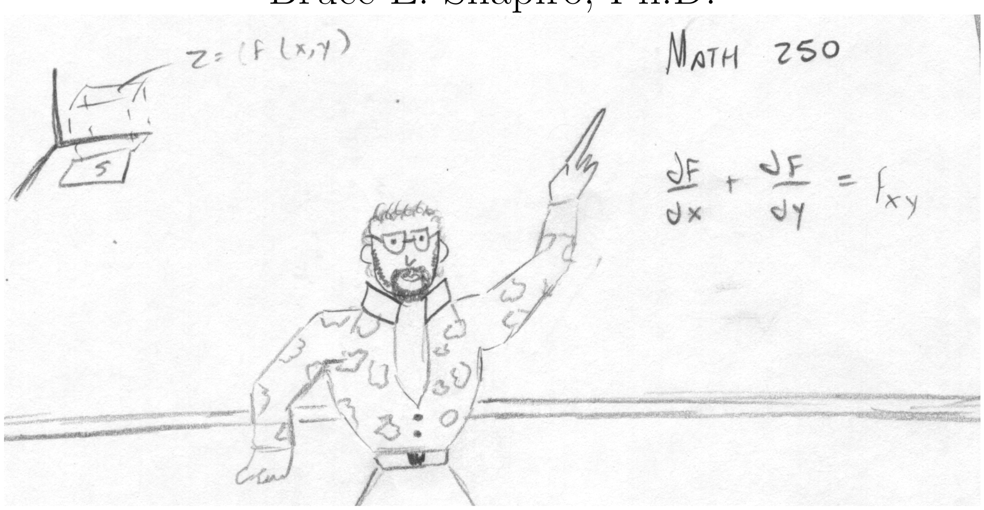
# Multivariate Calculus

in 25 Easy Lectures
(Revised: December 6, 2006)

Lecture Notes for Math 250[1]
California State University Northridge
Fall 2006
Class 13023

Bruce E. Shapiro, Ph.D.




[1]Lecture notes to accompany Varberg, Purcell & Rigdon, *Calculus, 8th Ed.*, in Math 250 at California State University Northridge. This document is intended solely for the use of students in Math 250, Class 13023, at California State University Northridge, during the the Fall 2006 Semester. Any other use is prohibited without prior permission. Please note that the contents of these lecture notes may not correspond to the material covered in other sections of Math 250 or to other editions of the textbook.  For further information, please contact: bruce.e.shapiro@csun.edu.


Typeset with $\LaTeX\,2_\varepsilon$ on December 6, 2006

# Contents

# Preface: A note to the Student

These lecture notes are meant to **accompany** the textbook, not to replace it. It has been my experience that students who attempt to use lecture notes (even mine, which I have to admit, are practically perfect) in lieu of the textbook receive abysmally poor final term grades. However, properly using the lecture notes can help you.

Here is how I would suggest you study in this class.

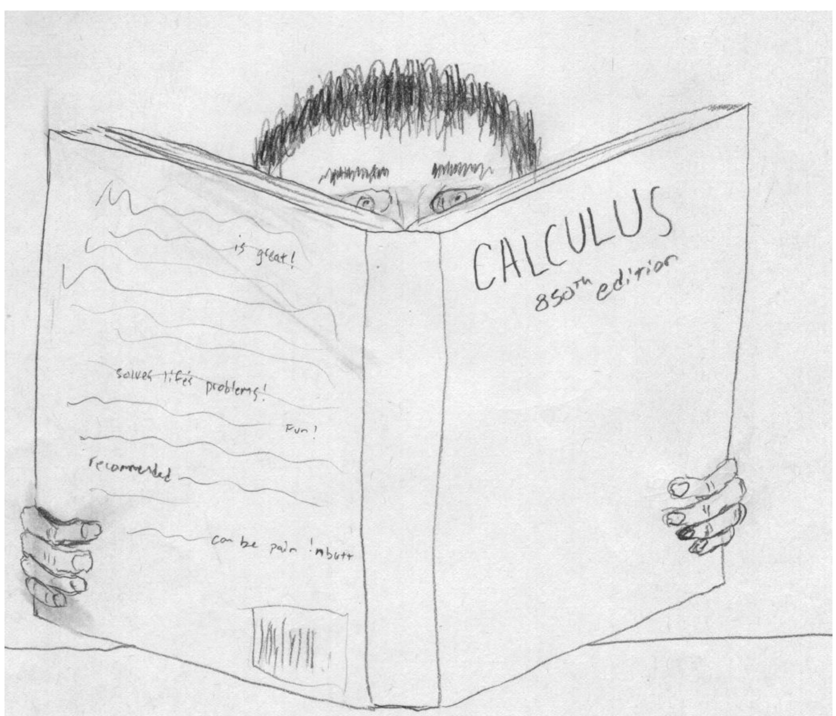
- Read the pertinent sections in the textbook prior to attending the class lecture.
- Come to class.
- Pay attention. Although I use the notes during class, I will not follow the notes verbatum, and I will most certainly not stand in front of the class and read the notes out loud to you. I will frequently also include additional material so you need to keep sharp.
- Take notes in class. Some students find it helpful to have a copy of these lecture notes in front of them, but others do not. It is a personal style, and you will have to decide what is right for you.
- Review your notes after class, before you try to do the homework. Compare your notes with the lecture notes; you can use the lecture notes to check to see if you copied the formulas correctly, or if you don't understand something in your notes.
- Reread or at least skim back over the chapter after reviewing your notes.
- Do the exercises at the end of the chapter. Note that this step comes AFTER you have read the chapter, not before.
- If you don't understand something, go back over the chapter, your class notes, and the lecture notes, in that order. If that doesn't help, ask me during office hours, during class, or via email for help.

