal{S}} \mathbf{F} \cdot d\mathbf{S}$.



Notice that, while $\mathbf{F}$ is rather complicated, its divergence is simple:

$$\operatorname{div} \mathbf{F} = 2 + 0 + 2 = 4,$$

a constant function over the region D enclosed by $\mathcal{S}$. In particular, the Divergence Theorem tells us that

$$\oiint_{\mathcal{S}} \mathbf{F} \cdot d\mathbf{S} = \iiint_D 4 \, dV = 4 \cdot (\text{Volume of } D) = 4 \cdot 12\pi = 48\pi.$$

Of course, we can also try a field where $\operatorname{div} \mathbf{F}$ is not constant on D ; for example, take

$$\mathbf{G} = (x^2 + \cos y)\mathbf{i} + (z^2 - \sin x)\mathbf{k} \Rightarrow \operatorname{div} \mathbf{G} = 2x + 2z,$$

with the same $\mathcal{S}$. Again, applying the Divergence Theorem simplifies the problem, since

$$\iint_{\mathcal{S}} \mathbf{G} \cdot d\mathbf{S} = \iiint_D (2x + 2z) dV = \iiint_D x dV + \iiint_D z dV.$$

By the radial symmetry of the cylinder, the first integral involving x is necessarily zero—you can check this by using polar coordinates, if you are hesitant. For the second integral, we use a trick involving averages to write

$$\begin{aligned} \iiint_D 2z dV &= 2 \cdot (\text{Volume of } D) \cdot \left(\frac{1}{(\text{Volume of } D)} \iiint_D z dV \right) \\ &= 2 \cdot (\text{Volume of } D) \cdot \bar{z}, \end{aligned}$$

where $\bar{z}$ denotes the average value of z over D . Since $0 \leq z \leq 3$, we know that $\bar{z} = \frac{3}{2}$. So,

$$\iint_{\mathcal{S}} \mathbf{F} \cdot d\mathbf{S} = 2 \cdot \frac{3}{2} \cdot 12\pi = 36\pi.$$

This shows that a seemingly complicated surface integral can, in certain cases, be reduced to an easier integral over a region. This method is generally useful when: (A) the region satisfies the hypothesis of the Divergence Theorem; and, (B) the divergence of $\mathbf{F}$ is much simpler (as a function) than $\mathbf{F}$ itself.

2. *To simplify and/or evaluate volume integrals (less common).*

Example 4.3.3. Suppose that $\mathcal{S} \subset \mathbb{R}^3$ is a cone whose base has area $A > 0$ and whose height is $h > 0$ (that is, the base may have arbitrary shape, but we know the height and the base area). Prove that the volume of the cone (regardless of the base shape) is

$$V = \frac{h \cdot A}{3}. \quad (4.3.2)$$

To show this, let us assume that the vertex of the cone is the origin, and the base is contained in the plane $z = h$ as in the picture. We then let $D \subset \mathbb{R}^3$ denote the region interior to $\mathcal{S}$.

To determine the volume of D , the trick is to apply the Divergence Theorem to a good choice of $\mathbf{F}$. It turns out that $\mathbf{F} = \frac{1}{3}(x\mathbf{i} + y\mathbf{j} + z\mathbf{k})$ ends up working for us. On one hand, its divergence is constant and equal to 1, so that

$$\iiint_D \operatorname{div} \mathbf{F} dV = \iiint_D 1 dV = \text{Volume of } D.$$

3. *Simplifying Integrals on Surfaces that Are Not Closed.* The Divergence Theorem, as it stands, is only applicable to *closed* surfaces. However, if the surface is not closed, we can try to “close” it by adding more pieces. Then the Divergence Theorem can be applied to the closed surface, and we can still benefit from the simplifications as long as the integrals on the added pieces are easier to evaluate than the original one. This is what we do in the next example.

Example 4.3.4. Evaluate $\iint_{\mathcal{S}} \mathbf{F} \cdot d\mathbf{S}$, if $\mathbf{F} = (xz + e^y)\mathbf{i} + (x^5y - z)\mathbf{j} + (x + 1)\mathbf{k}$ and $\mathcal{S}$ is the part of the sphere $x^2 + y^2 + z^2 = 1$ that lies above the plane $z = 0$, oriented so that the normal vector points away from the origin.

Let R be the top half of the ball $x^2 + y^2 + z^2 \leq 1$, and D be the disc $x^2 + y^2 \leq 1$ in the xy -plane. If we let $\mathcal{S}_1$ be the disc D oriented so that the normal vector is $-\mathbf{k}$, then $\mathcal{S}$ together with $\mathcal{S}_1$ form the oriented boundary of R . Applying the Divergence Theorem, we get

$$\iiint_R \nabla \cdot \mathbf{F} dV = \iint_{\mathcal{S}} \mathbf{F} \cdot d\mathbf{S} + \iint_D \mathbf{F} \cdot (-\mathbf{k}) dA.$$

We have $\nabla \cdot \mathbf{F} = z + x^5$, so that by symmetry and then spherical coordinates

$$\begin{aligned} \iiint_R \nabla \cdot \mathbf{F} &= \iiint_D (z + x^5) dV = \iiint_D z dV \\ &= \int_0^{2\pi} d\theta \int_0^{\pi/2} d\phi \int_0^1 R \cos \phi \cdot R^2 \sin \phi dR \\ &= 2\pi \int_0^2 R^3 dR \int_0^{\pi/2} \sin \phi \cos \phi d\phi \\ &= 2\pi \cdot \frac{R^4}{4} \Big|_0^1 \cdot \frac{\sin^2 \phi}{2} \Big|_0^{\pi/2} = 2\pi \cdot \frac{1}{4} \cdot 12 = \frac{\pi}{4}. \end{aligned}$$

Also,

$$\iint_D \mathbf{F} \cdot (-\mathbf{k}) = \iint_D (x - 1) dA = -\pi,$$

where we used that $\iint_D x dA = 0$ by symmetry. Hence

$$\iint_{\mathcal{S}} \mathbf{F} \cdot d\mathbf{S} = \frac{\pi}{4} - (-\pi) = \frac{5\pi}{4}.$$

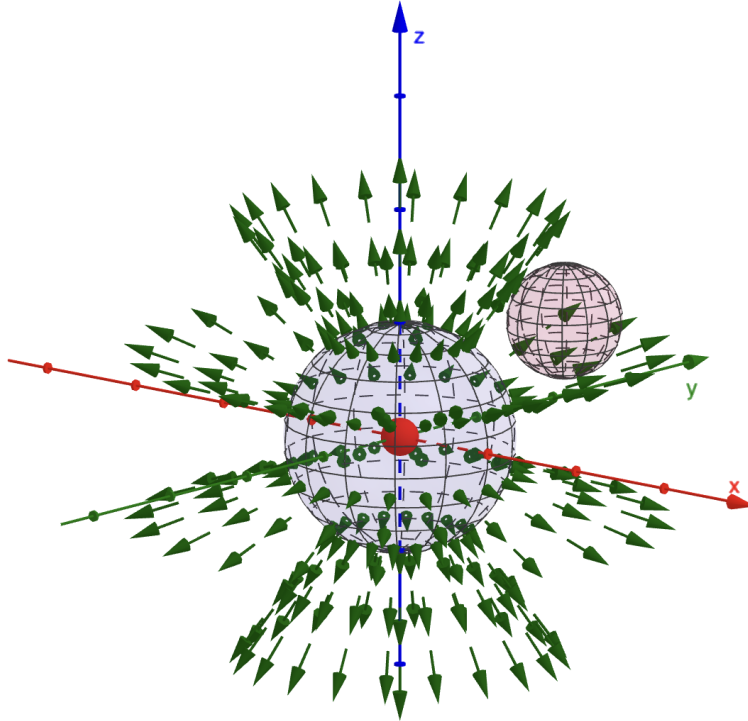
4. *Incompressible fields.* Suppose that $\mathcal{S}$ is a closed, piecewise smooth surface bounding a domain $D \subset \mathbb{R}^3$. Further suppose that $\operatorname{div} \mathbf{F} = 0$ at every single point $P \in D$. Then, the Divergence Theorem tells us that

$$\oiint_{\mathcal{S}} \mathbf{F} \cdot d\mathbf{S} = \iiint_D \operatorname{div} \mathbf{F} dV = 0.$$

However, it is important that the divergence of $\mathbf{F}$ vanish at every point of $P \in D$. Failure at even a single point can cause the above equality to fail. The following example demonstrates this phenomenon.

Example 4.3.5. Let $\mathbf{F}(\mathbf{r}) = \frac{c\mathbf{r}}{|\mathbf{r}|^3}$ in $\mathbb{R}^3$ for some $c \neq 0$ (e.g. a gravitational field). We proved in Section 4.1.1 that $\mathbf{F}$ is incompressible for all $\mathbf{r} \in \mathbb{R}^3 \setminus \{\mathbf{0}\}$.

Take two surfaces $\mathcal{S}_1$ and $\mathcal{S}_2$ that are closed, piecewise smooth, and bound domains $D_1, D_2 \subset \mathbb{R}^3$; we can assume that both D_1 and D_2 contain their boundaries as well. Assume further that D_1 contains the origin and D_2 does not. In the figure below, the vectors illustrate $\mathbf{F}$ and both D_1 and D_2 are balls.



Then:

- Since $\mathbf{0} \notin D_2$, we know that $\operatorname{div} \mathbf{F} = 0$ everywhere on D_2 , so that

$$\oiint_{\mathcal{S}_2} \mathbf{F} \cdot d\mathbf{S} = \iiint_{D_2} \operatorname{div} \mathbf{F} dV = 0.$$

This is generally true for any smooth, closed surface $\mathcal{S}$ whose interior region D does not contain the origin.

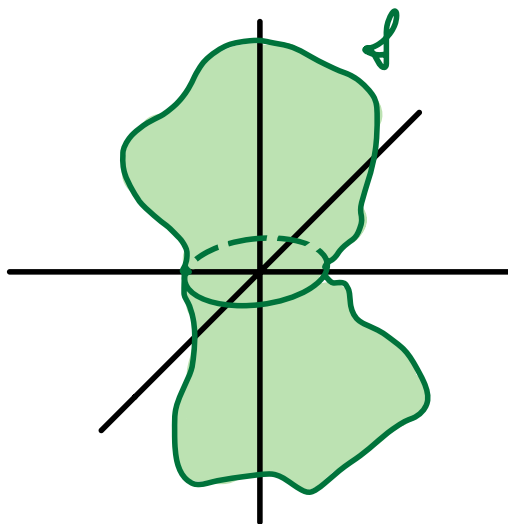
- Now, let us assume that $\mathcal{S}_1$ is the sphere $x^2 + y^2 + z^2 = a^2$ and $\mathbf{N}$ points outward. Notice that, not only does $\operatorname{div} \mathbf{F}$ fail to be zero at the origin; in fact, $\mathbf{F}$ has a singularity here. (By singularity, we mean that $|\mathbf{F}(\mathbf{r})| \rightarrow \infty$ as $\mathbf{r}$ approaches the origin.) Instead, we calculate the surface integral directly:

$$\oiint_{\mathcal{S}_1} \mathbf{F} \cdot \mathbf{N} dS = \oiint_{\mathcal{S}_1} \frac{c\mathbf{r}}{|\mathbf{r}|^3} \cdot \frac{\mathbf{r}}{|\mathbf{r}|} dS = \oiint_{\mathcal{S}_1} \frac{c}{a^2} dS,$$

using that $|\mathbf{r}| = a$ on the sphere $\mathcal{S}_1$. Therefore

$$\oiint_{\mathcal{S}_1} \mathbf{F} \cdot \mathbf{N} dS = \frac{c}{a^2} \cdot (\text{area of } \mathcal{S}_1) = 4\pi c.$$

Importantly: when $0 \in D_1$ and $\mathcal{S}_1$ is a sphere, we see that the value of the surface integral of $\mathbf{F}$ over $\mathcal{S}_1$ does not depend on the radius a ! Using the Divergence Theorem, we can show that this also continues to hold for *any closed, smooth surface* enclosing the origin. One such example is drawn below.



In terms of the δ_0 function introduced in the last lecture, this shows that

$$\operatorname{div} \mathbf{F} = 4\pi c \delta_0, \text{ whenever } \mathbf{F}(\mathbf{r}) = \frac{c\mathbf{r}}{|\mathbf{r}|^3}.$$

4.3.2 A Sketch of the Proof of Theorem 4.3.1

The main steps of the proof are similar to those for Green's Theorem. We start with two reductions.

1. First, we reduce to the case of regions that are x -simple, y -simple, and z -simple. This reduction is similar to the proof of Green's Theorem. In many cases, it suffices

to cut it up the region into finitely many pieces, prove the theorem for each of the domain-surface pairs, then add up integrals on both sides. The cutting process creates extra boundary surfaces (i.e. surfaces which are not contained in the original surface); however, the integrals on these extra surfaces will cancel out, since each surface will be counted twice with opposite orientations. This is one reason we *must* choose $\mathbf{N}$ to point outward on each component! In the general case, we may also need an additional limiting argument.

This step is presented here without proof. The remaining steps in the proof will be completely rigorous.

2. Now, suppose we are on one of these domain-surface pairs. The next step is to break the vector field $\mathbf{F}$ into its component pieces F_1, F_2 , and F_3 . We may then evaluate and compare the surface and volume integrals for each component, rather than tackling the integral of $\mathbf{F}$ all at once.

Having performed these reductions, it suffices to prove the following (much simpler) version of the Divergence Theorem.

Theorem 4.3.6. *(A special case of the Divergence Theorem) Suppose that $D \subset \mathbb{R}^3$ is the region given by*

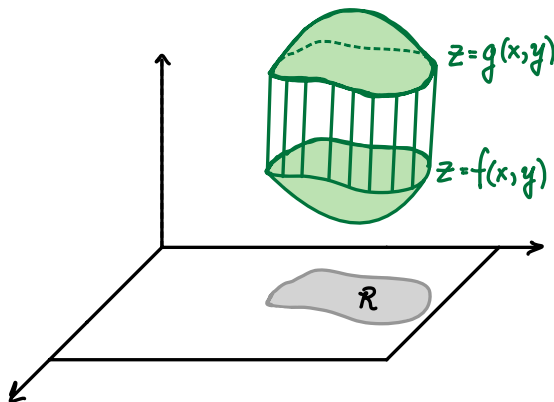
$$D = \{(x, y, z) : (x, y) \in R, f(x, y) \leq z \leq g(x, y)\},$$

where $R \subset \mathbb{R}^2$ is a 2-dimensional domain. Let $\mathbf{F} = F_3 \mathbf{k}$, and assume that f, g, F_3 are C^1 . Then

$$\iiint_D \frac{\partial F_3}{\partial z} dV = \oiint_S F_3 \mathbf{k} \cdot \mathbf{N} dS, \quad (4.3.3)$$

where S is the boundary surface of D .

Below is a small portrait of the geometric situation described:



Proof of Theorem 4.3.6. The main idea (as in the proof of Green's Theorem!) is to apply the Fundamental Theorem of Calculus in the z -direction. We first have to do some work before getting to the point where we can do that. We start with the surface integral in (4.3.3):

$$\oiint_{\mathcal{S}} F_3 \mathbf{k} \cdot \mathbf{N} dS = \iint_{\mathcal{S}_1} F_3 \mathbf{k} \cdot \mathbf{N} dS + \iint_{\mathcal{S}_2} F_3 \mathbf{k} \cdot \mathbf{N} dS + \iint_{\mathcal{S}_3} F_3 \mathbf{k} \cdot \mathbf{N} dS.$$

where $\mathcal{S}_1$ and $\mathcal{S}_2$ are the top and bottom surfaces given by the equations $z = g(x, y)$ and $z = f(x, y)$ (resp.) and $\mathcal{S}_3$ is the surface in-between. On the surface in-between (see the above drawing), we have

$$\iint_{\mathcal{S}_3} F_3 \mathbf{k} \cdot \mathbf{N} dS = 0,$$

since $\mathbf{N} \perp \mathbf{k}$ along the sides of $\mathcal{S}_3$. We are left with the top and bottom surfaces.

- For the top surface $\mathcal{S}_1$, let $\mathbf{r} = (x, y, g(x, y))$ so that $d\mathbf{S} = (-g_x \mathbf{i} - g_y \mathbf{j} + \mathbf{k})dA$. Then,

$$\iint_{\mathcal{S}_1} F_3 \mathbf{k} \cdot \mathbf{N} dS = \iint_R F_3(x, y, g(x, y)) dx dy.$$

- For the bottom surface $\mathcal{S}_2$, a similar calculation gives

$$\iint_{\mathcal{S}_2} F_3 \mathbf{k} \cdot \mathbf{N} dS = - \iint_R F_3(x, y, f(x, y)) dx dy,$$

and the minus sign is due to $\mathbf{N}$ pointing *downward* on the bottom surface.

Adding up and then applying the Fundamental Theorem of Calculus in the z variable, we get

$$\begin{aligned} \oiint_{\mathcal{S}} F_3 \mathbf{k} \cdot \mathbf{N} dS &= \iint_R \left[F_3(x, y, g(x, y)) - F_3(x, y, f(x, y)) \right] dx dy \\ &= \iint_R \left[\int_{f(x, y)}^{g(x, y)} \frac{\partial F_3}{\partial z} dz \right] dx dy \\ &= \iiint_D \frac{\partial F_3}{\partial z} dV, \end{aligned}$$

as claimed in (4.3.3). □

Practice Problems

1. Compute the flux of the vector field $\mathbf{F} = x\mathbf{i} + y\mathbf{j} - x\mathbf{k}$ out of a closed surface consisting of the part of the cylinder $x^2 + y^2 = 4$ between the planes $z = 0$ and $z = 3$, together with the discs at the top and bottom, with the normal vector pointing outward.
2. Evaluate $\iint_S \mathbf{F} \cdot d\mathbf{S}$, if $\mathbf{F} = (e^{y^2} - x^3)\mathbf{i} + (y^2 + e^{2z^2})\mathbf{j} + e^{x^2}\mathbf{k}$ and S is the sphere $x^2 + y^2 + z^2 = 1$, oriented so that the normal vector points outward.
3. Evaluate $\iint_S \mathbf{F} \cdot d\mathbf{S}$, if $\mathbf{F} = (x^2 + y^2)\mathbf{i} + z^2\mathbf{j} + \mathbf{k}$ and S is the half-sphere $x^2 + y^2 + z^2 = 1$, $z \geq 0$, oriented so that the normal vector points away from the origin. (Hint: complete the sphere to a closed surface, then apply the Divergence Theorem.)
4. Evaluate the integral $\iint_S \mathbf{F} \cdot d\mathbf{S}$, if $\mathbf{F} = y^2 \sin y \cos z \mathbf{i} + x^2 \cos x \sin z \mathbf{j} + (z^2 + 1)\mathbf{k}$ and S is the surface consisting of the cylinder $x^2 + y^2 = 1$, $0 \leq z \leq 2$, together with the disc at the top, oriented so that the normal vector points away from the origin.
5. Suppose that $g(x, y, z)$ is a smooth function on all of $\mathbb{R}^3$, that the level surface $S = \{(x, y, z) : g(x, y, z) = 0\}$ is closed, and that $\nabla g \neq 0$ on S .
 - (a) Prove that there is a point where the vector field $\mathbf{F} = \nabla g$ has nonzero divergence.
 - (b) Does there have to exist a point where $\mathbf{F} = \nabla g$ has positive divergence? If yes, explain why. If no, give a counterexample.
6. Assume that a solid D in $\mathbb{R}^3$ and its boundary S satisfy the assumptions of the Divergence Theorem. Let $f, g : \mathbb{R}^3 \rightarrow \mathbb{R}$ be C^2 functions. Prove that

$$\iint_S (f \nabla g - g \nabla f) \cdot d\mathbf{S} = \iiint_D (f \nabla^2 g - g \nabla^2 f) dV.$$

Here $\nabla^2 f = \nabla \cdot (\nabla f) = f_{xx} + f_{yy} + f_{zz}$.