Sometimes you will run into something in either the notes or the text that just plain does not make sense. If you can manage it, skip over the problem, or go on with the next step or sentence in what you are reading. Come back to it later, sometimes the answer will come to you.

And finally, even though the textbook is in its zillionth edition and I've been writing and revising these lecture notes since you were in kindergarten neither is perfect, and there are still plenty of typographical errors in each. Please let me know of any errors you discover. In particular, just prior to the fall 2006 term these notes were translated en-masse from one word processor (MS-Word) to another (LaTeX $2_\varepsilon$). Much of this translation was automated, and even computers sometimes makes mistakes ☺.

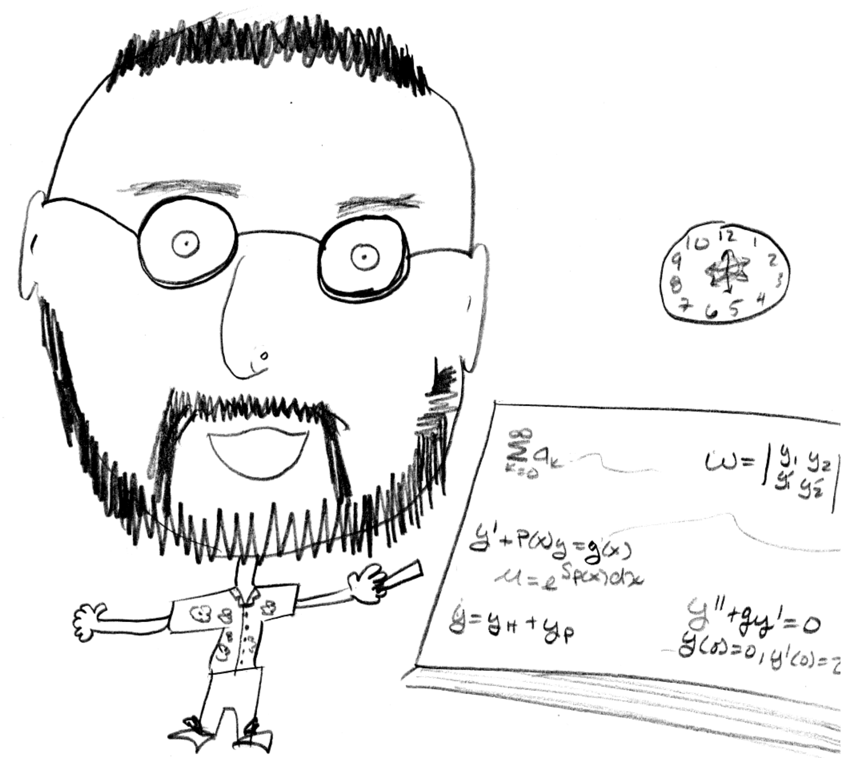
As your teacher I am here to help you. Feel free to contact me during my office hours with any questions or problems you have at all - that is what the office hours are for. And there is always email.

If you are a student in another Math 250 section at CSUN you are also welcome to use these notes. But you should check with your instructor first. There is no guarantee that he or she will approve of, agree with, or appreciate the material in these notes. Furthermore, there is virtually no chance at all that your teacher will cover the exact same material in the same order and with the same emphases placed on the different subtopics that I have chosen. These notes are provided without any guarantee whatsoever. Remember that you cannot use the excuse "but that's the way Shapiro did it in *his* lecture notes" in your class, especially if it turns out there is a bug in the notes. So be forewarned. And enjoy!