7. Let $\mathbf{F}$ be an incompressible smooth vector field such that its field lines are the hyperbolas

$$x^2 + y^2 = c(z^2 + 1), \quad ax + by = 0,$$

for $c \geq 0$ and $a, b \in \mathbb{R}$, all oriented towards increasing z . (For $c = 0$, this is the z -axis.) Suppose that

$$\iint_{D_A} \mathbf{F} \cdot \mathbf{k} dS = \ln(A + 1)$$

for all $A \geq 0$, where $D_A = \{(x, y, 0) : x^2 + y^2 \leq A^2\}$.

What is $\iint_S \mathbf{F} \cdot d\mathbf{S}$, if S is the disc $\{(x, y, 4) : x^2 + y^2 \leq 9\}$ with the normal vector pointing up? Justify your answer.

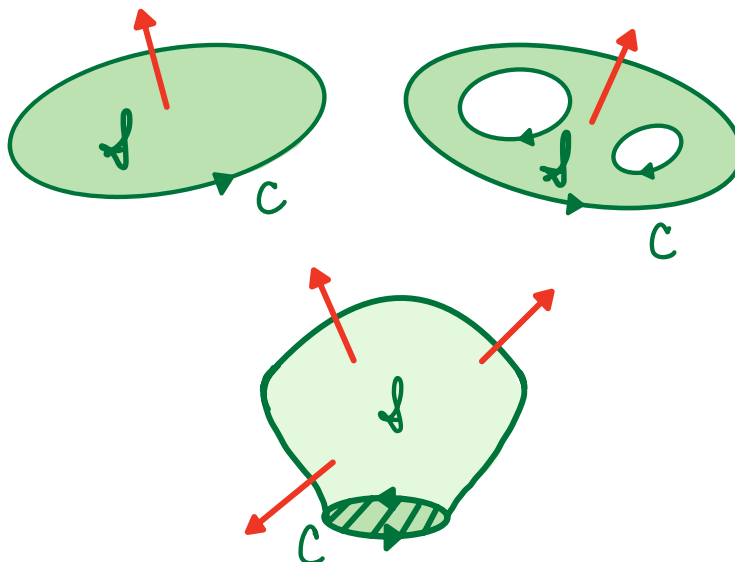
4.4 Stokes's Theorem

We conclude our discussion of fundamental theorems of vector calculus with the following result.

Theorem 4.4.1. (*Stokes's Theorem*) Let $\mathcal{S}$ be a piecewise smooth, oriented surface in $\mathbb{R}^3$, with normal $\mathbf{N}$, and suppose that the boundary $\mathcal{C}$ of $\mathcal{S}$ consists of a finite number of closed, piecewise smooth, simple curves with orientation inherited from $\mathcal{S}$. Let $\mathbf{F}$ be a C^1 vector field on an open set containing $\mathcal{S}$. Then,

$$\oint_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{r} = \iint_{\mathcal{S}} \operatorname{curl} \mathbf{F} \cdot \mathbf{N} \, dS. \quad (4.4.1)$$

When we say *inherited orientation*, we mean the boundary of $\mathcal{S}$ has counterclockwise orientation when viewed from the direction of $\mathbf{N}$ —and this orientation is *reversed* if $\mathcal{S}$ has “holes”. Some examples are shown below.



We note that Green's Theorem is a special case of Stokes's Theorem: to derive Green's Theorem from Stokes's Theorem, apply (4.4.1) with $\mathcal{S} = \mathcal{D}$ (a planar region) and $\mathbf{N} = \mathbf{k}$. In particular, if $\mathcal{C}$ is a planar curve, then the inherited orientation condition matches the assumption on the orientation of the boundary curves in Green's Theorem.

4.4.1 Stokes's Theorem and Evaluation of Integrals

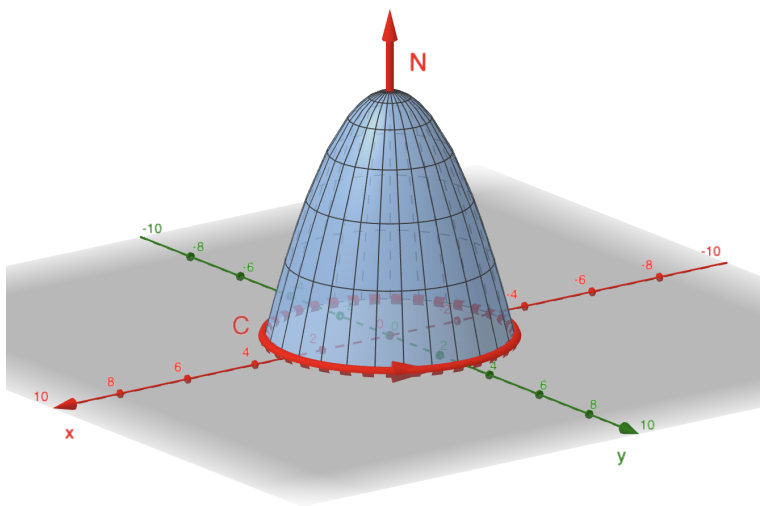
We start with examples illustrating the use of Stokes's Theorem to simplify integrals.

1. *Replacing a surface integral by a line integral.* There is no general rule, in Stokes's Theorem, as to which integral (surface or line) should be easier to evaluate. We will explore both directions. We start with an example where the line integral is simpler.

Example 4.4.2. Suppose that $\mathcal{S}$ is the part of the paraboloid $z = 9 - x^2 - y^2$ where $z \geq 0$, with $\mathbf{N}$ pointing upward, and

$$\mathbf{F} = (2z - y)\mathbf{i} + (x + z)\mathbf{j} + (3y - 2z)\mathbf{k}.$$

Importantly, we do not include the disc $x^2 + y^2 < 9$ in the surface $\mathcal{S}$! This is shown below. We want to evaluate $\iint_{\mathcal{S}} (\nabla \times \mathbf{F}) \cdot d\mathbf{S}$.



Notice that the boundary curve of $\mathcal{S}$ is the circle $x^2 + y^2 = 9$, $z = 0$, oriented counterclockwise relative to the vector $\mathbf{N} = \mathbf{k}$. Let's call this curve $\mathcal{C}$, and use the parametrization $x = 3 \cos t$, $y = 3 \sin t$ and $z = 0$ for $0 \leq t \leq 2\pi$. Then, for (x, y, z) in $\mathcal{C}$, we have

$$\mathbf{F} = -y\mathbf{i} + x\mathbf{j} + 3y\mathbf{k} = -3 \sin t \mathbf{i} + 3 \cos t \mathbf{j} + 9 \sin t \mathbf{k}.$$

Using (4.4.1), we then calculate the integral of the curl of $\mathbf{F}$ over $\mathcal{S}$ as

$$\begin{aligned} \iint_{\mathcal{S}} (\nabla \times \mathbf{F}) \cdot d\mathbf{S} &= \int_0^{2\pi} \left[\underbrace{(-3 \sin t)}_{-y} \underbrace{(-3 \sin t) dt}_{dx} + \underbrace{(3 \cos t)}_x \underbrace{(3 \cos t) dt}_{dy} \right] \\ &= \int_0^{2\pi} (9 \sin^2 t + 9 \cos^2 t) dt = \int_0^{2\pi} 9 dt \\ &= 18\pi. \end{aligned}$$

2. *Replacing a surface integral by an integral on a simpler surface with the same boundary.* Stokes's theorem tells us that the value of the integral $\iint_{\mathcal{S}} (\nabla \times \mathbf{F}) \cdot d\mathbf{S}$ for a given vector field $\mathbf{F}$ are entirely determined by the values of $\mathbf{F}$ along the boundary of $\mathcal{S}$! Specifically, we have the following corollary of Stokes's Theorem.

Corollary 4.4.3. Suppose that $\mathcal{S}_1$ and $\mathcal{S}_2$ are two oriented surfaces that share a common boundary $\mathcal{C}$ with the same induced orientation, and that they both satisfy the assumptions of Stokes's Theorem. Then

$$\iint_{\mathcal{S}_1} (\nabla \times \mathbf{F}) \cdot d\mathbf{S} = \oint_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{r} = \iint_{\mathcal{S}_2} (\nabla \times \mathbf{F}) \cdot d\mathbf{S}. \quad (4.4.2)$$

In particular, if (say) $\mathcal{S}_1$ is a complicated surface, but its boundary curve $\mathcal{C}$ is also the boundary of a simpler surface $\mathcal{S}_2$, then we can evaluate curl integrals over $\mathcal{S}_2$ as a substitute for curl integrals over $\mathcal{S}_1$.

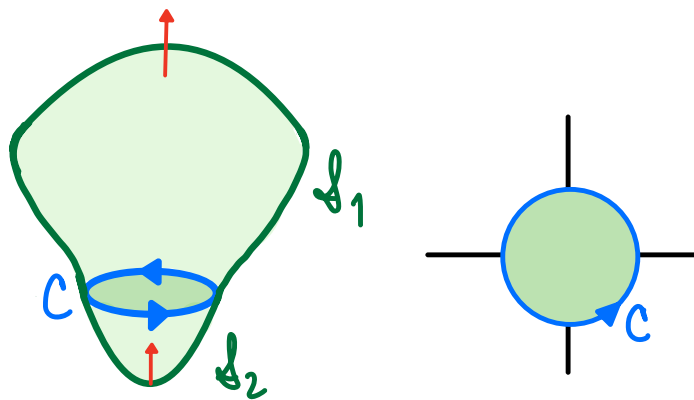


Figure 4.9: Two surfaces $\mathcal{S}_1$ and $\mathcal{S}_2$ with matching orientation (upward pointing normal) which share the same boundary curve $\mathcal{C}$.

Example 4.4.4. Evaluate the integral in Example 4.4.2 by replacing the surface $\mathcal{S}$ with a simpler surface.

The surface $\mathcal{S}$ is a part of a paraboloid bounded by the circle $\mathcal{C}$. However, $\mathcal{C}$ is also the boundary of the disc

$$\mathcal{D} := \{(x, y, 0) : x^2 + y^2 \leq 9\},$$

with normal $\mathbf{N} = \mathbf{k}$. Using the same $\mathbf{F}$ as in Example 4.4.2, we calculate

$$\nabla \times \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial_x & \partial_y & \partial_z \\ 2z - y & x + z & 3y - 2z \end{vmatrix} = 2\mathbf{i} + 2\mathbf{j} + 2\mathbf{k},$$

for every $(x, y, z) \in \mathcal{D}$. Integrating this directly over the paraboloid would require some work. However, if we integrate on the disc instead, then

$$\iint_{\mathcal{D}} (\nabla \times \mathbf{F}) \cdot d\mathbf{S} = 2 (\text{Area of } \mathcal{D}) = 2 \cdot 9\pi = 18\pi.$$

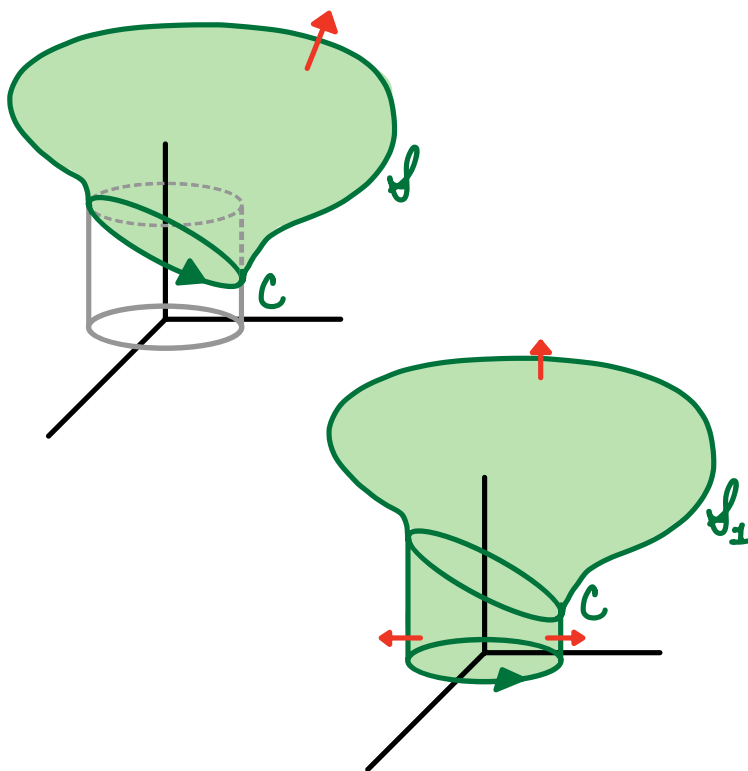
This agrees with our previous calculation using the line integral.

3. *Modifying the surface to get a simpler boundary curve, by adding pieces that do not contribute to the integral.* This is a little bit less common because additional assumptions on $\mathbf{F}$ are needed to make it work. Nonetheless, here is an example.

Example 4.4.5. Evaluate $\iint_S (\nabla \times \mathbf{F}) \cdot d\mathbf{S}$, if

$$\mathbf{F} = 2xz\mathbf{i} + (2yz + x^3)\mathbf{j} + (x^2 + y^2 + z^2)\mathbf{k}$$

and S is the surface shown below, with the boundary curve C on the cylinder $x^2 + y^2 = 4$ (shown in grey in the diagram), oriented with $\mathbf{N}$ pointing “up” so that C has counterclockwise orientation.



We could use Stokes's Theorem to convert our surface integral to $\oint_C \mathbf{F} \cdot d\mathbf{r}$. That, in general, would not help us much when we do not even know the equation for C ! But for this specific $\mathbf{F}$, we can work around it as follows. We claim that

$$\iint_S (\nabla \times \mathbf{F}) \cdot d\mathbf{S} = \iint_{S_1} (\nabla \times \mathbf{F}) \cdot d\mathbf{S},$$

with S_1 as shown above. To see this, we compute the curl:

$$\nabla \times \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial_x & \partial_y & \partial_z \\ 2xz & 2yz + x^3 & x^2 + y^2 + z^2 \end{vmatrix} = 3x^2\mathbf{k}.$$

In particular, since the normal vector of the cylinder is always parallel to the xy -plane, we know that $(\nabla \times \mathbf{F}) \cdot \mathbf{N} = 0$ on the cylinder portion of $\mathcal{S}$. In other words: the cylinder portion contributes nothing to the curl integral!

Now, the boundary of $\mathcal{S}_1$ is the circle

$$\mathcal{C}_1 := \{(x, y, 0) : x^2 + y^2 = 4\}.$$

While we could parametrize $\mathcal{C}_1$ by $\mathbf{r}(t) = \langle 2 \cos t, 2 \sin t, 0 \rangle$ for $0 \leq t \leq 2\pi$ and evaluate the integral directly, let us instead apply Green's Theorem to the disc

$$\mathcal{D} = \{(x, y) : x^2 + y^2 \leq 4\}.$$

Using symmetry and then polar coordinates, we get

$$\begin{aligned} \oint_{\mathcal{C}_1} \mathbf{F} \cdot d\mathbf{r} &= \iint_{\mathcal{D}} (\nabla \times \mathbf{F}) \cdot \mathbf{k} dA = \iint_{\mathcal{D}} 3x^2 dx dy \\ &= \frac{1}{2} \iint_{\mathcal{D}} 3(x^2 + y^2) dx dy \\ &= \frac{1}{2} \int_0^{2\pi} \int_0^2 3r^2 \cdot r dr d\theta = \frac{1}{2} \cdot 2\pi \cdot \frac{3r^4}{4} \Big|_0^2 \\ &= 12\pi. \end{aligned}$$

While the methods in the earlier examples work for *all* vector fields satisfying the hypothesis of Stokes's Theorem, here we also took advantage of the fact that the curl of $\mathbf{F}$ just happened to be parallel to the cylinder.

4.4.2 Proof of Stokes's Theorem

We seek to prove the equality

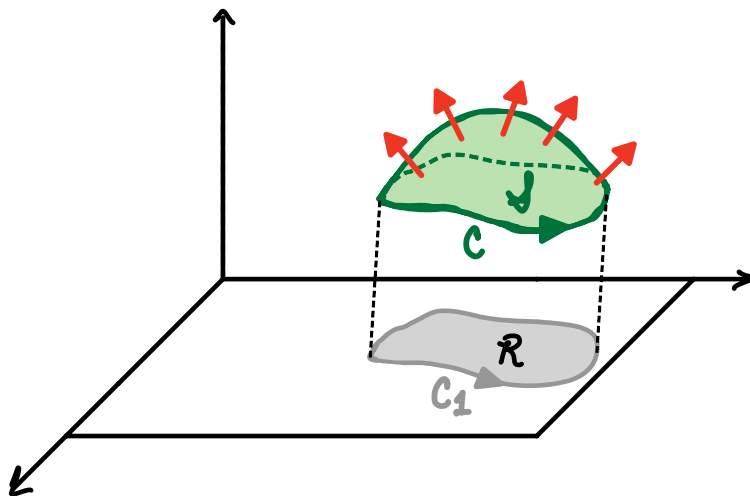
$$\oint_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{r} = \iint_{\mathcal{S}} \text{curl } \mathbf{F} \cdot \mathbf{N} dS, \quad (4.4.3)$$

whenever $\mathcal{S}$, $\mathcal{C}$, and $\mathbf{F}$ satisfy the assumptions of Stokes's Theorem.

To begin, subdivide the surface $\mathcal{S}$ into smaller surfaces so that each part has a one-to-one projection on some coordinate plane, and can be represented as a graph of a function. This gets us to the point where it suffices to consider surfaces $\mathcal{S}$ that are graphs of functions parametrized by

$$\mathbf{r}(x, y) = x\mathbf{i} + y\mathbf{j} + g(x, y)\mathbf{k},$$

with $\mathbf{N}$ pointing up, and with $(x, y) \in \mathcal{R}$ for some domain $\mathcal{R}$ in the xy -plane. An example of such a surface and its projection $\mathcal{R}$ onto the xy -axis is shown below.