The order in which the material is covered is more-or-less the same order as it is presented in the textbook. The following table provides a cross-reference between the notes and the textbook.

| **Lecture** | **Topic** | **Text Reference** |
|---|---|---|
| 1 | Cartesian Coordinates | 14.1 |
| 2 | Vectors | 14.2 |
| 3 | Cross Products | 14.3 |
| 4 | Lines and Curves | 14.4 |
| 5 | Velocity, Acceleration, and Curvature | 14.5 |
| 6 | Surfaces | 14.6 |
| 7 | Cylindrical and Spherical Coordinates | 14.7 |
| 8 | Functions of 2 Variables | 15.1 |
| 9 | Partial Derivatives | 15.2 |
| 10 | Limits and Continuity | 15.3 |
| 11 | Gradients | 15.4, 15.5 |
| 12 | Chain Rule | 15.6 |
| 13 | Tangent Planes | 15.7 |
| 14 | Unconstrained Optimization | 15.8 |
| 15 | Constrained Optimization | 15.9 |
| 16 | Double Integrals: Rectangles | 16.1, 16.2 |
| 17 | Double Integrals: General | 16.3 |
| 18 | Double Integrals: Polar | 16.4 |
| 19 | Surface Area | 16.6 |
| 20 | Triple Integrals | 16.7, 16.8 |
| 21 | Vector Fields | 17.1 |
| 22 | Line Integrals | 17.2, 17.3 |
| 23 | Green's Theorem | 17.4 |
| 24 | Flux Integrals | 17.5, 17.6 |
| 25 | Stokes' Theorem | 17.7 |


y'+P(x)y=g(x)
y=yH+yP

# Examples of Typical Symbols Used

Table 1: Symbols Used in this Document

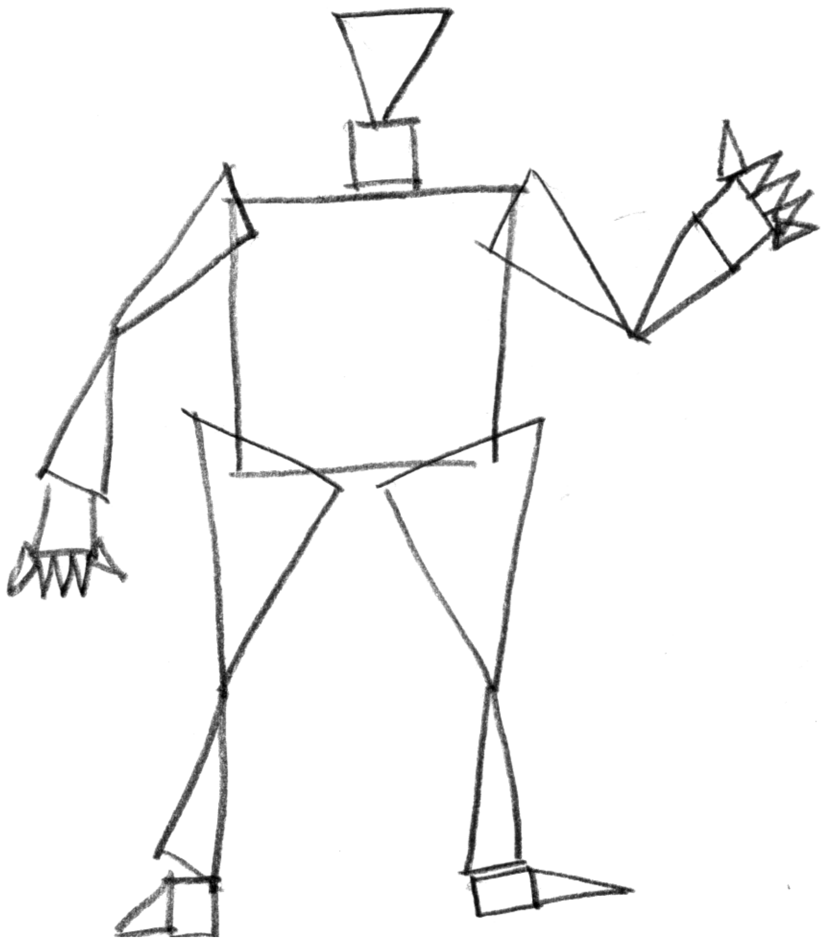
| **Symbol** | **Description** |
|---|---|
| $\blacksquare$ | Used to indicate the end of a proof or example |
| $\nabla$ | "Nabla" or "Del" gradient operator |
| $\nabla f$ | Gradient of a scalar function |
| $\nabla \cdot \mathbf{F}$ | divergence of a vector field |
| $\nabla \times \mathbf{F}$ | curl of a vector field |
| $a, b, \ldots$ | constants |
| $D_x f, D_y g, ..$ | partial derivative of $f$ w.r.t. $x$, or $g$ w.r.t. $y$ |
| $f_x, g_y, \ldots$ | |
| $\frac{\partial f}{\partial x}, \frac{\partial g}{\partial y}, \ldots$ | |
| $\partial S$ | boundary of surface C |
| $\partial f / \partial x$ | partial derivative |
| $D_{\mathbf{u}} f(\mathbf{P})$ | directional derivative of $f$ in the direction $\mathbf{u}$ at the point $\mathbf{P}$. |
| $f_{xx}, f_{yy}, f_{xy}$ | second partials, mixed partials |
| $f, g, h, \ldots$ | scalar functions |
| $f(x), g(y), \ldots$ | scalar functions of a single variable |
| $f(x, y), g(u, v), \ldots$ | scalar functions of two variables |
| $f(x, y, z), \ldots$ | scalar function of 3 variables |

Table 1: Symbols Used in this Document

| **Symbol** | **Description** |
|---|---|
| $\mathbf{F}, \mathbf{G}(x, y, ..),\ldots$ | Vector function, with or without explicit variables. |
| $f', f'(x), df/dx$ | the derivative of $f$ with respect to $x$ |
| $f : D \mapsto R$<br>$f : \mathbb{R}^n \mapsto \mathbb{R}^m$ | the function $f$ has a domain $D$ and a range $R$<br>the function $f$ has an n-dimensional domain<br>and an m-dimensioned range. |
| $\lim_{x\to y}$ | the limit as $x$ approaches $y$ |
| $\mathbf{u}, \mathbf{v}, \ldots$<br>$\mathbf{u} \cdot \mathbf{v},\ \mathbf{p} \cdot \mathbf{q},\ldots$<br>$\mathbf{u} \times \mathbf{v}$ | vectors<br>dot product between vectors<br>cross product between vectors |
| $\mathbf{P}, \mathbf{Q}, \ldots$ | points |
| $\mathbf{PQ}, \mathbf{RS},\ldots$ | line segments between points |
| $\overrightarrow{\mathbf{PQ}},\ldots$ | vector from $\mathbf{P}$ to $\mathbf{Q}$ |
| $\mathbb{R}, \mathbb{R}^2, \mathbb{R}^3$<br>$\mathbb{R}^n$ | real numbers, real line, 3D space<br>n-space |
| $x, y, z,\ldots$ | variables |
| $\|x\|, \|f(x)\|,\ldots$ | absolute values |
| $\\|\mathbf{v}\\|, \\|\mathbf{F}(x, y, z)\\|,\ldots$ | vector norms |
| $(x, y), \langle x, y\rangle, \begin{pmatrix} x \\ y \end{pmatrix}$ | two-dimensional vector with components $x$ and $y$ |
| $(x, y, z), \langle x, y, z\rangle, \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ | three-dimensional vector with components $x$, $y$, $z$ |
| $\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}, \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ | Matrices |

Table 1: Symbols Used in this Document

| **Symbol** | **Description** |
|---|---|
| $\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}, \begin{vmatrix} a & b \\ c & d \end{vmatrix}$ | Determinants |
| $\int, \iint, \iiint$ | single, double, triple integral |
| $\int_C, \iint_C, \iiint_C$ | integrals over a specific domain |
| $\int_a^b$ | definite integral |
| $\int_a^b \int_\alpha^\beta$ | iterated double integral |
| $\int_a^b \int_\alpha^\beta \int_\gamma^\mu$ | iterated double integral |



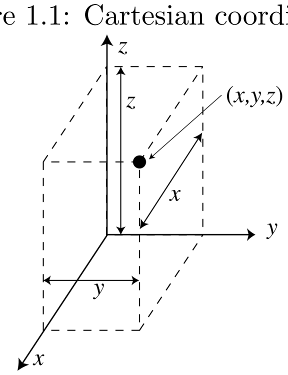
# Lecture 1

# Cartesian Coordinates

We express a **point** in three dimensions as an ordered sequence (x, y, z). In Cartesian coordinates, the third axis (the z-axis) is raised perpendicular to the x-y plane according to the **right-hand rule**:

1. Construct the x-y plane on a piece of paper.
2. Place your right hand on top of the paper at the origin with your fingers curling from the x-axis toward the y-axis.
3. Your thumb should be pointing upwards out of the paper. This is the direction of the z-axis.

Figure 1.1: Cartesian coordinates.