In the proof below, we will also assume that the function $g(x, y)$ is C^2 on some open set containing $\mathcal{R}$. This is stronger than our usual C^1 assumption, but we may reduce to this case using an approximation argument (the details of this part are omitted), and in any case the C^2 assumption will be satisfied in all applications of Stokes's Theorem in this course.

With the above parametrization,

$$\mathbf{n} = -g_x \mathbf{i} - g_y \mathbf{j} + \mathbf{k}, \text{ for all } (x, y) \in \mathcal{R}. \quad (4.4.4)$$

Let $\mathbf{F} = F_1 \mathbf{i} + F_2 \mathbf{j} + F_3 \mathbf{k}$. On the boundary curve $\mathcal{C}$, we know that $z = g(x, y)$, so that $dz = g_x dx + g_y dy$ by the Chain Rule. Therefore

$$\begin{aligned} \oint_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{r} &= \oint_{\mathcal{C}_1} F_1 dx + F_2 dy + F_3(g_x dx + g_y dy) \\ &= \oint_{\mathcal{C}_1} [(F_1 + F_3 g_x) dx + (F_2 + F_3 g_y) dy], \end{aligned} \quad (4.4.5)$$

where we understand that the component functions F_j are each being evaluated at the points $(x, y, g(x, y))$ for all points along $\mathcal{C}$. Applying Green's Theorem then gives

$$\oint_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{r} = \iint_{\mathcal{R}} \left[\frac{\partial}{\partial x} (F_2 + F_3 g_y) - \frac{\partial}{\partial y} (F_1 + F_3 g_x) \right] dA. \quad (4.4.6)$$

To compute the partial derivatives, we need to use the Chain Rule, keeping in mind that $z = g(x, y)$ so that each F_j depends on x and y both directly and indirectly through z . We get

$$\frac{\partial}{\partial x} (F_2 + F_3 g_y) = \frac{\partial F_2}{\partial x} + \frac{\partial F_2}{\partial z} g_x + \frac{\partial F_3}{\partial x} g_y + \underbrace{\frac{\partial F_3}{\partial z} g_x g_y + F_3 g_{yx}}_{(*)} \quad (4.4.7)$$

$$\frac{\partial}{\partial y}(F_1 + F_3 g_x) = \frac{\partial F_1}{\partial y} + \frac{\partial F_1}{\partial z} g_y + \frac{\partial F_3}{\partial y} g_x + \underbrace{\frac{\partial F_3}{\partial z} g_y g_x + F_3 g_{xy}}_{(**)}. \quad (4.4.8)$$

Since g is C^2 , we have equality of the mixed second order partials appearing in $(*)$ and $(**)$, so that we actually have $(*) = (**)$! Hence

$$(4.4.7) - (4.4.8) = \left(\frac{\partial F_2}{\partial x} + \frac{\partial F_2}{\partial z} g_x + \frac{\partial F_3}{\partial x} g_y \right) - \left(\frac{\partial F_1}{\partial y} + \frac{\partial F_1}{\partial z} g_y + \frac{\partial F_3}{\partial y} g_x \right).$$

Grouping by the terms g_x , g_y and reversing the signs of g_x and g_y , the above becomes

$$\left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z} \right) (-g_x) + \left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x} \right) (-g_y) + \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right),$$

so that

$$\begin{aligned} \oint_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{r} &= \iint_{\mathcal{R}} \left[\left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z} \right) (-g_x) \right. \\ &\quad \left. + \left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x} \right) (-g_y) + \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) \right] dx dy \\ &= \iint_{\mathcal{S}} \operatorname{curl} \mathbf{F} \cdot d\mathbf{S}, \end{aligned}$$

which is what we needed to prove. At the last step we used the normal vector computed in (4.4.4).

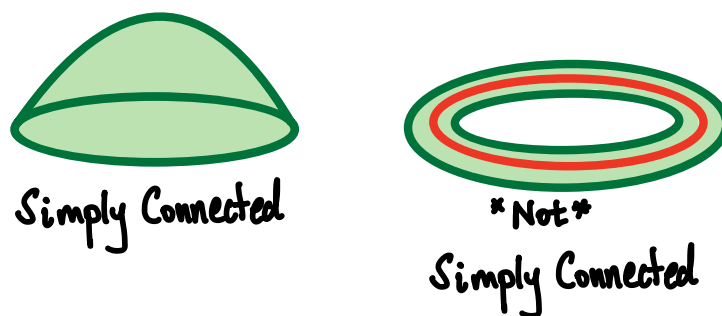
4.4.3 Conservative Vector Fields and Stokes's Theorem

Definition 4.4.6. A 3-dimensional domain $\mathcal{D} \subset \mathbb{R}^3$ is simply connected if every closed, simple, piecewise smooth curve $\mathcal{C}$ in $\mathcal{D}$ bounds a piecewise smooth surface $\mathcal{S}$ contained in $\mathcal{D}$.

Intuitively, $\mathcal{D}$ is simply connected if any closed curve in $\mathcal{D}$ can be contracted continuously to a point without leaving $\mathcal{D}$. (In topology, simple connectedness is usually defined without the smoothness conditions.) Thinking about simply connected domains as “free from holes” is less useful here, since a “hole” can refer to at least two different mathematical concepts in 3 dimensions. For example, a donut-shaped region is not simply connected (see the picture below), but a spherical annulus such as $\{(x, y, z) : 1 \leq x^2 + y^2 + z^2 \leq 4\}$ in $\mathbb{R}^3$ is simply connected.

We are now in a position to prove the following result.

Theorem 4.4.7. If $\mathcal{D} \subset \mathbb{R}^3$ is a simply connected domain, $\mathbf{F}$ is smooth on $\mathcal{D}$, and $\operatorname{curl} \mathbf{F} = \mathbf{0}$ everywhere on $\mathcal{D}$, then $\mathbf{F}$ is conservative.



Proof. By Stokes's Theorem and the simply connectedness of $\mathcal{D}$, for any closed, simple, piecewise smooth curve $\mathcal{C} \subset \mathcal{D}$, we have

$$\oint_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{r} = \iint_S \text{curl } \mathbf{F} \cdot d\mathbf{S} = 0.$$

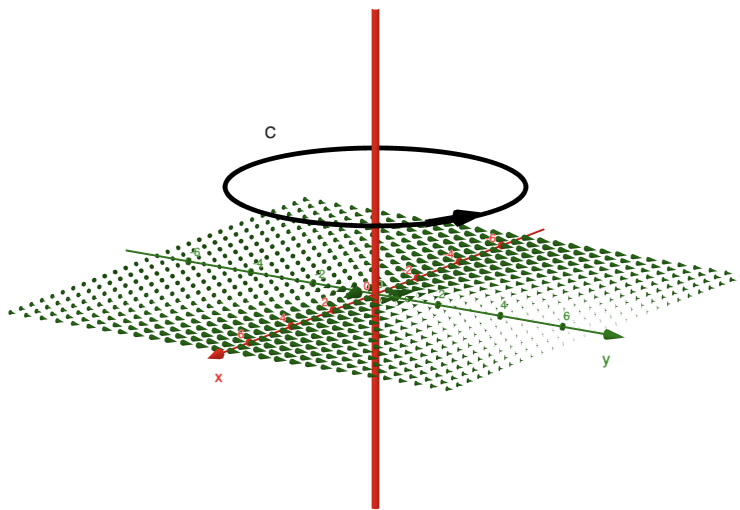
So, $\mathbf{F}$ is conservative on $\mathcal{D}$. □

The assumption that D is simply connected is necessary in the above theorem. The following example shows this.

Example 4.4.8. Consider the vector field

$$\mathbf{F}(x, y, z) := -\frac{y}{x^2 + y^2} \mathbf{i} + \frac{x}{x^2 + y^2} \mathbf{j},$$

defined on the domain $\mathcal{D} := \mathbb{R}^3 \setminus \{(0, 0, z) : z \in \mathbb{R}\}$. This is shown below. The function $\mathbf{F}$ satisfies $\text{curl } \mathbf{F} = \mathbf{0}$ everywhere on $\mathcal{D}$; however, $\mathcal{D}$ is not simply connected.



We have seen previously that

$$\oint_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{r} = 2\pi$$

for any circle $x^2 + y^2 = a^2$ in the xy -plane, oriented counterclockwise. The same calculation works for circles $x^2 + y^2 = a^2$, $z = b$, lying in a plane parallel to the xy -plane. Actually a stronger result is true: the same formula holds for any closed curve $\mathcal{C}_0$ in $\mathbb{R}^3$ circling the z -axis once counterclockwise! This can be proved by constructing a surface $\mathcal{S}$ “wrapped around” the z -axis, with boundary consisting of $\mathcal{C}_0$ and a circle $\mathcal{C}_1$ of the form $x^2 + y^2 = a^2$, $z = b$, and then applying Stokes’s Theorem to that surface. The figures below illustrate the construction of $\mathcal{S}$. We orient $\mathcal{S}$ so that $\mathbf{N}$ points away from the z -axis; then the oriented boundary of $\mathcal{S}$ consists of $\mathcal{C}_0$ oriented counterclockwise and $\mathcal{C}_1$ oriented clockwise (can you explain why?). By Stokes’s Theorem, we then have

$$0 = \iint_{\mathcal{S}} \operatorname{curl} \mathbf{F} \cdot d\mathbf{S} = \oint_{\mathcal{C}_0} \mathbf{F} \cdot d\mathbf{r} + \oint_{\mathcal{C}_1} \mathbf{F} \cdot d\mathbf{r} = \oint_{\mathcal{C}_0} \mathbf{F} \cdot d\mathbf{r} - 2\pi,$$

so that the last integral is 2π . This proves the claim.

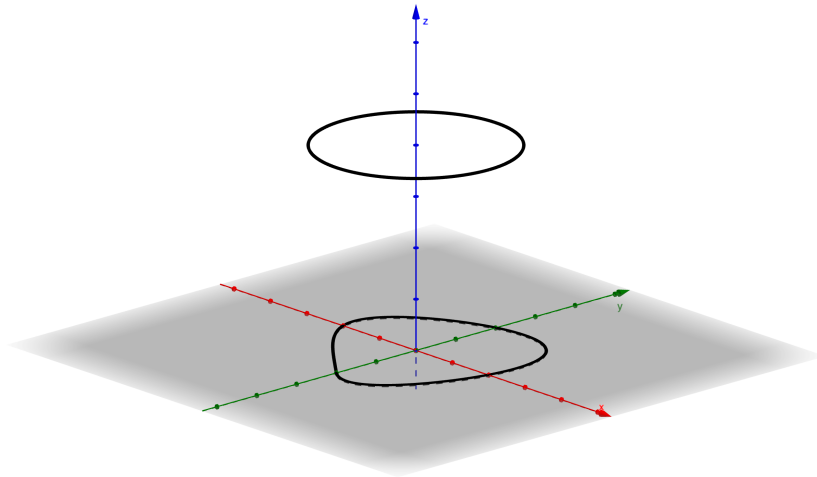


Figure 4.10: Two curves in $\mathbb{R}^3$. The lower curve $\mathcal{C}_0$ is contained in the plane $z = 0$ and wraps around the z -axis once counterclockwise. The upper curve is the circle $\mathcal{C}_1 := \{(\cos t, \sin t, 1) : 0 \leq t \leq 2\pi\}$.

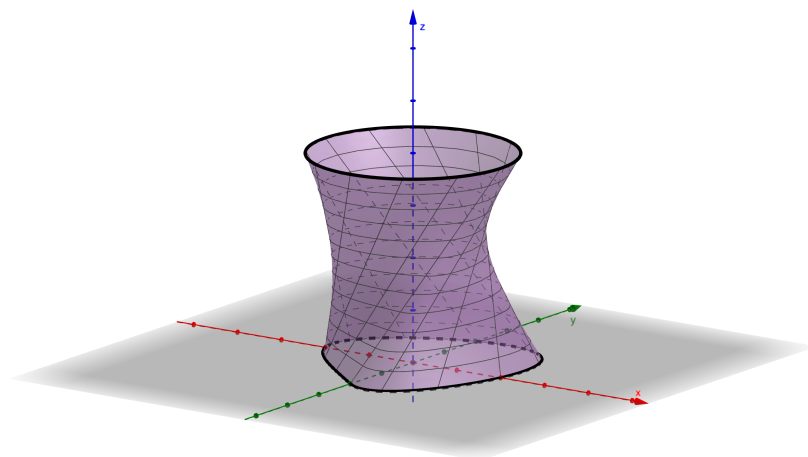


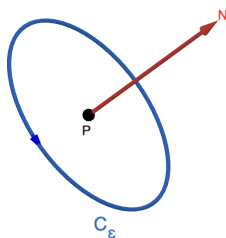
Figure 4.11: A surface $\mathcal{S}$ bounded by the curves $\mathcal{C}_0$ and $\mathcal{C}_1$. We chose $\mathcal{C}_0$ to lie in the xy -plane so that it would be easier to draw a picture, but this is not necessary for the construction.

4.4.4 The infinitesimal version

As a last consideration, let $\mathcal{C}_\epsilon$ denote a circle of small radius $\epsilon > 0$, in a plane perpendicular to $\mathbf{N}$ and oriented counterclockwise when viewed from the direction of $\mathbf{N}$, and let $\mathcal{S}_\epsilon$ denote the disc bounded by $\mathcal{C}_\epsilon$. Then, by Stokes's Theorem

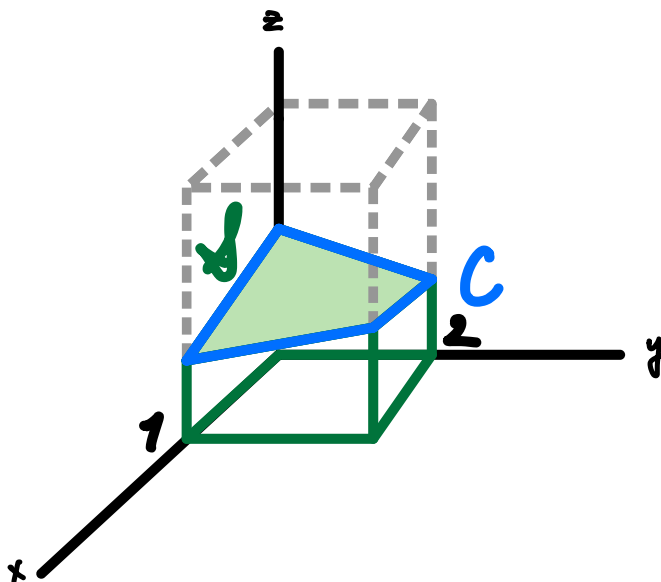
$$\oint_{\mathcal{C}_\epsilon} \mathbf{F} \cdot d\mathbf{r} = \iint_{\mathcal{S}_\epsilon} \text{curl } \mathbf{F} \cdot \mathbf{N} \, dS \approx \left(\text{curl } \mathbf{F}(P) \cdot \mathbf{N} \right) \left(\text{Area of } \mathcal{S}_\epsilon \right),$$

where P denotes the centre of the circle $\mathcal{C}_\epsilon$, as shown below. In this sense, the curl measures the “rotating” effect of $\mathbf{F}$, and the rotation is *strongest* whenever $\mathbf{N}$ is parallel to $\text{curl } \mathbf{F}$. This matches the formula from Section 4.1.



Worked Problems

1. Evaluate $\oint_C \mathbf{F} \cdot d\mathbf{r}$ if C is the intersection of the graph $z = g(x, y)$ and the boundary of $0 \leq x \leq 1$ and $0 \leq y \leq 2$, oriented counterclockwise, and $\mathbf{F} = e^{-x}\mathbf{i} + e^x\mathbf{j} + e^z\mathbf{k}$.



Let $\mathcal{S}$ be the surface bounded by the curve C . We parametrize $\mathcal{S}$ by

$$\mathbf{r}(x, y) = \langle x, y, g(x, y) \rangle, \text{ for } 0 \leq x \leq 1, 0 \leq y \leq 2.$$

Then, $d\mathbf{S} = (-g_x\mathbf{i} - g_y\mathbf{j} + \mathbf{k})dxdy$ and also

$$\nabla \times \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial_x & \partial_y & \partial_z \\ e^{-x} & e^x & e^z \end{vmatrix} = e^x\mathbf{k}.$$

Applying Stokes's Theorem, we get

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \iint_S e^x\mathbf{k} \cdot d\mathbf{S} = \int_0^1 \int_0^2 e^x dy dx = 2e^x \Big|_0^1 = 2(e - 1).$$

2. Evaluate $\iint_S (\nabla \times \mathbf{F}) \cdot d\mathbf{S}$, if $\mathbf{F} = (x^3 + y)e^z\mathbf{i} + y^2e^{2z}\mathbf{j} + e^z\mathbf{k}$ and $\mathcal{S}$ is the half-sphere $x^2 + y^2 + z^2 = 9$, $z \geq 0$, oriented so that the normal vector points away from the origin.

We will apply Stokes's Theorem. The curl of $\mathbf{F}$ is

$$\nabla \times \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial_x & \partial_y & \partial_z \\ (x^3 + y)e^z & y^2e^{2z} & e^z \end{vmatrix} = -2y^2e^{2z}\mathbf{i} + (x^3 + y)e^z\mathbf{j} - e^z\mathbf{k}.$$

The easiest way to solve the problem is to replace $\mathcal{S}$ by the disk $D = \{(x, y, 0) : x^2 + y^2 \leq 9\}$, with $\mathbf{N} = \mathbf{k}$, which has the same oriented boundary curve. On D , we have $(\nabla \times \mathbf{F}) \cdot \mathbf{N} = -e^z = -1$, so that

$$\iint_{\mathcal{S}} (\nabla \times \mathbf{F}) \cdot d\mathbf{S} = \iint_D (\nabla \times \mathbf{F}) \cdot d\mathbf{S} = -(\text{area of } D) = -9\pi.$$

Practice Problems

1. Evaluate $\iint_{\mathcal{S}} \text{curl } \mathbf{F} \cdot d\mathbf{S}$, if $\mathbf{F} = (x^4 - 2y)\mathbf{i} + (\cos y)\mathbf{j} + z \sin(xy)\mathbf{k}$ and $\mathcal{S}$ is the part of the paraboloid $z = 9 - x^2 - y^2$ that lies above the xy -plane.
2. Use Stokes's Theorem to evaluate $\int_C (x-z)dx + (x+y)dy + (y+z)dz$, if C is the ellipse where the plane $z = y$ intersects the cylinder $x^2 + y^2 = 1$, oriented counterclockwise as viewed from above.
3. Evaluate the integral $\int_C \mathbf{F} \cdot d\mathbf{r}$, where C is the curve of intersection of the cylinder $x^2 + y^2 = 1$ and the paraboloid $z = 100 - (x-2)^2 - (y+4)^2$, oriented counterclockwise when viewed from above, and $\mathbf{F} = (3x^2y + y^3)\mathbf{i} + z^4\mathbf{k}$.
4. Let R be the cube whose one face is the square in the xy -plane with vertices $(\pm 1, \pm 1, 0)$, and the opposite face lies in the plane $z = 2$ and has vertices $(\pm 1, \pm 1, 2)$. Let

$$\mathbf{F} = xyz^2\mathbf{i} + xy^2\mathbf{j} + x^3yz\mathbf{k}.$$

Evaluate $\iint_{\mathcal{S}} (\nabla \times \mathbf{F}) \cdot d\mathbf{S}$, if $\mathcal{S}$ is the surface consisting of the 4 vertical faces and the top of R (without the bottom), oriented so that the normal vector points away from the origin.