**Definition 1.1 (Distance Formula)** *The distance between two points* $\mathbf{P}_1 = (x_1, y_1, z_1)$ *and* $\mathbf{P}_2 = (x_2, y_2, z_2)$ *is calculated using the following distance formula*

$$|\mathbf{PQ}| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \tag{1.1}$$

**Example 1.1** *Find the distance between* $\mathbf{P} = (7, 5, -3)$ *and* $\mathbf{Q} = (12, 8, 6)$.

According to the distance formula,

$$\begin{aligned} |PQ| &= \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \\ &= \sqrt{(7-12)^2 + (5-8)^2 + (-3-6)^2} \\ &= \sqrt{(-5)^2 + (-3)^2 + (-9)^2} \\ &= \sqrt{25 + 9 + 81} = \sqrt{115} \approx 10.72 \end{aligned}$$

Therefore the distance is approximately 10.72.■

**Example 1.2** *Show that the triangle with vertices given by* $\mathbf{P} = (2, 1, 6)$, $\mathbf{Q} = (4, 7, 9)$ *and* $\mathbf{R} = (8, 5, -6)$ *is a right triangle.*

We use the fact that a right triangle must satisfy the Pythagorean theorem. The lengths of the three sides are:

$$\begin{aligned} |PQ| &= \sqrt{(4-2)^2 + (7-1)^2 + (9-6)^2} = \sqrt{4 + 36 + 9} = \sqrt{49} = 7 \\ |PR| &= \sqrt{(8-2)^2 + (5-1)^2 + (-6-6)^2} = \sqrt{36 + 16 + 144} = \sqrt{196} = 14 \\ |QR| &= \sqrt{(8-4)^2 + (5-7)^2 + (-6-9)^2} = \sqrt{16 + 4 + 225} = \sqrt{245} \end{aligned}$$

Since

$$|PQ|^2 + |PR|^2 = 49 + 196 = 245 = |QR|^2$$

we know that the triangle satisfies the Pythagorean theorem, and therefore it must be a right triangle. ■

**Definition 1.2** *A* ***sphere*** *is locus of all points that are some distance* $r$ *from some fixed point* $\mathbf{C}$. *The number* $r$ *is called the* ***radius*** *of the sphere.*

Suppose that $C = (x_0, y_0, z_0)$. Then the distance between $\mathbf{C}$ and any other point $\mathbf{P} = (x, y, z)$ is

$$|PC| = \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2} \tag{1.2}$$

If $\mathbf{P}$ is a distance $r$ from $\mathbf{C}$, then

$$r = \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2} \tag{1.3}$$

That is the condition that all points a distance $r$ from $\mathbf{C}$ must satisfy. Equivalently, the squares of both sides of (1.3) are equal to one another,

$$r^2 = (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 \tag{1.4}$$

**Definition 1.3** *Equation* (1.4) *is called the* ***standard Equation of a Sphere*** *of radius* $r$ *and center* $C = (x_0, y_0, z_0)$.

**Example 1.3** . *Find the center and radius of the sphere given by*

$$x^2 + y^2 + z^2 + 8x - 4y - 22z + 77 = 0 \tag{1.5}$$

Completing the squares in equation (1.5),

$$\begin{aligned} 0 &= x^2 + 8x + y^2 - 4y + z^2 - 22z + 77 \\ &= x^2 + 8x + 4^2 - 4^2 + y^2 - 4y + (-2)^2 - (-2)^2 + z^2 - 22z + (-11)^2 - (-11)^2 + 77 \\ &= (x+4)^2 - 16 + (y-2)^2 - 4 + (z-11)^2 - 121 + 77 \\ &= (x+4)^2 + (y-2)^2 + (z-11)^2 - 64 \end{aligned}$$

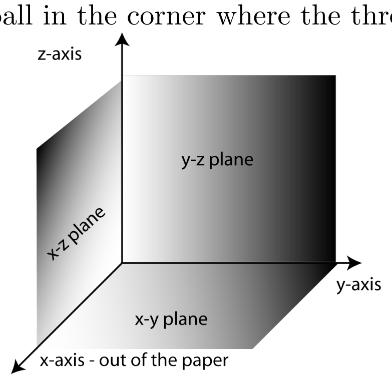
Rearranging,

$$(x+4)^2 + (y-2)^2 + (z-11)^2 = (8)^2 \tag{1.6}$$

Comparing equations (1.6) and (1.4) we conclude that the center of the sphere is at $\mathbf{C}_0 = (-4, 2, 11)$ and its radius is $r = 8$. ■

**Example 1.4** *Find the equation of a sphere that is tangent to the three coordinate planes whose radius is 6 and whose center is in the first octant* (see text, page 599,#27).[1]

Figure 1.2: Place the ball in the corner where the three walls come together.




Now imagine placing a ball that is 6 inches in diameter right in the corner. If you push the ball right about against both walls and the floor it will be 6 inches from each wall. Hence the center of the ball will be at

$$C = (6, 6, 6)$$

Since the radius of the ball is 6, the equation of the sphere is

$$(x-6)^2 + (y-6)^2 + (z-6)^2 = 36. \blacksquare$$

[1] The first octant is that region of space where $x > 0$, $y > 0$, and $z > 0$.

**Example 1.5** *Suppose that two spheres of equal radii have their centers at*

$$\begin{aligned} \mathbf{P} &= (-3, 1, 2) \\ \mathbf{Q} &= (5, -3, 6) \end{aligned}$$

*Find the equations of the two spheres if the two spheres are just touching (tangent) at precisely one point.* (See the text, page 599 #26.)

Since the two spheres are tangent at a single point, the radius is one half the distance between the two centers (think of two bowling balls that are just touching one another). Hence

$$\begin{aligned} 2r &= |PQ| = \sqrt{(-3-5)^2 + (1 - -3)^2 + (2-6)^2} \\ &= \sqrt{(-8)^2 + (4)^2 + (-4)^2} = \sqrt{64+16+16} = \sqrt{96} \end{aligned}$$

and therefore

$$r = \frac{1}{2}\sqrt{96} \Rightarrow r^2 = \frac{96}{4} = 24$$

The equation of the sphere centered at $\mathbf{P} = (-3, 1, 2) is$

$$(x+3)^2 + (x-1)^2 + (x-2)^2 = 24$$

while the equation of the sphere whose center is $\mathbf{Q} = (5, -3, 6)$ is

$$(x-5)^2 + (y+3)^2 + (z-6)^2 = 24. \blacksquare$$

**Theorem 1.1 Midpoint formula**. *The coordinates of a point halfway between the points*

$$P = (x_1, y_1, z_1)$$

*and*

$$Q = (x_2, y_2, z_2)$$

*are given by the formula*

$$(m_1, m_2, m_3) = \left( \frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2}, \frac{z_1 + z_2}{2} \right) \tag{1.7}$$

**Example 1.6** *Find the equations of the sphere that has the line segment between (-2,3,6) and (4,-1,5) as its diameter.*

The center of the sphere must be the center of the line segment between the two points (-2, 3, 6) and (4, -1, 5),

$$C = \left( \frac{-2+4}{2}, \frac{3 + -1}{2}, \frac{6+5}{2} \right) = (1, 1, 5.5)$$

The diameter of the sphere is the length of the line segment,

$$d = \sqrt{(-2-4)^2 + (3 - -1)^2 + (6-5)^2} = \sqrt{36+16+1} = \sqrt{53}$$

The radius is one-half the diameter, $r = \sqrt{53}/2$ The equation of the sphere is then

$$(x-1)^2 + (y-1)^2 + (z-5.5)^2 = 53/4. \blacksquare$$

**Definition 1.4** *A* **linear equation** *is any equation of the form*

$$Ax + By + Cz = D \tag{1.8}$$

*where A, B, C and D are constants and A, B and C are not all zero (D may be zero).*

**Theorem 1.2** *An equation is linear if and only if it is the equation of a plane, i.e., linear equations are equations of planes, and all equations of planes are linear equations.*

**Theorem 1.3 Properties of a linear equation**. *Suppose that* $Ax+By+Cz = D$*. Then*

1. *If* $A \neq 0$*, the x-intercept is at* $D/A$*. If* $A = 0$*, the plane is parallel to the x-axis.*
2. *If* $B \neq 0$*, The y-intercept is at* $D/B$*. If* $B = 0$*, the plane is parallel to the y-axis.*
3. *If* $C \neq 0$*, The z-intercept is at* $D/C$*. If* $C = 0$*, the plane is parallel to the z-axis*

If $D \neq 0$then the equation of a plane is

$$\frac{A}{D}x + \frac{B}{D}y + \frac{C}{D}z = 1$$

which we can rewrite as

$$\frac{x}{D/A} + \frac{y}{D/B} + \frac{z}{D/C} = 1$$

Let

$$\begin{aligned} a &= D/A \\ b &= D/B \\ c &= D/C \end{aligned}$$

Then

$$\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1$$

where $a$ is the $x$-intercept, $b$ is the $y$-intercept, and $c$ is the $z$-intercept.

**Example 1.7** *Sketch the plane* $-3x + 2y + (3/2)z = 6$

The x-intercept is at 6/(-3)=-2, which is the point (-2, 0, 0)

The y-intercept is at 6/2=3, which is the point (0, 3, 0)

The z-intercept is at 6/(3/2)=4, which is the point (0, 0, 4)

The plane is sketched in figure 1.3. ■

Figure 1.3: The plane in example 1.7 is sketched by first finding the three coordinate intercepts and then drawing a triangle to connect them.