Chapter 5

Differential Forms

5.1 An Introduction to Differential Forms

We will end this course with a short introduction to differential forms. Our goal is to develop a unified framework for the main concepts (line integrals, flux integrals) and theorems (Green's, Divergence, Stokes's) that we have covered. It turns out that all of those can be viewed as special cases of “Generalized Stokes's Theorem”, which also extends to arbitrary dimensions.

We will work from the bottom up, using coordinate systems and concrete examples at every step. This will allow us to ground the abstract theory in our earlier work in 2 and 3 dimensions. This also means that we define some of the main concepts (wedge product, exterior derivative) in terms of how they are computed. The same concepts could also be defined in a coordinate-free way (in terms of their functional properties), but that would require building up more theory and then doing additional work to derive the coordinate representation. There is not enough time for it in this course, but we encourage the interested student to consider further study in differential geometry.

5.1.1 Manifolds: Definitions and Examples

We will provide a rigorous definition of a manifold a little bit later; for the purpose of this introduction, it suffices to think of manifolds as a common term for objects like curves in $\mathbb{R}^n$, surfaces in $\mathbb{R}^3$, hypersurfaces in $\mathbb{R}^n$, and so on. It also includes 0-dimensional objects (points) and n -dimensional domains in $\mathbb{R}^n$.

A manifold might or might not have a boundary. For example, the boundary of a ball

$$\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq a^2\}$$

is the sphere

$$\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = a^2\}.$$

Stokes's Theorem involves a surface $\mathcal{S}$ and a curve C on its boundary. The boundary of a curve C from P to Q consists of its endpoints P, Q . Generally, if our manifold is k -dimensional, its boundary (if non-empty) has dimension $k - 1$. An example of a manifold with no boundary is a closed surface or curve.

Let's first look at a few types of integrals on manifolds that we have seen.

- If $f : D \rightarrow \mathbb{R}$ is a function defined on a domain $D \subset \mathbb{R}^n$, we could fix a point $P \in D$ and consider $f(P)$ (the value of f at P) as the “integral” of f on the zero-dimensional set P . In fact, using definitions from analysis courses such as Math 320, $f(P)$ is the integral of f with respect to the “delta-measure” at P —an idea we encountered in connection with the Divergence Theorem and Stokes's Theorem for surfaces.
- In single-variable calculus, we consider integrals $\int_a^b f(x)dx$ of functions $f : [a, b] \rightarrow \mathbb{R}$, where $[a, b] \subset \mathbb{R}$ is an interval. Moreover, we have the Fundamental Theorem of Calculus: if $f = F'$ then

$$\int_a^b f(x)dx = F(b) - F(a).$$

This gives a relation between the integral of a function F on the 0-dimensional set of endpoints $\{a, b\}$ and the integral of another function f , related to F through differentiation, on a manifold of dimension 1. (The set of endpoints $\{a, b\}$ is “oriented” in the sense that we take the *difference* $F(b) - F(a)$, so that it matters which term has the minus sign.)

- If $\mathbf{F}$ is a vector field on a domain $D \subset \mathbb{R}^n$, and if C is an oriented curve in D connecting the point P to the point Q , we can consider the integral

$$\int_C \mathbf{F} \cdot d\mathbf{r}.$$

Moreover, if $\mathbf{F} = \nabla f$ is conservative, then

$$\int_C \mathbf{F} \cdot d\mathbf{r} = f(Q) - f(P),$$

a relation similar to the Fundamental Theorem of Calculus.

- *Green's Theorem*: If $\mathbf{F}$ is a vector field and if D is a 2-dimensional domain bounded by a closed curve C , then

$$\iint_D \mathbf{curl} \mathbf{F} \cdot \mathbf{k} dA = \int_C \mathbf{F} \cdot d\mathbf{r},$$

a relation between a one-dimensional line integral of a vector field and a two-dimensional integral of a function involving the derivatives of the component functions of that vector field. The Fundamental Theorem of Calculus was featured prominently in the proof.

- *The Divergence Theorem:* If $\mathbf{F}$ is a vector field on $\mathbb{R}^3$ and if D is a domain bounded by a closed surface $\mathcal{S}$, then

$$\iiint_D \operatorname{div} \mathbf{F} \, dV = \iint_{\mathcal{S}} \mathbf{F} \cdot d\mathbf{S}.$$

Again, this relates an integral of a vector field on the boundary of a domain (this time, 2-dimensional) to the integral of a function derived from it on that domain (3-dimensional).

The justification for introducing differential forms is to establish a common framework for all of these. A *differential form* is going to be “something we can integrate on a manifold”. Differential forms will have a certain dimensionality that has to be matched to the dimension of the manifold where we want to integrate them. For example, functions on domains or vector fields on curves can be rewritten as differential forms satisfying these dimensionality conditions.

We will work towards *Generalized Stokes’s Theorem*—a relation between integrals

$$\int_{\partial\mathcal{M}} \omega = \int_{\mathcal{M}} d\omega,$$

where ω is a differential form, $d\omega$ is another differential form derived from it, $\mathcal{M}$ is a manifold, and $\partial\mathcal{M}$ is the boundary of $\mathcal{M}$. Without thinking (for now) about what a differential form is or what it means to integrate it, you can check that all of the above theorems have this general form.

5.1.2 Differential k -forms

We introduce the basic language used to define differential forms and operations on them. While the general theory works in $\mathbb{R}^n$ for any $n \geq 1$, the case $n = 4$ is the lowest dimension in which you will see new phenomena that we did not cover previously. In this section, we will usually work out 2-dimensional and 3-dimensional cases of our general formulas in detail, to ensure that our abstract general theory matches the theorems we already know in $\mathbb{R}^2$ and $\mathbb{R}^3$. Additionally, we will often refer to examples in dimension $n = 4$, to make sure we go beyond our previous work.

We start with k -forms with constant coefficients. In general, we will use $\mathbf{x} = (x_1, x_2, \dots, x_n)$ to denote vectors in $\mathbb{R}^n$, but we will also continue to use the coordinates (x, y) in $\mathbb{R}^2$ and (x, y, z) in $\mathbb{R}^3$ as needed.

- *1-forms* with constant coefficients are linear mappings $\mathbb{R}^n \rightarrow \mathbb{R}$ (in functional analysis, such mappings are called *linear functionals*), written as $\omega = a_1 dx_1 + \cdots + a_n dx_n$, acting on vectors via

$$dx_i(\mathbf{v}) = v_i$$

and then extended via linearity so that $(a_1 dx_1 + \cdots + a_n dx_n)(\mathbf{v}) = \sum a_i v_i$. For example, if $\omega = 3dx_1 - 5dx_4$ in $\mathbb{R}^4$, and if $\mathbf{v} = \langle 1, 2, 3, 4 \rangle$, then

$$\omega(\mathbf{v}) = 3 \cdot 1 - 5 \cdot 4 = 3 - 20 = -17.$$

- *2-forms* with constant coefficients are functions $\phi : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ (so that ϕ acts on two vectors simultaneously) that are *bilinear*:

$$\phi(a\mathbf{v} + b\mathbf{w}, \mathbf{u}) = a\phi(\mathbf{v}, \mathbf{u}) + b\phi(\mathbf{w}, \mathbf{u}),$$

$$\phi(\mathbf{u}, a\mathbf{v} + b\mathbf{w}) = a\phi(\mathbf{u}, \mathbf{v}) + b\phi(\mathbf{u}, \mathbf{w}),$$

and *antisymmetric*:

$$\phi(\mathbf{v}, \mathbf{u}) = -\phi(\mathbf{u}, \mathbf{v}).$$

The elementary 2-forms are

$$dx_i \wedge dx_j(\mathbf{u}, \mathbf{v}) = \begin{vmatrix} u_i & v_i \\ u_j & v_j \end{vmatrix} = \begin{vmatrix} u_i & u_j \\ v_i & v_j \end{vmatrix}.$$

By the basic properties of determinants, we have

$$dx_i \wedge dx_j = -dx_j \wedge dx_i, \quad dx_i \wedge dx_i = 0. \quad (5.1.1)$$

(A determinant is 0 if two of its rows or columns are identical, and changes sign if two of its rows or columns are interchanged.)

Example 5.1.1. If $\mathbf{v} = \langle 1, 2, 3, 4 \rangle$ and $\mathbf{w} = \langle -5, 6, -7, 8 \rangle$, then

$$(2dx_1 \wedge dx_4 - dx_2 \wedge dx_3)(\mathbf{v}, \mathbf{w}) = 2 \begin{vmatrix} 1 & -5 \\ 4 & 8 \end{vmatrix} - \begin{vmatrix} 2 & 6 \\ 3 & -7 \end{vmatrix} = 24.$$

Example 5.1.2. Let $\phi = dx_2 \wedge dx_3 + dx_3 \wedge dx_1 + dx_1 \wedge dx_2$ in $\mathbb{R}^3$, then

$$\phi(\mathbf{u}, \mathbf{v}) = \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix} + \begin{vmatrix} u_3 & u_1 \\ v_3 & v_1 \end{vmatrix} + \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix}.$$

The formal similarity to the cross product of two vectors is 100% intended; in fact, we will use differential forms to generalize the concept of the cross-product to higher dimensions.

- *k-forms* with constant coefficients are defined as mappings

$$\phi : \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{k \text{ times}} \rightarrow \mathbb{R},$$

generalizing the above to k vector-valued variables. We continue to require *multilinearity*,

$$\begin{aligned} \phi(\mathbf{v}_1, \dots, \mathbf{v}_{i-1}, a\mathbf{v}_i + b\mathbf{v}'_i, \mathbf{v}_{i+1}, \dots, \mathbf{v}_k) \\ = a\phi(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_k) + b\phi(\mathbf{v}_1, \dots, \mathbf{v}'_i, \dots, \mathbf{v}_k) \end{aligned}$$

and *antisymmetry*,

$$\phi(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_{i'}, \dots, \mathbf{v}_k) = -\phi(\mathbf{v}_1, \dots, \mathbf{v}_{i'}, \dots, \mathbf{v}_i, \dots, \mathbf{v}_k).$$

In words: multilinearity requires ϕ to be linear in each variable; antisymmetry requires ϕ to change signs if any two variables are exchanged (e.g. first vector is switched with the last; 2nd vector switched with 8th vector). We use $\Lambda_k(\mathbb{R}^n)$ to denote the linear space of all k -forms with constant coefficients on $\mathbb{R}^n$.

Elementary k forms are the basic building blocks for k -forms. They are:

$$dx_{i_1} \wedge \cdots \wedge dx_{i_k}(\mathbf{v}_1, \dots, \mathbf{v}_k) := \begin{vmatrix} v_{1i_1} & v_{2i_1} & \cdots & v_{ki_1} \\ v_{1i_2} & v_{2i_2} & \cdots & v_{ki_2} \\ \vdots & \vdots & \ddots & \vdots \\ v_{1i_k} & v_{2i_k} & \cdots & v_{ki_k} \end{vmatrix}$$

and any k -form can be written as a linear combination of such forms. Another way to write the above formula is

$$dx_{i_1} \wedge \cdots \wedge dx_{i_k}(\mathbf{v}_1, \dots, \mathbf{v}_k) = \det(dx_{i_s}(\mathbf{v}_t))_{1 \leq s, t \leq k},$$

if we remember that the 1-form dx_i were defined as $dx_i(\mathbf{x}) = x_i$. We give an example below to illustrate this.

Example 5.1.3. Let $\mathbf{u} = \langle 1, 2, 3, 4 \rangle$, $\mathbf{v} = \langle 5, 6, 7, 8 \rangle$, and $\mathbf{w} = \langle 1, -1, 1, -1 \rangle$, then

$$dx_1 \wedge dx_3 \wedge dx_4(\mathbf{u}, \mathbf{v}, \mathbf{w}) = \begin{vmatrix} 1 & 5 & 1 \\ 3 & 7 & 1 \\ 4 & 8 & -1 \end{vmatrix}$$

As in (5.1.1), an elementary k -form is zero if two of the indices i_j are equal, and changes sign if two of the expressions dx_{i_j} and $dx_{i_{j'}}$ are interchanged. For example, $dx \wedge dy \wedge dx = 0$ and $dx \wedge dy \wedge dz = -dz \wedge dy \wedge dx$.

Example 5.1.4. The only nonzero elementary n -form in $\mathbb{R}^n$, up to $\pm$ signs, is $dx_1 \wedge dx_2 \wedge \cdots \wedge dx_n$. If $\mathbf{v}_1, \dots, \mathbf{v}_n$ are vectors in $\mathbb{R}^n$, then

$$\mathbf{v}_1, \dots, \mathbf{v}_n \text{ are linearly independent} \Leftrightarrow dx_1 \wedge \cdots \wedge dx_n(\mathbf{v}_1, \dots, \mathbf{v}_n) \neq 0.$$

(See the determinant formula above.)

A general **differential k -form** is a linear combination of the elementary k -forms defined above, but we allow variable coefficients which are functions of $\mathbf{x} = (x_1, \dots, x_n)$. For example,

$$x^2 dx \wedge dy, \quad ydx - zdy, \quad x_2 \cos(x_3) dx_1 \wedge dx_3 \wedge dx_4$$

are differential forms. Unless we say explicitly otherwise, we will assume that all coefficient functions are C^∞ on their domain D , meaning that all partial derivatives of any order exist and are continuous on D . Many of the results below are valid under weaker assumptions, for example Proposition 5.2.4 below remains true if the coefficient functions are C^2 instead of C^∞ . The point of the C^∞ assumption is to allow us to focus on the concepts without getting distracted by counting derivatives.

When $k = 0$, a differential 0-form is simply a C^∞ function $D \rightarrow \mathbb{R}$. For $k \geq 1$, a general differential k -form on a domain $D \subset \mathbb{R}^n$ can be written as

$$\Phi(\mathbf{x}) = \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} a_{i_1 \dots i_k}(\mathbf{x}) dx_{i_1} \wedge \cdots \wedge dx_{i_k},$$

where the coefficients $a_{i_1 \dots i_k}$ are C^∞ functions on D . Notice that we only use the indices $1 \leq i_1 < i_2 < \cdots < i_k \leq n$ in the summation. This is because, first, $dx_{i_1} \wedge \cdots \wedge dx_{i_k} = 0$ if any two of the indices are equal. Second, for any remaining non-zero terms, we can rearrange the dx_i 's so that the indices come in increasing order, at the cost of possibly multiplying that term by -1 . For instance,

$$x_1^2 dx_3 \wedge dx_4 \wedge dx_1 = -x_1^2 dx_1 \wedge dx_4 \wedge dx_3 = x_1^2 dx_1 \wedge dx_3 \wedge dx_4,$$

where at each step we swapped two of the dx_i 's and changed the sign of the form.

Similarly to the coefficients, we will also allow the *arguments* of differential forms to depend on $\mathbf{x}$, so that differential k -form may act on a k -tuple of vector fields (rather than just k -tuples of fixed vectors). For example, if $\mathbf{F} = e^x \mathbf{i} - e^y \mathbf{j} + e^{2z} \mathbf{k}$ and $\mathbf{G} = x\mathbf{i} + x^2\mathbf{j} + x^3\mathbf{k}$, then

$$ydx \wedge dz(\mathbf{F}, \mathbf{G}) = y \begin{vmatrix} e^x & e^{2z} \\ x & x^3 \end{vmatrix} = y(x^3 e^x - x e^{2z}).$$

5.1.3 Exterior (wedge) product

If $\phi \in \Lambda_k(\mathbb{R}^n)$ and $\psi \in \Lambda_l(\mathbb{R}^n)$ (so that ϕ is a k -form and ψ is an l -form on $\mathbb{R}^n$), we define their wedge product $\phi \wedge \psi \in \Lambda_{k+l}(\mathbb{R}^n)$ as follows:

- if ϕ and ψ are elementary forms, just concatenate them and write a wedge in between,
- use linearity to extend it to more general forms.

For example, if $\phi = dx_1 \wedge dx_2$ and $\psi = dx_3 \wedge dx_4 \wedge dx_5$, then

$$\phi \wedge \psi = dx_1 \wedge dx_2 \wedge dx_3 \wedge dx_4 \wedge dx_5.$$

A couple of (slightly) more complicated examples are below.

Example 5.1.5. Suppose $\phi = a_1 dx_1 + a_2 dx_2 + a_3 dx_3$ and $\psi = b_1 dx_1 + b_2 dx_2$. Then,

$$\begin{aligned} \phi \wedge \psi &= \cancel{a_1 b_1 dx_1 \wedge dx_1}^0 + a_2 b_1 dx_2 \wedge dx_1 + a_3 b_1 dx_3 \wedge dx_1 \\ &\quad + a_1 b_2 dx_1 \wedge dx_2 + \cancel{a_2 b_2 dx_2 \wedge dx_2}^0 + a_3 b_2 dx_3 \wedge dx_2 \\ &= (a_1 b_2 - a_2 b_1) dx_1 \wedge dx_2 + a_3 b_1 dx_3 \wedge dx_1 + a_3 b_2 dx_3 \wedge dx_2. \end{aligned}$$

Example 5.1.6. Let $\phi = a_1 dx_1 + a_2 dx_2 + a_3 dx_3$ and

$$\psi = b_1(dx_2 \wedge dx_3) + b_2(dx_3 \wedge dx_1) + b_3(dx_1 \wedge dx_2).$$

Then,

$$\phi \wedge \psi = (a_1 b_1 + a_2 b_2 + a_3 b_3)(dx_1 \wedge dx_2 \wedge dx_3).$$

(Exercise: verify this!)

Here are a couple of examples with variable coefficients.

Example 5.1.7. We work in $\mathbb{R}^3$, with coordinates $\mathbf{r} = (x, y, z)$. Then $\Phi = 3xz dx \wedge dy + y^2 dx \wedge dz$ is a differential 2-form in 3 variables, and $\Psi := e^x dx + e^z dz$ is a differential 1-form in 3 variables. Their wedge product is

$$\begin{aligned} \Phi \wedge \Psi &= (3xz dx \wedge dy + y^2 dx \wedge dz) \wedge (e^x dx + e^z dz) \\ &= \cancel{3xe^x dx \wedge dy \wedge dx}^0 + \cancel{y^2 e^x dx \wedge dz \wedge dx}^0 \\ &\quad + 3xe^z dx \wedge dy \wedge dz + \cancel{y^2 e^z dx \wedge dz \wedge dz}^0 \\ &= 3xe^z dx \wedge dy \wedge dz. \end{aligned}$$

Three terms in this calculation cancel to zero because $dx \wedge dx = dz \wedge dz = 0$. Since Φ was a 2-form and Ψ was a 1-form, the resulting wedge product $(\Phi \wedge \Psi)$ is a 3-form.