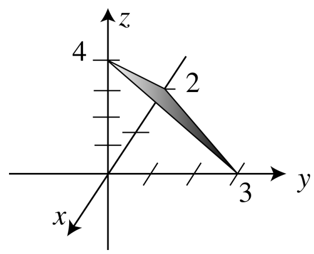
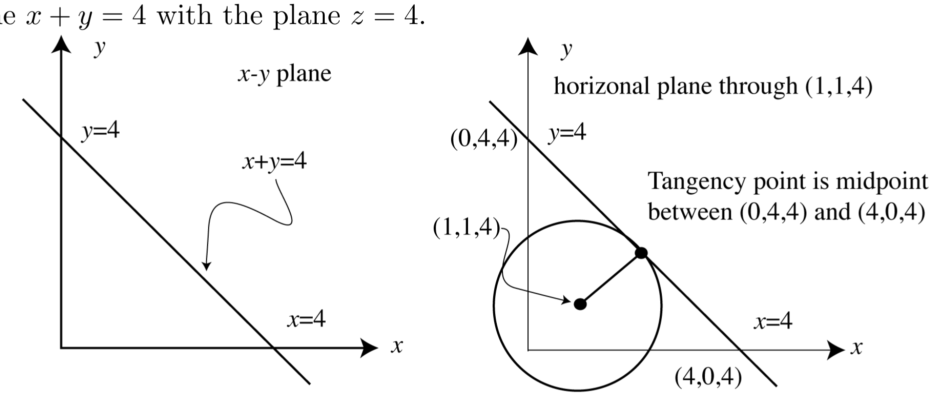



Figure 1.4: Left: the line $x + y = 4$ in the $xy$-plane. Right: cross section of the plane $x + y = 4$ with the plane $z = 4$.



**Example 1.8** *Find the equation of sphere that is tangent to the plane* $x + y = 4$ *with center at the point* $\mathbf{P} = (1, 1, 4)$*.* (See the text, page 599 #28.)

In two dimensions (the $xy$-plane), the equation $x+y = 4$ describes a line of slope $-1$ with $y$-intercept of 4. This line also crosses the $x$-axis (has $x$-intercept) at $x = 4$ (see figure 1.4.) In three dimensions, this equation $x + y = 4$ represents a vertical plane whose projection onto the $xy$-plane is the line $x + y = 4$.

To find the equation of the desired sphere, consider the horizontal plane passing through $\mathbf{P}$, namely, the plane $z = 4$. This plane sits parallel to and directly above the $xy$-plane, and from an illustrative point of view (see figure 1.4) looks just like the $xy$-plane. The center of the sphere will lay at (1, 1, 4), or at the point (1,1) of the equivalent $xy$-plane look-alike illustration. The cross section of the sphere with this plane is a circle of radius $r$ (whose value is not yet determined) that is tangent to the line $x + y = 4$ in this plane. The point of tangency is point on the line that is closest to the point (1,1). By symmetry, this must be the midpoint of the $x$-intercept and $y$-intercept, which is

$$\left(\frac{0+4}{2}, \frac{4+0}{2}, \frac{4+4}{2}\right) = (2, 2, 4)$$

The radius is the distance between the center at (1,1,4) and the point of tangency at (2,2,4):

$$r = \sqrt{(1-2)^2 + (1-2)^2 + (4-4)^2} = \sqrt{2}$$

Hence the equation of the sphere is

$$(x-1)^2 + (y-1)^2 + (z-4)^2 = 2. \blacksquare$$

# Lecture 2

# Vectors in 3D

## Properties of Vectors

**Definition 2.1** *A* **displacement vector v** *from point* **A** *to point* **B** *is an arrow pointing from* **A** *to* **B**, *and is denoted as*

$$\mathbf{v} = \vec{\mathbf{AB}} \tag{2.1}$$

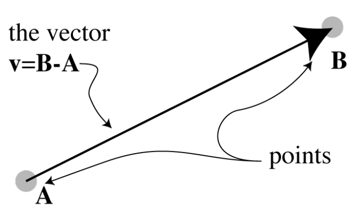
In general, we will use the same notation (e.g., a boldface letter such as **v**) to denote a vector that we use to describe a point (such as **A**) as well as a matrix. In most cases (but not always) we will use upper case letters for points (or matrices) and lower case letters for vectors. Whenever there is any ambiguity we will write a small arrow over the symbol for the vector, as in $\vec{\mathbf{v}}$, which means the same thing as **v**. A small arrow over a pair of points written next to each other, as in $\vec{\mathbf{AB}}$ is used to denote the displacement vector pointing from **A** to **B**. If **v** is a vector then $v$ denotes its magnitude:

**Definition 2.2** *The* **length** *or* **magnitude of a vector v** *is the distance measured from one end point to the other, and is denoted by the following equivalent notations:*

$$v = |\mathbf{v}| = \|\mathbf{v}\| \tag{2.2}$$

In print the notation **v** is more common for a vector; in handwritten documents (and some textbooks) it is usual to write $\vec{v}$ for a vector.

Figure 2.1: Concept of a vector as the difference between two points.