Example 5.1.8. Let $\Phi = xy dx + y dz$ and $\Psi = x dx + z^2 dy$ be differential 1-forms in 3 variables. Then, their wedge product is the 2-form

$$\begin{aligned} \Phi \wedge \Psi &= (xy dx + y dz) \wedge (x dx + z^2 dy) \\ &= xyz dx \wedge dx + xy z^2 dx \wedge dy + y z^2 dz \wedge dy, \end{aligned}$$

where again we used that $dx \wedge dx = 0$.

Worked Problems

1. Let $\mathbf{u} = (1, 2, -1, 3)$, $\mathbf{v} = (3, 0, 1, 0)$, $\mathbf{w} = (2, 2, 1, -1)$. Evaluate:

(a) $dx_1 \wedge dx_3 \wedge dx_4(\mathbf{u}, \mathbf{v}, \mathbf{w})$,

$$\begin{aligned} dx_1 \wedge dx_2 \wedge dx_4(\mathbf{u}, \mathbf{v}, \mathbf{w}) &= \begin{vmatrix} 1 & 3 & 2 \\ -1 & 1 & 1 \\ 3 & 0 & -1 \end{vmatrix} \\ &= 3 \begin{vmatrix} 3 & 2 \\ 1 & 1 \end{vmatrix} - \begin{vmatrix} 1 & 3 \\ -1 & 1 \end{vmatrix} \\ &= 3(1) - 1(4) \\ &= -1 \end{aligned}$$

where we expanded the 3×3 determinant along the third row.

(b) $2dx_1 \wedge dx_2(\mathbf{u}, \mathbf{v}) - dx_3 \wedge dx_4(\mathbf{u}, \mathbf{w})$.

$$\begin{aligned} 2dx_1 \wedge dx_2(\mathbf{u}, \mathbf{v}) - dx_3 \wedge dx_4(\mathbf{u}, \mathbf{w}) &= 2 \begin{vmatrix} 1 & 3 \\ 2 & 0 \end{vmatrix} - \begin{vmatrix} -1 & 1 \\ 3 & -1 \end{vmatrix} \\ &= 2(-6) - (-2) \\ &= -12 + 2 = 10 \end{aligned}$$

2. Calculate:

(a) $\Phi(\mathbf{F}, \mathbf{G})$, if $\Phi = ydx \wedge dy - z^2dx \wedge dz$, $\mathbf{F} = \cos x\mathbf{i} + \sin x\mathbf{j}$, and $\mathbf{G} = xy\mathbf{j} + yz\mathbf{k}$.

$$\begin{aligned} \Phi(\mathbf{F}, \mathbf{G}) &= ydx \wedge dy(\mathbf{F}, \mathbf{G}) - z^2dx \wedge dz(\mathbf{F}, \mathbf{G}) \\ &= y \begin{vmatrix} \cos x & 0 \\ \sin x & xy \end{vmatrix} - z^2 \begin{vmatrix} \cos x & 0 \\ 0 & yz \end{vmatrix} \\ &= xy^2 \cos x - yz^3 \cos x. \end{aligned}$$

(b) $\Psi(\mathbf{F}, \mathbf{G}, \mathbf{H})$, if $\Psi = dx \wedge dy \wedge dz$, $\mathbf{F}$ and $\mathbf{G}$ are as in (a), and $\mathbf{H} = 2y\mathbf{i} - x\mathbf{k}$.

$$\begin{aligned} \Psi(\mathbf{F}, \mathbf{G}, \mathbf{H}) &= dx \wedge dy \wedge dz(\mathbf{F}, \mathbf{G}, \mathbf{H}) = \begin{vmatrix} \cos x & 0 & 2y \\ \sin x & xy & 0 \\ 0 & yz & -x \end{vmatrix} \\ &= \cos x \begin{vmatrix} xy & 0 \\ yz & -x \end{vmatrix} + 2y \begin{vmatrix} \sin x & xy \\ 0 & yz \end{vmatrix} \\ &= -x^2y \cos x + 2y^2z \sin x. \end{aligned}$$

3. Calculate and simplify the wedge product $\phi \wedge \psi$. Write the dx_i 's in the answers in increasing subscript order.

(a) $\phi = 2dx_1 \wedge dx_3 - 5dx_3 \wedge dx_4$, $\psi = -dx_1 \wedge dx_2 + 4dx_2 \wedge dx_4$

$$\begin{aligned}\phi \wedge \psi &= (2dx_1 \wedge dx_3 - 5dx_3 \wedge dx_4) \wedge (-dx_1 \wedge dx_2 + 4dx_2 \wedge dx_4) \\ &= 5dx_3 \wedge dx_4 \wedge dx_1 \wedge dx_2 + 8dx_1 \wedge dx_3 \wedge dx_2 \wedge dx_4 \\ &= -5dx_1 \wedge dx_4 \wedge dx_3 \wedge dx_2 - 8dx_1 \wedge dx_2 \wedge dx_3 \wedge dx_4 \\ &= 5dx_1 \wedge dx_2 \wedge dx_3 \wedge dx_4 - 8dx_1 \wedge dx_2 \wedge dx_3 \wedge dx_4 \\ &= -3dx_1 \wedge dx_2 \wedge dx_3 \wedge dx_4.\end{aligned}$$

The other terms vanish because they each have two appearances of some dx_i .

(b) $\phi = dx_1 \wedge dx_3 \wedge dx_4$, $\psi = 7dx_2 - dx_3$

$$\begin{aligned}\phi \wedge \psi &= (dx_1 \wedge dx_3 \wedge dx_4) \wedge (7dx_2 - dx_3) \\ &= 7dx_1 \wedge dx_3 \wedge dx_4 \wedge dx_2 \\ &= -7dx_1 \wedge dx_2 \wedge dx_4 \wedge dx_3 \\ &= -7dx_1 \wedge dx_2 \wedge dx_3 \wedge dx_4.\end{aligned}$$

Practice Problems

- Let $\mathbf{u} = (1, -1, 0, 5)$, $\mathbf{v} = (3, 2, 6, -1)$, $\mathbf{w} = (5, -2, 3, 1)$. Evaluate $dx_1 \wedge dx_3 \wedge dx_4(\mathbf{u}, \mathbf{v}, \mathbf{w})$.
- Let $\mathbf{u} = (2, 1, 4, -1)$, $\mathbf{v} = (1, 3, 5, 0)$, $\mathbf{w} = (2, 7, -2, 3)$. Evaluate:
 - $dx_1 \wedge dx_2 \wedge dx_4(\mathbf{u}, \mathbf{v}, \mathbf{w})$,
 - $dx_2 \wedge dx_3(\mathbf{u}, \mathbf{v}) - 2dx_1 \wedge dx_4(\mathbf{u}, \mathbf{w})$.
- Calculate and simplify the wedge product $\phi \wedge \psi$. Write the dx_i 's in the answers in increasing subscript order.
 - $\phi = dx_1 \wedge dx_3 \wedge dx_4$, $\psi = a_1dx_1 + a_2dx_2 + a_3dx_3$.
 - $\phi = dx_2 \wedge dx_3 + dx_1 \wedge dx_4$, $\psi = 4dx_1 \wedge dx_2 - dx_1 \wedge dx_3 + 5dx_2 \wedge dx_3$.

4. Calculate and simplify the wedge product $\phi \wedge \psi$. Write the dx_i 's in the answers in increasing subscript order.
- (a) $\phi = 3dx_2 \wedge dx_3 - 7dx_2 \wedge dx_4$, $\psi = 2dx_1 + 3dx_2 - dx_4$.
- (b) $\phi = dx_2 \wedge dx_3 \wedge dx_4$, $\psi = 3dx_1 - 5dx_2$
- (c) $\phi = 2dx_1 \wedge dx_3$, $\psi = -2dx_1 \wedge dx_4 + 3dx_2 \wedge dx_4$.
5. Compute the exterior product $\Phi_1 \wedge \Phi_2$, if $\Phi_1 = x_1 dx_1 \wedge dx_2 \wedge dx_4$ and $\Phi_2 = x_4 dx_1 + x_2 dx_3 + x_3 dx_4$ in $\mathbb{R}^4$. Give your answer in the form $f(x_1, x_2, x_3, x_4) dx_1 \wedge dx_2 \wedge dx_3 \wedge dx_4$.

5.2 Exterior Derivative

5.2.1 Defining the exterior derivative

We now define an operation that will generalize several of the differential operators (such as the gradient, curl, and divergence) we used previously.

Definition 5.2.1. *The **exterior derivative** of a differential 0-form $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ is the differential 1-form*

$$df(\mathbf{x}) := \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i.$$

Formally, this is the same as the differential of a function f that we used in earlier parts of this course, but now we interpret it as a differential form.

For a differential k -form $\Phi(\mathbf{x}) = a(\mathbf{x}) dx_{i_1} \wedge \cdots \wedge dx_{i_k}$, its exterior derivative $d\Phi$ is the differential $(k+1)$ -form given by

$$d\Phi(\mathbf{x}) := da(\mathbf{x}) \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k},$$

with $da(\mathbf{x})$ defined as above. We then extend it to general k -forms by linearity.

For 1-forms, this means $d(f dx_i) = df \wedge dx_i$, with df being the differential of the function f . This should get us a 2-form. For 2-forms, we have $d(f dx_i \wedge dx_j) = df \wedge dx_i \wedge dx_j$ (a 3-form), and so on. We then use linearity to extend it to the general case. Here is a concrete example.

Example 5.2.2. *We have*

$$d(xy) = ydx + xdy \qquad d(x^2 + y^2) = 2xdx + 2ydy,$$

and

$$\begin{aligned}
 d(xydx + (x^2 + y^2)dy) &= d(xy) \wedge dx + d(x^2 + y^2) \wedge dy \\
 &= (ydx + xdy) \wedge dx + (2xdx + 2ydy) \wedge dy \\
 &= xdy \wedge dx + 2xdx \wedge dy \\
 &= xdx \wedge dy,
 \end{aligned}$$

where at the last step we used that $dy \wedge dx = -(dx \wedge dy)$. Note that if we take another exterior derivative, we have

$$d(xdx \wedge dy) = dx \wedge dx \wedge dy = 0.$$

The exterior derivative generalizes several of the definitions we have used before, if vector fields are replaced by differential forms with the same component functions. This will be important in matching our previous fundamental theorems (i.e. Green's, Divergence, Stokes) to the appropriate special cases in Generalized Stokes's Theorem.

(1) Gradient: If f is a 0-form on $\mathbb{R}^n$, then

$$df = \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i$$

is a differential 1-form with coefficients $\frac{\partial f}{\partial x_i}$. Note that these coefficients match the corresponding components of the gradient vector field ∇f . If we apply the differential 1-form df to a unit vector $\mathbf{u}$, we get

$$df(\mathbf{u}) = \sum_{i=1}^n \frac{\partial f}{\partial x_i} u_i = D_{\mathbf{u}}f,$$

the directional derivative of f in the direction of $\mathbf{u}$.

(2) Curl in $\mathbb{R}^3$: Let $\mathbf{F} = F_1\mathbf{i} + F_2\mathbf{j} + F_3\mathbf{k}$ be a smooth vector field on $\mathbb{R}^3$. Can we match $\mathbf{F}$ to a differential form on $\mathbb{R}^3$ with the same coefficients F_1, F_2, F_3 ? One natural choice is

$$\Phi = F_1dx + F_2dy + F_3dz,$$

but a second choice that might be a little bit less obvious is to use the 2-form

$$F_1dy \wedge dz + F_2dz \wedge dx + F_3dx \wedge dy.$$

We illustrate the usefulness of both types of matching as follows. Let $\mathbf{G} = \text{curl } \mathbf{F}$. In components, we have $\mathbf{G} = G_1\mathbf{i} + G_2\mathbf{j} + G_3\mathbf{k}$. Then, with Φ defined as above, we have

$$d\Phi = G_1dy \wedge dz + G_2dz \wedge dx + G_3dx \wedge dy.$$

We show this calculation in the Worked Problems Section below. Therefore, curl is a special case of exterior derivative if we use the first kind of matching for $\mathbf{F}$ and the second kind for $\text{curl } \mathbf{F}$.

(3) Divergence in $\mathbb{R}^3$: Suppose that $\mathbf{G}$ is a smooth vector field on $\mathbb{R}^3$. Let

$$\Psi = G_1 dy \wedge dz + G_2 dz \wedge dx + G_3 dx \wedge dy.$$

Then

$$d\Psi = (\text{div } \mathbf{G})(dx \wedge dy \wedge dz).$$

We will show the calculation in the Worked Problems Section below. (Exercise: how would you modify this example to make it work in $\mathbb{R}^2$?)

5.2.2 Closed and exact differential k -forms

We are ready for our first common generalization of two different results about vector fields.

Definition 5.2.3. A differential k -form Φ is called **closed** if $d\Phi = 0$ and **exact** if there is some $(k-1)$ -form Ψ such that $\Phi = d\Psi$.

Proposition 5.2.4. Let Φ be a differential k -form on some domain $D \subset \mathbb{R}^n$, and a and b are real numbers. Then

$$d^2(\Phi) := d(d\Phi) = 0, \tag{5.2.1}$$

where 0 denotes the zero differential form. Put another way, every exact form is closed.

Outline of proof. The proof uses the equality of the mixed partials. By linearity, it is enough to prove (5.2.1) when $\Phi = a(\mathbf{x})dx_{i_1} \wedge \cdots \wedge dx_{i_k}$. Suppose for the moment that Φ is a 0-form, so that $\Phi = a(\mathbf{x})$ is a smooth function on $\mathbb{R}^n$. Then

$$d\Phi = da = \sum_{i=1}^n \frac{\partial a}{\partial x_i} dx_i.$$

By direct computation (you should check this!),

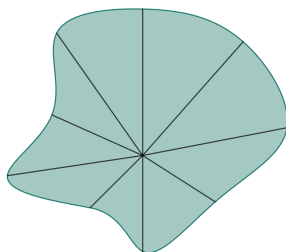
$$d(d\Phi) = \sum_{i < j} \left(\frac{\partial^2 a}{\partial x_i \partial x_j} - \frac{\partial^2 a}{\partial x_j \partial x_i} \right) dx_i \wedge dx_j,$$

with coefficients identically zero by the equality of mixed partials. Hence $d(d\Phi) = 0$. The general case is similar, but with more “trailing” dx_i ’s that do not change the calculation. We leave it as an exercise for you to write out the details. $\square$

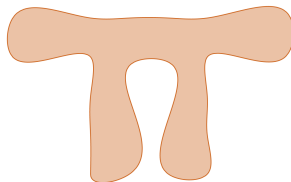
For the converse, we have the following result, which is known as *Poincaré's Lemma*.

Theorem 5.2.5. *Let Φ be a closed differential k -form on a **starlike** domain $D \subset \mathbb{R}^k$ (see below). Then Φ is exact on D .*

The domain D is starlike if there exists a point $P \in D$ such that every other point in D can be joined to P by a line segment that lies entirely in D . There may be more than one such point—for example, if D is convex, we may take P to be any point in D . However, starlike domains do not have to be convex.



(a) A starlike domain



(b) A simply connected, non-starlike domain

Figure 5.1: Starlike is a stronger condition than being simply connected.

We omit the proof of the theorem. Instead, we discuss how the above results apply to vector fields in 3 dimensions.

- (1) Conservative fields: Let f be a smooth function on $\mathbb{R}^3$, and let $\mathbf{F} = \nabla f$. Define the 1-form

$$\Phi = df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz = F_1dx + F_2dy + F_3dz.$$

Proposition 5.2.4 tells us that $d\Phi = d(df) = 0$. But we have seen that

$$d\Phi = G_1dy \wedge dx + G_2dz \wedge dx + G_3dx \wedge dy,$$

where $\mathbf{G} = \text{curl } \mathbf{F}$. Hence, $d\Phi = 0$ means that $\text{curl } \mathbf{F} = \mathbf{0}$.

Proposition 5.2.4 thus includes, as a special case, the fact that conservative vector fields have zero curl. Poincaré's Lemma says that, on a starlike domain D , the converse is true: if $\mathbf{F}$ has zero curl, it must in fact be conservative.

- (2) Divergence and Curl: Let $\mathbf{F}$ be a smooth vector field in $\mathbb{R}^3$, and let $\mathbf{G} = \text{curl } \mathbf{F}$. We have seen that the differential forms

$$\Phi = F_1 dx + F_2 dy + F_3 dz, \text{ and } \Psi := G_1 dy \wedge dx + G_2 dz \wedge dx + G_3 dx \wedge dy.$$

satisfy $d\Phi = \Psi$. Proposition 5.2.4 then says that

$$d\Psi = d(d\Phi) = 0.$$

But, we have calculated that $d\Psi = (\text{div } \mathbf{G}) dx \wedge dy \wedge dz$! So, in this case, Proposition 5.2.4 shows that the divergence of curl is zero. For starlike domains D , Poincaré's Lemma provides a converse: if a smooth vector field $\mathbf{G}$ satisfies $\text{div } \mathbf{G} = 0$ on D , then $\mathbf{G} = \text{curl } \mathbf{F}$ for some smooth vector field $\mathbf{F}$ on D .

There are important special cases when the starlike assumption in Poincaré's Lemma can be weakened and the conclusion still holds. As an example, we already know that if $n = 2$ or 3 and if the domain $D \subset \mathbb{R}^n$ is simply connected, then $\text{curl } \mathbf{F} = 0$ on D implies that $\mathbf{F}$ is conservative on D . The property of being simply connected is strictly weaker than the property of being starlike, in the sense that all starlike domains are simply connected, but the converse is not true. Starlike domains are a general class of domains for which Poincaré's Lemma holds for *all differential forms*, regardless of order and ambient dimension.

Worked Problems

1. Calculate the exterior derivative of the differential forms. Write the dx_i 's in the answers in increasing subscript order.

(a) $\Phi = x_1^2 e^{x_3} dx_1 \wedge dx_2 + (x_2 + x_3) dx_3 \wedge dx_4$

$$\begin{aligned} d\Phi &= d(x_1^2 e^{x_3} dx_1 \wedge dx_2 + (x_2 + x_3) dx_3 \wedge dx_4) \\ &= d(x_1^2 e^{x_3}) \wedge dx_1 \wedge dx_2 + d(x_2 + x_3) \wedge dx_3 \wedge dx_4 \\ &= (2x_1 e^{x_3} dx_1 + x_1^2 e^{x_3} dx_3) \wedge dx_1 \wedge dx_2 + (dx_2 + dx_3) \wedge dx_3 \wedge dx_4 \\ &= x_1^2 e^{x_3} dx_3 \wedge dx_1 \wedge dx_2 + dx_2 \wedge dx_3 \wedge dx_4 \\ &= x_1^2 e^{x_3} dx_1 \wedge dx_2 \wedge dx_3 + dx_2 \wedge dx_3 \wedge dx_4. \end{aligned}$$

(b) $\Phi = x_1 x_2^2 \cos(x_3) dx_1 \wedge dx_2 \wedge dx_3 + x_2^2 \sin(x_3) dx_1 \wedge dx_3 \wedge dx_4.$

$$\begin{aligned} d\Phi &= d(x_1 x_2^2 \cos(x_3)) \wedge dx_1 \wedge dx_2 \wedge dx_3 + d(x_2^2 \sin(x_3)) \wedge dx_1 \wedge dx_3 \wedge dx_4 \\ &= 2x_2 \sin(x_3) dx_2 \wedge dx_1 \wedge dx_3 \wedge dx_4 \\ &= -2x_2 \sin(x_3) dx_1 \wedge dx_2 \wedge dx_3 \wedge dx_4. \end{aligned}$$

2. Let $\mathbf{F} = F_1\mathbf{i} + F_2\mathbf{j} + F_3\mathbf{k}$ be a smooth vector field, and define the differential 1-form $\Phi = F_1dx + F_2dy + F_3dz$. Calculate $d\Phi$ and identify its associated component functions G_1, G_2 and G_3 as coefficients of $dy \wedge dz, dz \wedge dx$ and $dx \wedge dy$.

This is the calculation for Special Case (2) of Subsection 5.2.1. Recall that, for each $i = 1, 2, 3$, we have

$$dF_i = \frac{\partial F_i}{\partial x}dx + \frac{\partial F_i}{\partial y}dy + \frac{\partial F_i}{\partial z}dz,$$

and

$$d\Phi = dF_1 \wedge dx + dF_2 \wedge dy + dF_3 \wedge dz.$$

Using linearity of the wedge product, the above becomes

$$\begin{aligned} d\Phi &= \cancel{\frac{\partial F_1}{\partial x}dx \wedge dx}^0 + \frac{\partial F_1}{\partial y}dy \wedge dx + \frac{\partial F_1}{\partial z}dz \wedge dx \\ &\quad + \frac{\partial F_2}{\partial x}dx \wedge dy + \cancel{\frac{\partial F_2}{\partial y}dy \wedge dy}^0 + \frac{\partial F_2}{\partial z}dz \wedge dy \\ &\quad + \frac{\partial F_3}{\partial x}dx \wedge dz + \frac{\partial F_3}{\partial y}dy \wedge dz + \cancel{\frac{\partial F_3}{\partial z}dz \wedge dz}^0. \end{aligned}$$

We then re-write the above to match the basis $dy \wedge dz, dz \wedge dx$ and $dx \wedge dy$, keeping in mind that $dz \wedge dy = -dy \wedge dz$ and similarly for other variables:

$$d\Phi = \underbrace{\left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}\right)}_{G_1} dy \wedge dz + \underbrace{\left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}\right)}_{G_2} dz \wedge dx + \underbrace{\left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)}_{G_3} dx \wedge dy.$$

But the coefficients G_1, G_2, G_3 are the same as the components of $\text{curl } \mathbf{F}$! Letting $\mathbf{G} = G_1\mathbf{i} + G_2\mathbf{j} + G_3\mathbf{k} = \text{curl } \mathbf{F}$, this identifies $\mathbf{G}$ (using our second kind of matching) with the exterior derivative $d\Phi$.

3. Suppose that

$$\Theta = H_1dy \wedge dz + H_2dz \wedge dx + H_3dx \wedge dy,$$

is a differential 2-form with smooth component functions H_i . Calculate the differential 3-form $d\Theta$. We use this calculation in Special Case (3) of Subsection 5.2.1.

We know that

$$d\Theta := dH_1 \wedge dy \wedge dz + dH_2 \wedge dz \wedge dx + dH_3 \wedge dx \wedge dy.$$

Now, we have $dH_i := \frac{\partial H_i}{\partial x}dx + \frac{\partial H_i}{\partial y}dy + \frac{\partial H_i}{\partial z}dz$, with terms involving each of the elementary forms dx , dy and dz . But since $dx \wedge dx = dy \wedge dy = dz \wedge dz = 0$, the only nonzero terms in $d\Theta$ are

$$d\Theta := \frac{\partial H_1}{\partial x}dx \wedge dy \wedge dz + \frac{\partial H_2}{\partial y}dy \wedge dz \wedge dx + \frac{\partial H_3}{\partial z}dz \wedge dx \wedge dy.$$

To simplify this, we notice that

$$dy \wedge dz \wedge dx = -dy \wedge dx \wedge dz = +dx \wedge dy \wedge dz,$$

and similarly

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

Therefore

$$d\Theta = \left(\frac{\partial H_1}{\partial x} + \frac{\partial H_2}{\partial y} + \frac{\partial H_3}{\partial z} \right) dx \wedge dy \wedge dz.$$

So, if we define the vector field $\mathbf{H} = \langle H_1, H_2, H_3 \rangle$, then

$$d\Theta = (\operatorname{div} \mathbf{H}) dx \wedge dy \wedge dz.$$

Practice Problems

1. Let $\Psi = x_1^2 dx_1 \wedge dx_3 - x_3 dx_2 \wedge dx_4$ in $\mathbb{R}^4$.

(a) Compute $\Psi \wedge \Phi$, if $\Phi = x_3^5 dx_2 \wedge dx_4$. Give your answer in the form $f(x_1, x_2, x_3, x_4) dx_1 \wedge dx_2 \wedge dx_3 \wedge dx_4$.

(b) Find $d\Psi$. Simplify your answer.

2. Calculate the exterior derivative of the differential forms:

(a) $\Phi = xydy + x^2z^2dz$

(b) $\Phi = (x_1 + x_2^2)dx_1 \wedge dx_2 + x_2 \cos(x_3)dx_2 \wedge dx_4$

3. Calculate the exterior derivative of the differential forms:

(a) $\Phi = 5x_2 dx_1 \wedge dx_3 + (3x_3 - x_1) dx_3 \wedge dx_4$

(b) $\Psi = x_1 x_2^2 dx_1 \wedge dx_3 \wedge dx_4 + x_3^2 x_4 dx_2 \wedge dx_3 \wedge dx_4$.

(c) $\Phi = e^{x_1 - x_3} dx_1 \wedge dx_2 \wedge dx_4 + e^{4x_2 - 5x_3} dx_1 \wedge dx_3 \wedge dx_4$.

4. Find the exterior derivative of the 3-form $x_1 x_2^2 x_3^4 dx_1 \wedge dx_3 \wedge dx_4$ in $\mathbb{R}^4$. Give your answer in the form $g(x_1, x_2, x_3, x_4) dx_1 \wedge dx_2 \wedge dx_3 \wedge dx_4$.
5. Let $\Psi = xydx + z^2dy$ and $\Phi = (2y + z)dx \wedge dy + (xy + z^2)dy \wedge dz$.
- (a) Find $\Psi \wedge \Phi$. Simplify your answer.
- (b) Find $d\Phi$ and $d\Psi$.

5.3 Integrating Differential Forms

Differential forms exist to be integrated on manifolds. In this section, we will talk about how this is done.

5.3.1 Manifolds

For our purposes here, a **smooth parametrized k -manifold** in $\mathbb{R}^n$ will be given by a vector-valued function $\mathbf{x} : U \rightarrow \mathbb{R}^n$, where U is a domain in $\mathbb{R}^k$, $\mathbf{x}$ is C^∞ and one-to-one on U except possibly for the boundary, and its derivative matrix

$$D\mathbf{x}(\mathbf{u}) = \begin{vmatrix} \frac{\partial x_1}{\partial u_1} & \cdots & \frac{\partial x_1}{\partial u_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial u_1} & \cdots & \frac{\partial x_n}{\partial u_k} \end{vmatrix}$$

has rank k (maximal) at every point $\mathbf{u} \in D$. This last condition (maximal rank) means that, if we fix a point $\mathbf{u}_0 \in D$, and calculate the derivative matrix $D\mathbf{x}(\mathbf{u}_0)$ at $\mathbf{u}_0$, then the range of this matrix has dimension k .

This is the set-up that we will use to define integration of differential forms. If $n = 3$ and $k = 1$, then we just have one parameter t and a smooth parametrized curve $\mathbf{r}(t)$ with $\mathbf{r}'(t) \neq 0$. If $n = 3$ and $k = 2$, this is our previous definition of parametrized surfaces with two parameters u, v .

There are many other ways to define a manifold. One is to use a system of equations $f_1(\mathbf{x}) = 0, \dots, f_j(\mathbf{x}) = 0$. Generically (in some neighbourhood of a point satisfying the given equations, with the precise conditions given by the Implicit Function Theorem), each equation reduces the dimension of the manifold by 1. A useful term to describe this is the *codimension*: each equation increases the codimension by 1, so that a manifold described by a nondegenerate system of j equations in $\mathbb{R}^n$ has codimension j and dimension $n - j$.

Another way to define manifolds is more geometrical: essentially, a k -manifold without boundary is a subset of $\mathbb{R}^n$ which “locally” (in a neighbourhood of any fixed point) looks like $\mathbb{R}^k$. A k -manifold with boundary is similar, except at the boundary points

the manifold locally looks like a half-space $\{\mathbf{x} \in \mathbb{R}^k : x_k \geq 0\}$. (The precise meaning of “locally looks like” is provided by terms like *homeomorphism* or *diffeomorphism*. You don’t have to know, in this class, what these are, but you will hear more about it if you take further courses in differential geometry.)

Each definition can be extended to include **piecewise smooth** manifolds. This does not contribute anything conceptually new at this point, so we will skip it here.

5.3.2 Tangent vectors and volume elements

Let $\mathbf{x}(\mathbf{u})$ be a smooth parametrized manifold as above. Then we can find k tangent vectors to the manifold by evaluating the derivatives:

$$\mathbf{T}_i = \left\langle \frac{\partial x_1}{\partial u_i}, \dots, \frac{\partial x_n}{\partial u_i} \right\rangle.$$

(For surfaces in $\mathbb{R}^3$, these are the vectors $\mathbf{r}_u$ and $\mathbf{r}_v$, which we considered regularly throughout this course.) The vectors $\mathbf{T}_1, \dots, \mathbf{T}_k$ are the columns of the derivative matrix $D\mathbf{x}$. Under the maximal rank condition, they are linearly independent at each point $P = \mathbf{x}(\mathbf{u})$ and span a k -dimensional affine space called the *tangent space* to the manifold at P .

The **k -volume element** on a manifold is the volume of the k -dimensional parallelogram spanned by the tangent vectors. This is again a generalization of the work we did for surfaces in $\mathbb{R}^3$. We will skip this part here, but you can check any reasonable textbook on the subject to learn more.

5.3.3 Integrating Differential Forms

Recall that a differential k -form acts on k -tuples of vectors by setting up a $k \times k$ determinant. So, there is a natural way to set up integration of k -forms on k -manifolds: just have the form act on the k tangent vectors $\mathbf{T}_1, \dots, \mathbf{T}_k$ defined above. Evaluating the determinant, we get a function of $\mathbf{u}$ with values in $\mathbb{R}$. Then integrate that function on U . This is in fact what we will do. We will also look at how this recovers some of the integrals we considered earlier, such as flux and circulation.

Remark 5.3.1. *For now, we will ignore issues related to orientation. This is, in fact, the entire focus of the next section.*

If our parametrized manifold $\mathcal{M}$ is given by a function $\mathbf{x} : U \rightarrow \mathbb{R}^n$, where $U \subset \mathbb{R}^k$, and if Φ is a differential k -form defined on $\mathcal{M}$, we define

$$\int_{\mathcal{M}} \Phi = \int_U \Phi(\mathbf{T}_1, \dots, \mathbf{T}_k) du_1 \dots du_k,$$

where $\mathbf{T}_1, \dots, \mathbf{T}_k$ are the tangent vectors defined above.

Let us see some examples and special cases.

- (1) **Volume integration in $\mathbb{R}^n$.** Let $\mathcal{M} = U$ be an open set in $\mathbb{R}^n$, parametrized by $\mathbf{x}(\mathbf{u}) = \mathbf{u}$. This is an n -manifold in $\mathbb{R}^n$. The tangent vectors are $\mathbf{T}_i = \frac{\partial \mathbf{u}}{\partial u_i} = \mathbf{e}_i$.

Let $\Phi = f(x_1, \dots, x_n) dx_1 \wedge \dots \wedge dx_n$ be a differential n -form. Then

$$\Phi(\mathbf{T}_1, \dots, \mathbf{T}_n) = f(x_1, \dots, x_n) \begin{vmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{vmatrix} = f(u_1, \dots, u_n)$$

so that

$$\int_{\mathcal{M}} \Phi = \int_U f(u_1, \dots, u_n) du_1 \dots du_n$$

is the usual n -dimensional integral of functions.

- (2) **Line integrals.** Let $\mathcal{M}$ be a smooth curve in $\mathbb{R}^n$, parametrized by $\mathbf{r}(t)$ with $t \in [a, b]$. There is only one tangent vector $\mathbf{r}'(t)$.

Let $\Phi = F_1 dx_1 + \dots + F_n dx_n$ be a 1-form. Then

$$\begin{aligned} \int_{\mathcal{M}} \Phi &= \int_a^b [F_1 dx_1(\mathbf{r}'(t)) + \dots + F_n dx_n(\mathbf{r}'(t))] dt \\ &= \int_a^b [F_1 x'_1(t) + \dots + F_n x'_n(t)] dt \\ &= \int_a^b F_1 dx_1 + \dots + F_n dx_n, \end{aligned}$$

generalizing vector line integrals to n dimensions.

- (3) **Flux integrals.** Let $\mathcal{S}$ be a 2-dimensional surface in $\mathbb{R}^3$, parametrized by $\mathbf{r}(u, v)$ with $(u, v) \in D$. The tangent vectors are $\mathbf{r}_u$ and $\mathbf{r}_v$.

For a vector field $\mathbf{F}$ in $\mathbb{R}^3$, define the 2-form $\Phi = F_1 dy \wedge dz + F_2 dz \wedge dx + F_3 dx \wedge dy$. Then

$$\begin{aligned} \int_{\mathcal{S}} \Phi &= \int_{\mathcal{S}} F_1 dy \wedge dz + F_2 dz \wedge dx + F_3 dx \wedge dy \\ &= \iint_D F_1(\mathbf{r}(u, v)) dy \wedge dz(\mathbf{r}_u, \mathbf{r}_v) + \dots \end{aligned}$$

Evaluating the 2-form, we get

$$\int_{\mathcal{S}} \Phi = \iint_D \left[F_1(\mathbf{r}(u, v)) \begin{vmatrix} y_u & z_u \\ y_v & z_v \end{vmatrix} + \dots \right] dudv = \iint_D \mathbf{F} \cdot \mathbf{n} dudv,$$

where

$$\mathbf{n} = \mathbf{r}_u \times \mathbf{r}_v = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_u & y_u & z_u \\ x_v & y_v & z_v \end{vmatrix}$$

(Exercise: fill in the blanks and check that the calculation is correct!) This matches our formula for evaluating $\iint_S \mathbf{F} \cdot d\mathbf{S}$, if $\mathbf{F} = F_1\mathbf{i} + F_2\mathbf{j} + F_3\mathbf{k}$.

- (4) **Higher-dimensional flux integrals.** Let's try to extend the concept of flux integrals to n dimensions. Let $\mathcal{M}$ be an $(n-1)$ -dimensional surface in $\mathbb{R}^n$, parametrized by $\mathbf{x}(\mathbf{u})$ with $\mathbf{u} \in U$, where U is a domain in $\mathbb{R}^{n-1}$. The tangent vectors are $\mathbf{T}_i = \frac{\partial \mathbf{x}}{\partial u_i}$. The vector

$$\mathbf{n} = \begin{vmatrix} \mathbf{e}_1 & \cdots & \mathbf{e}_n \\ \frac{\partial x_1}{\partial u_1} & \cdots & \frac{\partial x_n}{\partial u_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_1}{\partial u_{n-1}} & \cdots & \frac{\partial x_n}{\partial u_{n-1}} \end{vmatrix}$$

is normal to $\mathcal{M}$. (Exercise: check that it is orthogonal to each $\mathbf{T}_i$.) Then we can set up integrals of $(n-1)$ -forms Φ on $\mathcal{M}$ so that

$$\int_{\mathcal{M}} \Phi = \int_U \mathbf{F} \cdot \mathbf{n} \, d\mathbf{u},$$

where $\mathbf{F}$ is the vector field with the components $F_1, \dots, F_n$ equal to the corresponding components of Φ .

Worked Problems

In the hands-on examples below, we ignore orientation issues for now. If you use a different ordering of the tangent vectors in your solution, your answer may differ by a $\pm$ sign. (Exercise: complete the calculation of the integrals!)

1. Integrate $\Phi = x_1x_3dx_1 \wedge dx_2 + x_3dx_1 \wedge dx_4$ on the 2-manifold $\mathbf{x} = \langle t, 3s+t, s, s-t \rangle$, $0 \leq t \leq 1$, $0 \leq s \leq 2$, in $\mathbb{R}^4$.

The tangent vectors are $\mathbf{x}_t = \langle 1, 1, 0, -1 \rangle$ and $\mathbf{x}_s = \langle 0, 3, 1, 1 \rangle$. Let us evaluate the 2-forms first:

$$x_1x_3dx_1 \wedge dx_2 = st \begin{vmatrix} 1 & 1 \\ 0 & 3 \end{vmatrix} = 3st$$

$$x_3dx_1 \wedge dx_4 = s \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = s.$$

The integral is

$$\int_0^1 \int_0^2 (3st + s) ds dt = 5$$

If you re-ordered your vectors $\mathbf{x}_t$ and $\mathbf{x}_s$ when evaluating the wedge product, you would have obtained an integral of -5 , rather than $+5$.

2. Integrate $\Psi = x_2 x_3 dx_1 \wedge dx_2 \wedge dx_4$ on the part of the paraboloid $\mathbf{x} = \langle s, t, u, s^2 + t^2 + u^2 \rangle$ given by $0 \leq s \leq 1$, $0 \leq t \leq 2$, $0 \leq u \leq 3$ in $\mathbb{R}^4$.

The tangent vectors are

$$\mathbf{x}_s = \langle 1, 0, 0, 2s \rangle$$

$$\mathbf{x}_t = \langle 0, 1, 0, 2t \rangle$$

$$\mathbf{x}_u = \langle 0, 0, 1, 2u \rangle$$

so that

$$x_2 x_3 dx_1 \wedge dx_2 \wedge dx_4(\mathbf{x}_s, \mathbf{x}_t, \mathbf{x}_u) = tu \begin{vmatrix} 1 & 0 & 2s \\ 0 & 1 & 2t \\ 0 & 0 & 2u \end{vmatrix} = tu \begin{vmatrix} 1 & 2t \\ 0 & 2u \end{vmatrix} = 2tu^2.$$

Hence, we evaluate the definite integral

$$\int_0^3 \int_0^2 \int_0^1 2u^2 t ds dt du = 36.$$

Again, note: if you re-ordered the vectors $\mathbf{x}_s, \mathbf{x}_t$ and $\mathbf{x}_u$ in your determinant, you may have obtained an answer of -36 instead.

Practice Problems

For now, ignore orientation issues. Your answer may differ by a $\pm$ sign, depending on how you order the tangent vectors.

1. Compute $\int_M \Psi$, if M is the parametrized manifold

$$\mathbf{r}(u, v) = (u^3 - v, v^3, uv^2, u^2v), \quad 0 \leq u \leq 3, \quad 0 \leq v \leq 2,$$

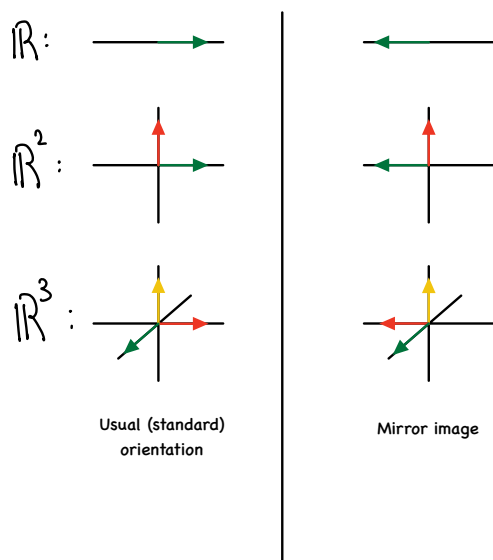
and $\Psi = dx_1 \wedge dx_2 + 2dx_3 \wedge dx_4$.

2. Integrate $\Phi = 4dx_1 \wedge dx_2 - x_2^2 dx_3 \wedge dx_4$ on the 2-manifold $\mathbf{x} = \langle t, s, t \cos s, t \sin s \rangle$, $0 \leq t \leq 1$, $0 \leq s \leq 2\pi$, in $\mathbb{R}^4$.

5.4 Orientation on Manifolds

5.4.1 Defining orientation on manifolds

In dimensions $n = 1, 2, 3$, we have always worked with the **oriented** space $\mathbb{R}^n$ in the sense that we assumed a certain ordering of the basis vectors (for example, the vectors $\mathbf{i}$ and $\mathbf{j}$ in $\mathbb{R}^2$ are arranged in the counterclockwise order). Switching to a different coordinate system preserves orientation if the new basis vectors are arranged the same way. You can think about this in terms of manipulating the basis vectors. If you want to preserve orientation, you can stretch the basis vectors or rotate them continuously, but with restrictions: you are not allowed to shrink any basis vector to zero length, and you are not allowed to rotate any one of the basis vectors so that it crosses the subspace spanned by the remaining vectors.



A more precise way to define this, and one that works in higher dimensions as well, is to look at the sign of the determinant of a linear mapping. Let's say that $\mathbf{u}_1, \dots, \mathbf{u}_n$ and $\mathbf{v}_1, \dots, \mathbf{v}_n$ are the basis vectors in two coordinate systems. Define a linear mapping $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$ so that $L(\mathbf{u}_j) = \mathbf{v}_j$ for $j = 1, \dots, n$. We will then say that the two coordinate systems define the same orientation if the determinant of L is positive, and that they define opposite orientations if the determinant is negative.

If the linear mapping L is just a permutation of the basis vectors, then it preserves orientation if the permutation involves an even number of transpositions of pairs of vectors, and reverses it if it involves an odd number of transpositions. (Permutations can be written as superpositions of transpositions of pairs in many different ways. A result in combinatorics says that for a given permutation, the number of transpositions is either always odd or always even. This is called the **sign** of the permutation: a

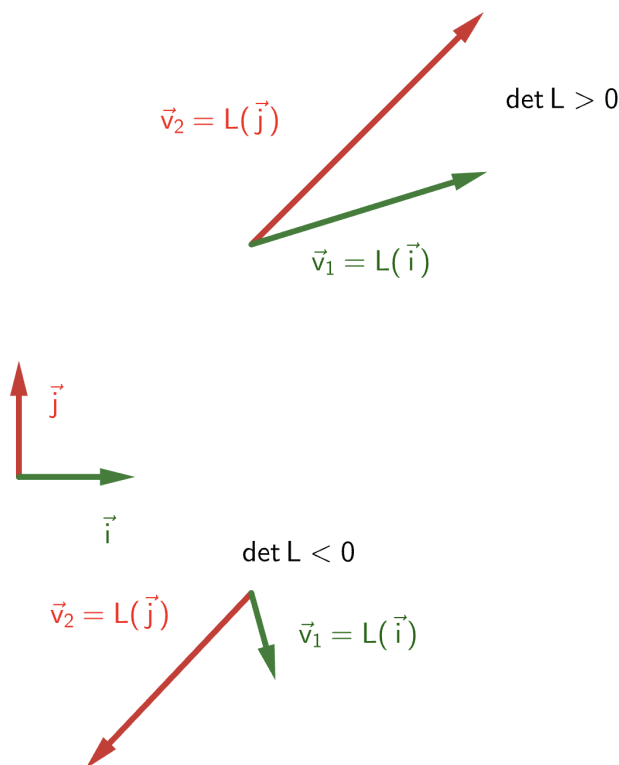


Figure 5.2: The original basis vectors $\mathbf{i}$ and $\mathbf{j}$, as well as two linear transformations applied to them: one with positive determinant, the other, negative.

permutation is odd if the number of transpositions involved in it is always odd, and even if the number of transpositions is even.)

Note that the differential form $dx_1 \wedge \cdots \wedge dx_n$ captures the same information. We have

$$dx_1 \wedge \cdots \wedge dx_n(\mathbf{e}_1, \dots, \mathbf{e}_n) = 1.$$

If we interchange two of the basis vectors, this reverses the orientation; it also changes the differential form evaluation to -1 (because two of the rows in the determinant get interchanged). If we interchange two pairs, we return to the original orientation, and the form now evaluates to $+1$; and so on.

That's how we get to differential forms defining orientation on manifolds. On a k -manifold, a small neighbourhood of any non-boundary point looks locally like $\mathbb{R}^k$, so we can define orientation *on that neighbourhood* like we would in $\mathbb{R}^k$ – by choosing a basis of tangent vectors in the tangent space at that point, and then choosing an ordering of them. However, this works only locally, and we want to orient the entire manifold so that the orientations on different small neighbourhoods match continuously. A k -differential form with possibly non-constant coefficients allows us to provide such

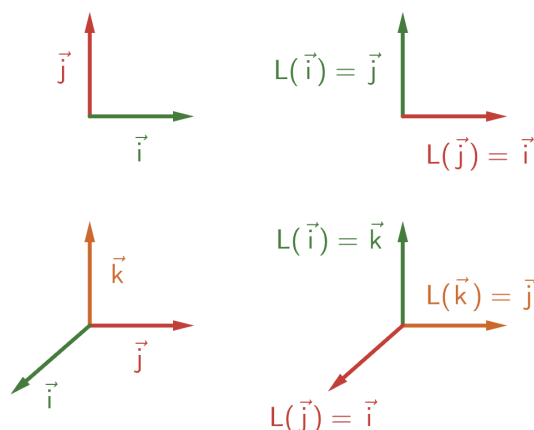


Figure 5.3: Two different linear transformations. The first linear transformation (2-dimensional) does *not* preserve orientation, since one transposition occurs. However, the second transformation (3-dimensional) does preserve orientation, since two transformations (an even number) of transpositions occurred.

consistency on manifolds which are not flat.

Let ω be a smooth differential k -form, defined everywhere on a k -manifold $\mathcal{M}$ in $\mathbb{R}^n$. Assume that $\mathcal{M}$ has a smooth parametrization $\mathbf{x} : U \rightarrow \mathbb{R}^n$, where $U \subset \mathbb{R}^k$. Fix some $\mathbf{x} = \mathbf{x}(\mathbf{u}) \in \mathcal{M}$. We are interested in how ω acts on vectors tangent to $\mathcal{M}$ at that $\mathbf{x}$.

Let $\mathbf{T}_1, \dots, \mathbf{T}_k$ be the tangent vectors to $\mathcal{M}$ at $\mathbf{x}$ that we defined above. From our definition of a smooth manifold, we know that they are linearly independent. Let $\mathcal{T}_{\mathbf{x}}$ be the linear space spanned by these vectors. (This is the k -dimensional space of all directions tangent to $\mathcal{M}$ at the given point.)

We now want to zoom in on that space. First, we evaluate all coefficient functions of ω at $\mathbf{x}$. This gives us a k -form $\omega_{\mathbf{x}}$ with constant coefficients, still defined on $\mathbb{R}^n$. Then, we restrict $\omega_{\mathbf{x}}$ to $\mathcal{T}_{\mathbf{x}}$. This is still linear in each variable and antisymmetric, hence a k -form with constant coefficients on a k -dimensional space.

We can define a new coordinate system $\mathbf{y} = (y_1, \dots, y_k)$ on $\mathcal{T}_{\mathbf{x}}$ (for example, with $\mathbf{T}_1, \dots, \mathbf{T}_k$ as the basis vectors). In these coordinates, $\Omega_{\mathbf{x}}$ has the form

$$\Omega_{\mathbf{x}} = a_{\mathbf{x}} dy_1 \wedge \dots \wedge dy_k \quad (5.4.1)$$

for some constant $a \in \mathbb{R}$. We will say that $\Omega_{\mathbf{x}}$ is **non-degenerate** if $a_{\mathbf{x}} \neq 0$; since $\mathbf{T}_1, \dots, \mathbf{T}_k$ are linearly independent, this is true if and only if

$$\Omega_{\mathbf{x}}(\mathbf{T}_1, \dots, \mathbf{T}_k) \neq 0$$

(see Example 5.1.4). In that case, the sign of $\Omega_{\mathbf{x}}(\mathbf{T}_1, \dots, \mathbf{T}_k)$ captures the ordering of the tangent vectors $\mathbf{T}_1, \dots, \mathbf{T}_k$ at $\mathbf{x}$.

To do this consistently throughout the entire manifold, we need a “global” version, as follows.

Definition 5.4.1. Let ω be a smooth differential k -form on a k -manifold $\mathcal{M}$ in $\mathbb{R}^n$. We say that ω **orients** $\mathcal{M}$ if the restricted k -form $\Omega_{\mathbf{x}}$, defined above, is non-degenerate for each $\mathbf{x} \in \mathcal{M}$. We say that $\mathcal{M}$ is **orientable** if such a form ω exists.

In this global picture, we go back from $\Omega_{\mathbf{x}}$ to ω as follows. First, since $\Omega_{\mathbf{x}}$ was just the restriction of $\omega_{\mathbf{x}}$ to tangent directions, we have $\Omega_{\mathbf{x}}(\mathbf{T}_1, \dots, \mathbf{T}_k) = \omega_{\mathbf{x}}(\mathbf{T}_1, \dots, \mathbf{T}_k)$. Second, we now allow $\mathbf{x}$ to vary, so that both the coefficients of ω and the tangent vectors $\mathbf{T}_j$ are variable as we move around the manifold. Thus, ω orients $\mathcal{M}$ if and only if

$$\omega(\mathbf{T}_1, \dots, \mathbf{T}_k) \neq 0 \text{ for all } \mathbf{u} \in U,$$

and the given parametrization is **consistent** (or compatible) with the orientation given by ω if the tangent vectors $\mathbf{T}_1, \dots, \mathbf{T}_k$ satisfy

$$\omega(\mathbf{T}_1, \dots, \mathbf{T}_k) > 0 \text{ for all } \mathbf{u} \in U. \quad (5.4.2)$$

(If they don't, we need to reorder the variables so that the above condition is satisfied.)

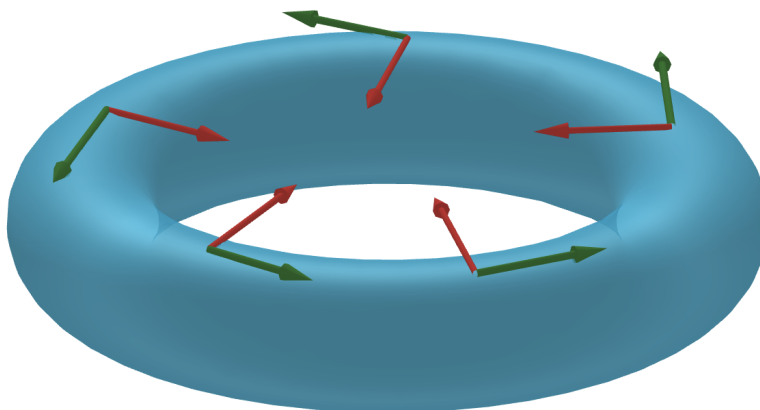


Figure 5.4: The torus in $\mathbb{R}^3$ (pictured above) is an example of a “non-flat” manifold where we can define orientation and keep it consistent throughout the manifold.

Example 5.4.2. Let $\mathcal{M} = \{(x, y, z, g(x, y, z)) : (x, y, z) \in U\}$ be the graph of a function of 3 variables in $\mathbb{R}^4$. With $\mathcal{M}$ parametrized by the (x, y, z) variables, the tangent vectors are

$$\mathbf{T}_x = (1, 0, 0, g_x), \quad \mathbf{T}_y = (0, 1, 0, g_y), \quad \mathbf{T}_z = (0, 0, 1, g_z).$$

We claim that the 3-form $\omega = dx \wedge dy \wedge dz$ orients $\mathcal{M}$. Indeed,

$$\omega(\mathbf{T}_x, \mathbf{T}_y, \mathbf{T}_z) = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1,$$

which is never zero. The same calculation shows that the parametrization above is consistent with the orientation given by ω . (Exercise: think about how this would extend to functions of n variables!)

Example 5.4.3. To orient manifolds that are not graphs of functions, we usually need forms ω with variable coefficients. For instance, let $\mathbf{r}(t, z) = \cos t \mathbf{i} + \sin t \mathbf{j} + z \mathbf{k}$ be the parametrization of a cylinder in $\mathbb{R}^3$. The two tangent vectors are

$$\mathbf{r}_t = -\sin t \mathbf{i} + \cos t \mathbf{j}, \quad \mathbf{r}_z = \mathbf{k}.$$

If we try to use the elementary 2-forms for orientation, we get

$$dx \wedge dy(\mathbf{r}_t, \mathbf{r}_z) = 0, \quad dx \wedge dz(\mathbf{r}_t, \mathbf{r}_z) = -\sin t, \quad dy \wedge dz(\mathbf{r}_t, \mathbf{r}_z) = \cos t,$$

so that none of them work. However, if $\omega = xdy \wedge dz - ydx \wedge dz$, then

$$\omega(\mathbf{r}_t, \mathbf{r}_z) = \cos^2 t + \sin^2 t = 1,$$

so that this form does provide a valid orientation. A geometric interpretation is as follows. Let $\mathbf{N} = \cos t \mathbf{i} + \sin t \mathbf{j}$ be the normal vector to the cylinder. Then the form ω above can also be written in the determinant form

$$\omega(\mathbf{v}_1, \mathbf{v}_2) = \begin{vmatrix} \cos t & v_{11} & v_{21} \\ \sin t & v_{12} & v_{22} \\ 0 & v_{13} & v_{23} \end{vmatrix},$$

with columns $\mathbf{N}, \mathbf{v}_1, \mathbf{v}_2$. In particular,

$$\omega(\mathbf{r}_t, \mathbf{r}_z) = \begin{vmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1,$$

corresponding to the fact that $\mathbf{N}, \mathbf{r}_t, \mathbf{r}_z$ are three orthonormal vectors oriented the same way as the $\mathbf{i}, \mathbf{j}, \mathbf{k}$ triple.

Some important special cases we should keep in mind are the following.

- **Orientation on $\mathbb{R}^n$:** We will always use the n -form $dx_1 \wedge \cdots \wedge dx_n$ to orient solid n -dimensional domains in $\mathbb{R}^n$ including the entire space. In $\mathbb{R}^2$ and $\mathbb{R}^3$, this means $dx \wedge dy$ and $dx \wedge dy \wedge dz$, respectively. This corresponds to the standard (counterclockwise for $n = 2, 3$) orientation we have been using all along.
- **Orientation on curves:** if $\mathbf{r}(t)$ is a smooth parametrized curve, then $\mathbf{r}'(t)$ provides orientation. In the language of differential forms, the 1-form $\omega_{\mathbf{r}}(\mathbf{v}) = \mathbf{r}' \cdot \mathbf{v}$ orients the curve. Recall that, for a smooth parametrized curve, we assumed that $\mathbf{r}'(t) \neq \mathbf{0}$ for all t , so that $\mathbf{r}' \cdot \mathbf{v} \neq 0$ for any vector $\mathbf{v}$ tangent to the curve at $\mathbf{r}(t)$.

- **Orientation on surfaces in $\mathbb{R}^3$:** this is a generalization of Example 5.4.3. If $\mathbf{n}(\mathbf{x})$ is the normal vector to $\mathcal{S}$ at $\mathbf{x}$, let

$$\omega_{\mathbf{x}}(\mathbf{v}_1, \mathbf{v}_2) = |\mathbf{n}(\mathbf{x}) \quad \mathbf{v}_1 \quad \mathbf{v}_2| \quad (5.4.3)$$

(the determinant of the matrix with columns $\mathbf{n}(\mathbf{x})$, $\mathbf{v}_1$, $\mathbf{v}_2$). If $\mathcal{S}$ is orientable, then $\mathbf{n}(\mathbf{x})$ can be chosen consistently on $\mathcal{S}$ so that it is never zero—and therefore (5.4.3) defines a nonvanishing differential form. The determinant in (5.4.3) is then positive if and only if the vectors $\mathbf{n}(\mathbf{x})$, $\mathbf{v}_1$, $\mathbf{v}_2$ are oriented the same way as the vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$ in $\mathbb{R}^3$, meaning that $\mathbf{v}_1, \mathbf{v}_2$ are oriented counterclockwise when viewed from the direction of $\mathbf{n}(\mathbf{x})$. A parametrization $\mathbf{r}(t, s)$ is consistent with this orientation if the same is true of the $\mathbf{T}_t, \mathbf{T}_s$ vectors (i.e. if they are oriented counterclockwise when viewed from the direction of $\mathbf{n}(\mathbf{x})$).

- **Orientation on hypersurfaces of dimension $n - 1$ in $\mathbb{R}^n$:** the last definition generalizes to all $n \geq 2$. Suppose that $\mathcal{M}$ is a manifold of dimension $n - 1$ in $\mathbb{R}^n$, and that we can choose a non-zero normal vector $\mathbf{n}$ continuously throughout $\mathcal{M}$. Then orientation on $\mathcal{M}$ can be defined using the form

$$\omega_{\mathbf{x}}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{n-1}) = |\mathbf{n}(\mathbf{x}) \quad \mathbf{v}_1 \quad \mathbf{v}_2 \quad \dots \quad \mathbf{v}_{n-1}| \quad (5.4.4)$$

In many situations, there is a natural choice of $\mathbf{n}$. For example, if $\mathcal{M}$ is given by an equation $g(x_1, \dots, x_n) = 0$ with $\nabla g \neq 0$ on $\mathcal{M}$, then we can choose $\mathbf{n} = \nabla g$. If $\mathcal{M}$ is the boundary of a solid region $D \subset \mathbb{R}^n$, we may choose $\mathbf{n}$ to point away from D . As in $\mathbb{R}^3$, the determinant in (5.4.4) is positive if and only if the ordering of $\mathbf{n}, \mathbf{v}_1, \dots, \mathbf{v}_{n-1}$ matches the standard orientation in $\mathbb{R}^n$.

5.4.2 Inherited orientation on boundary

Suppose that $\mathcal{M}$ is a manifold of dimension k in $\mathbb{R}^n$, oriented by a k -form ω , and that its boundary $\partial\mathcal{M}$ is a $(k - 1)$ -manifold. Then the *inherited orientation* on $\partial\mathcal{M}$ is given by the form

$$\partial\omega_{\mathbf{x}}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{k-1}) := \omega_{\mathbf{x}}(\mathbf{n}(\mathbf{x}), \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{k-1}), \quad (5.4.5)$$

where $\mathbf{n}(\mathbf{x})$ is the vector “tangent to $\mathcal{M}$ ” (more precisely, it would be tangent to $\mathcal{M}$ if $\mathcal{M}$ were extended past its boundary), normal to $\partial\mathcal{M}$, and pointing away from $\mathcal{M}$.

- Suppose that $\mathcal{M} = D$ is a solid region in $\mathbb{R}^n$. We use our standard orientation in $\mathbb{R}^n$, given by the form $\omega = dx_1 \wedge \dots \wedge dx_n$. Then the formula (5.4.5) defining $\partial\omega$ on $\partial\mathcal{M}$ matches (5.4.4) if we let $\mathbf{n}(\mathbf{x})$ point away from D . This is consistent with our earlier convention for surfaces bounding solid regions in $\mathbb{R}^3$.

- Let $\mathcal{M} = \mathcal{S}$ be a 2-dimensional surface in $\mathbb{R}^3$, bounded by a curve $\mathcal{C}$. Assume that $\mathcal{S}$ is orientable according to our definition in Chapter 3, so that it has a unit normal vector $\mathbf{N}$ varying smoothly throughout the manifold. We then have the 2-form

$$\omega_{\mathbf{x}}(\mathbf{v}_1, \mathbf{v}_2) = |\mathbf{N} \cdot \mathbf{v}_1 \times \mathbf{v}_2| \quad (5.4.6)$$

as in (5.4.3), providing counterclockwise orientation on $\mathcal{S}$ when viewed from the direction of $\mathbf{N}$. We then imagine $\mathcal{S}$ extending beyond its boundary, and orient the pair $(\mathbf{n}, \mathbf{T})$ consistently with this orientation, where $\mathbf{n}$ is as described above and $\mathbf{T}$ indicates the orientation of the boundary curve. You can check that this matches our earlier convention.

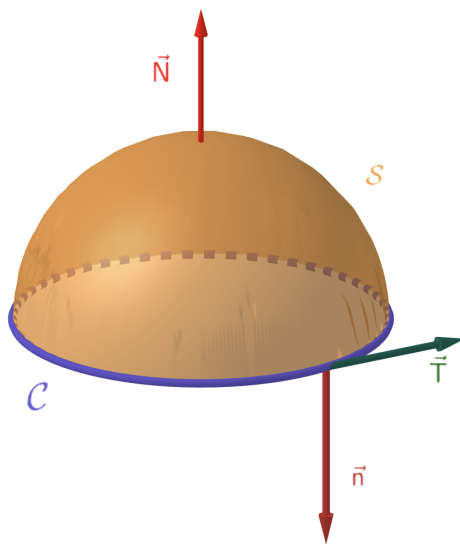


Figure 5.5: Inherited orientation on the boundary of the surface.

5.4.3 Integration on oriented manifolds

Our integration formula from Section 5.3 was as follows: if $\mathcal{M}$ is parametrized by a function $\mathbf{x} : U \rightarrow \mathbb{R}^n$, where $U \subset \mathbb{R}^k$, and if Φ is a differential k -form defined on $\mathcal{M}$, then

$$\int_{\mathcal{M}} \Phi = \int_U \Phi(\mathbf{T}_1, \dots, \mathbf{T}_k) du_1 \dots du_k,$$

where $\mathbf{T}_1, \dots, \mathbf{T}_k$ are the tangent vectors defined above.

But the thing is, how do we know the ordering of the $\mathbf{T}_1, \dots, \mathbf{T}_k$ vectors in Φ ? For example, if a 2-manifold is parametrized by (s, t) , we could use (s, t) in that order, or

we could switch them and use (t, s) instead. Such a change in parametrization would affect the order of two of the tangent vectors, therefore the sign of $\Phi(\mathbf{T}_1, \dots, \mathbf{T}_k)$ and, ultimately, the sign of the integral.

When we introduce orientation on the manifold, *we fix the order in which we should write the tangent vectors*. Specifically, if ω is the form providing orientation on $\mathcal{M}$, we want to label our parameters $u_1, \dots, u_k$ (therefore the tangent vectors $\mathbf{T}_1, \dots, \mathbf{T}_k$) so that $\omega(\mathbf{T}_1, \dots, \mathbf{T}_k) > 0$. If this holds, then the parametrization is consistent with the orientation given by ω . This is the order of tangent vectors we need to use in our integration formula. We illustrate this in the Worked Problems section.

Remark 5.4.1. *We assumed earlier that the tangent vectors $\mathbf{T}_1, \dots, \mathbf{T}_k$ are linearly independent everywhere on $\mathcal{M}$ except possibly on its boundary. We also assumed that ω is nonzero on $\mathcal{M}$. Therefore $\omega(\mathbf{T}_1, \dots, \mathbf{T}_k)$ cannot change sign on $\mathcal{M}$, so that it must be either always positive or always negative on $\mathcal{M}$ except possibly for boundary points. We will see this in the examples below.*

Worked Problems

In the Worked Problems in Section 5.3, our solutions ignored orientation issues. We now revisit the same questions with attention paid to the orientation of the underlying manifolds.

1. Check that the 2-manifold $\mathbf{x} = \langle t, 3s + t, s, s - t \rangle$, $0 \leq t \leq 1$, $0 \leq s \leq 2$, in $\mathbb{R}^4$, is oriented by the 2-form $\omega = dx_1 \wedge dx_2$. Integrate $\Phi = x_1 x_3 dx_1 \wedge dx_2 + x_3 dx_1 \wedge dx_4$ on this manifold, with this choice of orientation.

The tangent vectors are $\mathbf{x}_t = \langle 1, 1, 0, -1 \rangle$ and $\mathbf{x}_s = \langle 0, 3, 1, 1 \rangle$. Since

$$dx_1 \wedge dx_2(\mathbf{x}_t, \mathbf{x}_s) = \begin{vmatrix} 1 & 1 \\ 0 & 3 \end{vmatrix} = 3$$

is never 0, the form orients the manifold; moreover, since it evaluates to a positive value on $(\mathbf{x}_t, \mathbf{x}_s)$ taken in this order, this is the order of the tangent vectors we want to use.

The rest of the solution is as before. We evaluate the 2-forms:

$$x_1 x_3 dx_1 \wedge dx_2(\mathbf{x}_t, \mathbf{x}_s) = ts \begin{vmatrix} 1 & 1 \\ 0 & 3 \end{vmatrix} = 3ts,$$

$$x_3 dx_1 \wedge dx_4(\mathbf{x}_t, \mathbf{x}_s) = s \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = s,$$

so that the integral is

$$\int_0^1 \int_0^2 (3ts + s) ds dt = 5.$$

(Additional note: if we wanted to orient $\mathcal{M}$ by $-\omega$ instead, we would have to reverse the order of the t and s variables.)

2. Integrate $\Psi = x_2x_3dx_1 \wedge dx_2 \wedge dx_4$ on the part of the paraboloid $\mathbf{x} = \langle s, t, u, s^2+t^2+u^2 \rangle$ given by $0 \leq s \leq 1$, $0 \leq t \leq 2$, $0 \leq u \leq 3$ in $\mathbb{R}^4$, with orientation given by the form $\omega = dx_1 \wedge dx_2 \wedge dx_3$.

The tangent vectors are $\mathbf{x}_s = \langle 1, 0, 0, 2s \rangle$, $\mathbf{x}_t = \langle 0, 1, 0, 2t \rangle$, and $\mathbf{x}_u = \langle 0, 0, 1, 2u \rangle$. Then

$$dx_1 \wedge dx_2 \wedge dx_3(\mathbf{x}_s, \mathbf{x}_t, \mathbf{x}_u) = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1 > 0,$$

so that the above order of the s, t, u variables works for us. Now

$$x_2x_3dx_1 \wedge dx_2 \wedge dx_4(\mathbf{x}_s, \mathbf{x}_t, \mathbf{x}_u) = tu \begin{vmatrix} 1 & 0 & 2s \\ 0 & 1 & 2t \\ 0 & 0 & 2u \end{vmatrix} = tu(2u) = 2tu^2$$

so that the integral is

$$\int_0^1 \int_0^2 \int_0^3 2tu^2 du dt ds = 36.$$

Calculation omitted here (but not in your own work)!

(Additional note: the above calculation shows that the form $\omega' = dx_1 \wedge dx_2 \wedge dx_4$ could also be used to orient the manifold, with $\omega'(\mathbf{x}_s, \mathbf{x}_t, \mathbf{x}_u) = 2u > 0$ except for the part of boundary where $u = 0$. This would still work for the purposes of integration here, but there are times (as in the Generalized Stokes's Theorem, to come next time) when we want a form that does not vanish on the boundary.)

Practice Problems

1. Compute $\int_M \Psi$, if $\Psi = x_1^2 dx_1 \wedge dx_3 - x_3 dx_2 \wedge dx_4$ and M is the parametrized manifold $\mathbf{r}(u, v) = (u, v, uv, u^2 - v^2)$, $0 \leq u \leq 1$, $0 \leq v \leq 2$, oriented consistently with this parametrization.
2. Compute $\int_M \Psi$, if M is the parametrized manifold $\mathbf{r}(u, v) = (u^2, v^2, u + v^3, u^3 - v)$, $0 \leq u \leq 1$, $0 \leq v \leq 1$, oriented consistently with this parametrization, and $\Psi = 3x_1 dx_2 \wedge dx_4 + x_2 dx_3 \wedge dx_4$.
3. Compute $\int_M \Psi$, if $\Psi = 2x_4 dx_1 \wedge dx_2 \wedge dx_4$ and M is the parametrized manifold $\mathbf{x}(u, v, t) = (u \cos t, u \sin t, u + v, v)$, $0 \leq u \leq 1$, $0 \leq v \leq 2$, $0 \leq t \leq 2\pi$, oriented consistently with this parametrization.

4. Let M be the manifold parametrized by

$$\mathbf{r}(u, v) = (u + v, u - v, u \cos v, u \sin v),$$

with $0 \leq u \leq 2$, $0 \leq v \leq \pi$.

(a) Prove that the differential form $\omega = dx_1 \wedge dx_2$ orients M .

(b) Compute $\int_M \Phi$, if $\Phi = x_1 x_2 dx_3 \wedge dx_4$ and M is oriented by the form ω from (a).

5.5 Generalized Stokes's Theorem

Generalized Stokes's Theorem states that, with all the assumptions and notation we have introduced,

$$\int_{\partial \mathcal{M}} \omega = \int_{\mathcal{M}} d\omega, \quad (5.5.1)$$

where ω is a differential form, $d\omega$ is its exterior derivative, $\mathcal{M}$ is an oriented manifold, and $\partial \mathcal{M}$ is the boundary of $\mathcal{M}$ with orientation inherited from M .

We will skip the proof of the theorem; if your Calculus textbook covers differential forms, it will likely include a sketch of it. In general terms, the proof has the same main ingredients as the cases we have seen in 2 and 3 dimensions: break up the manifold (and therefore the integrals) into smaller pieces where one can use a single local coordinate system, write out the iterated integrals, use the Fundamental Theorem of Calculus in one of the variables, and cancel out the integrals along the new boundary pieces that we introduced in the subdivision of the manifold.

Instead, we will revisit the integral theorems we have seen so far and take a look at how they fit into the framework of the Generalized Stokes Theorem.

5.5.1 Stokes's Theorem

We can obtain our classical Stokes's Theorem from Generalized Stokes's Theorem by specializing to the case when $n = 3$, $\mathcal{M} = \mathcal{S}$ is a 2-dimensional surface, and its boundary $\partial \mathcal{M} = \mathcal{C}$ is a closed curve. We also assume that $\mathcal{C}$ is oriented consistently with $\mathcal{M}$. We have already seen that the general definition of “inherited orientation” for manifolds is consistent with how we previously oriented $\mathcal{S}$ and $\mathcal{C}$.

Let $\Phi = F_1 dx + F_2 dy + F_3 dz$. Then

$$\int_{\partial \mathcal{M}} \Phi = \int_{\mathcal{C}} F_1 dx + F_2 dy + F_3 dz = \int_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{r},$$

where $\mathbf{F} = F_1 \mathbf{i} + F_2 \mathbf{j} + F_3 \mathbf{k}$.

In the Worked Problems of Section 5.2, we verified that

$$d\Phi = \Psi = G_1 dy \wedge dz + G_2 dz \wedge dx + G_3 dx \wedge dy,$$

where $\mathbf{G} = G_1\mathbf{i} + G_2\mathbf{j} + G_3\mathbf{k} = \text{curl } \mathbf{F}$. We also checked in Section 5.3 that

$$\int_{\mathcal{M}} \Psi = \int_{\mathcal{S}} \mathbf{G} \cdot d\mathbf{S}.$$

Hence, the Generalized Stokes Theorem tells us that

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_{\mathcal{S}} \mathbf{G} \cdot d\mathbf{S},$$

which is Stokes's Theorem for surfaces. Of course, as discussed in Chapter 4, Green's Theorem is just a special case where $\mathcal{S}$ is a domain in the xy -plane.

Remark 5.5.1. For 2-dimensional surfaces in $\mathbb{R}^n$ with boundary $\mathcal{C}$, equation (5.5.1) in Generalized Stokes's Theorem relates the line integral $\int_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{x}$ (now with n components) to the surface integral of the 2-form $d\Psi$ —a differential form which can be viewed as an n -dimensional generalization of curl.

5.5.2 Divergence Theorem.

Let $\mathcal{M} = D$ be a domain in $\mathbb{R}^3$ bounded by a closed surface $\mathcal{S}$. Apply (5.5.1) to the form

$$\Psi = G_1 dy \wedge dz + G_2 dz \wedge dx + G_3 dx \wedge dy.$$

We get that

$$\int_{\mathcal{S}} \mathbf{G} \cdot d\mathbf{S} = \int_{\partial\mathcal{M}} \Psi = \int_{\mathcal{M}} d\Psi.$$

In the Worked Problems of Section 5.2, we saw that $d\Psi = \text{div } \mathbf{G} \, dx \wedge dy \wedge dz$, so that the above identity recovers the Divergence Theorem. The n -dimensional case can be set up similarly if we apply the Generalized Stokes Theorem to an appropriate differential $(n-1)$ -form with the same components as our vector field.

5.5.3 Line integrals of conservative vector fields.

Let $\mathcal{M}$ be a smooth curve in $\mathbb{R}^n$, parametrized by $\mathbf{x}(t)$ with $t \in [a, b]$. Also, let $\theta = f$ be a 0-form (a smooth function). Then

$$d\theta = df = \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i$$

and

$$\int_{\mathcal{M}} d\theta = \int_C \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i = \int_C \nabla f \cdot d\mathbf{x} = f(\mathbf{x}(b)) - f(\mathbf{x}(a)).$$

This is consistent with the identity (5.5.1) if we think of the endpoints of the curve as its boundary, “oriented” in the sense that the initial point $\mathbf{x}(a)$ comes with a $-$ sign and the terminal point $\mathbf{x}(b)$ comes with a $+$ sign. The “integration” in 0 dimension just means evaluating a function at a point.

5.5.4 One final example

Finally, we give an example in 4-dimensions which fully goes beyond anything covered by the previous three “classical” theorems, showing indeed that the Generalized Stokes Theorem supersedes the previous results (in the sense that it includes them as special cases, but also provides new results in higher dimensions).

Example 5.5.1. In $\mathbb{R}^4$, consider the 3-form

$$\Phi = x_1^3 dx_2 \wedge dx_3 \wedge dx_4 - x_2^3 dx_3 \wedge dx_4 \wedge dx_1 + x_3 dx_4 \wedge dx_1 \wedge dx_2.$$

We use the Generalized Stokes Theorem to evaluate the integral of Φ on the oriented boundary on D , where

$$D = \{x \in \mathbb{R}^4 : x_1^2 + x_2^2 \leq 4, x_3^2 + x_4^2 \leq 9\}.$$

Note that D “looks” something like the intersection of two “hyper-cylinders” of radii $r_1 = 2$ and $r_2 = 3$. This plays into our final calculation (which is most easily done in polar coordinates).

We start by applying the identity (5.5.1) to obtain that:

$$\int_{\partial D} \Phi = \int_D d\Phi.$$

For the integral of $d\Phi$ over D , we then proceed with standard integration of a differential form over a manifold. Namely, a calculation shows that

$$\begin{aligned} \int_D d\Phi &= \int_D 3x_1^2 dx_1 \wedge dx_2 \wedge dx_3 \wedge dx_4 - 3x_2^2 dx_2 \wedge dx_3 \wedge dx_4 \wedge dx_1 \\ &\quad + dx_3 \wedge dx_4 \wedge dx_1 \wedge dx_2 \\ &= \int_D (3x_1^2 + 3x_2^2 + 1) dx_1 \wedge dx_2 \wedge dx_3 \wedge dx_4 \\ &= \int \dots \int_D (3x_1^2 + 3x_2^2 + 1) dx_1 dx_2 dx_3 dx_4 \end{aligned}$$

(Make sure that you can explain each step of this calculation in detail!)

We now write the integral on D as an iterated integral. The integral in x_1 and x_2 can be evaluated using polar coordinates $x_1 = r \cos \theta$, $x_2 = r \sin \theta$:

$$\begin{aligned} \iint_{x_1^2 + x_2^2 \leq 4} (3x_1^2 + 3x_2^2 + 1) dx_1 dx_2 &= \int_0^2 \int_0^{2\pi} (3r^2 + 1) r d\theta dr \\ &= 2\pi \int_0^2 (3r^3 + r) dr = 2\pi \left(\frac{3r^4}{4} + \frac{r^2}{2} \right) \Big|_0^2 \\ &= 2\pi(12 + 2) = 28\pi. \end{aligned}$$

We finish with

$$\iint_{x_3^2 + x_4^2 \leq 9} 28\pi dx_3 dx_4 = 9\pi \cdot 28\pi = 252\pi^2.$$