If the point $\mathbf{A} = (x_a, y_a, z_a)$ and the point $\mathbf{B} = (x_b, y_b, z_b)$ then we define the **components** of the vector $\mathbf{v} = \vec{\mathbf{AB}}$ as $\mathbf{v} = (v_x, v_y, v_z)$, where

$$\mathbf{v} = (v_x, v_y, v_z) = (x_b - x_a, y_b - y_a, z_b - z_a) \tag{2.3}$$

We observe that *the difference of two points is a vector*, and write this as

$$\mathbf{v} = \vec{\mathbf{AB}} = \mathbf{B} - \mathbf{A} \tag{2.4}$$

There is no equivalent concept of the sum of two points, although we will see that it is possible to add two vectors.

The text uses angle brackets $\langle\rangle$ to denote a vector in terms of its components:

$$\langle a, b, c\rangle = (a, b, c) \tag{2.5}$$

The use of parenthesis is more common, although angle brackets are used whenever there is some possibility of confusion between vectors and points.

**Theorem 2.1** *The magnitude of a vector* $\mathbf{v} = (v_x, v_y, v_z)$ *is given by*

$$\|\mathbf{v}\| = \sqrt{v_x^2 + v_y^2 + v_z^2} \tag{2.6}$$

*Proof.* According to the distance formula, the distance from $\mathbf{A}$ to $\mathbf{B}$ is

$$\|\mathbf{v}\| = \sqrt{(x_b - x_a)^2 + (y_b - y_a)^2 + (z_b - z_a)^2} = \sqrt{v_x^2 + v_y^2 + v_z^2}. \blacksquare$$

**Definition 2.3** *Two vectors* $\mathbf{v}$ *and* $\mathbf{w}$ *are said to be* **equal** *if they have the same magnitude and direction i.e., they have the same length and are parallel, and we write* $\mathbf{v} = \mathbf{w}$

**Theorem 2.2** *Two vectors* $\mathbf{v} = (v_x, v_y, v_z)$ *and* $\mathbf{w} = w_x, w_y, w_z$ *are equal if and only if their components are equal, namely the following three conditions all hold:*

$$v_x = w_x, v_y = w_y, v_z = w_z \tag{2.7}$$

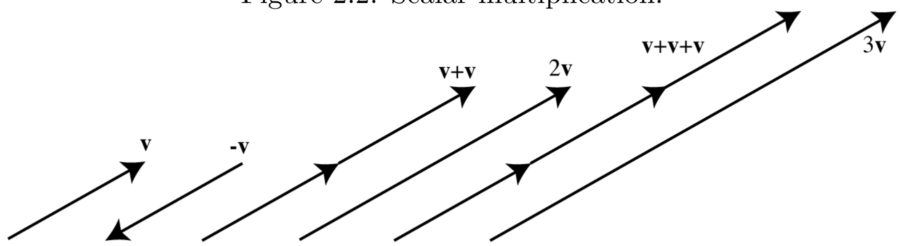
**Definition 2.4** *The* **zero vector***, denoted by* $\mathbf{0}$*, is a vector with zero magnitude (and undefined direction).*

## Operations on Vectors

**Definition 2.5 Scalar Multiplication** *or* **multiplication of a vector by a scalar** *is defined as follows. Suppose that* $a \in \mathbb{R}$ *is any real number and* $\mathbf{v} = v_x, v_y, v_z \in \mathbb{R}^3$ *is a vector. Then*

$$a\mathbf{v} = a(v_x, v_y, v_z) = (av_x, av_y, av_z) \tag{2.8}$$

Figure 2.2: Scalar multiplication.



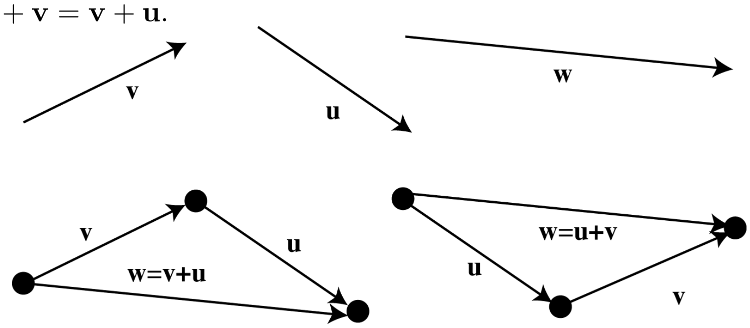

Scalar multiplication changes the length of a vector, as illustrated in figure 2.2. As a consequence of definitions 2.3 and 2.4, we also conclude that

$$\mathbf{0} = (0, 0, 0) \tag{2.9}$$

**Definition 2.6 Vector additive inverse** *or* **the negative of a vector**. *Suppose that* $\mathbf{v} = (v_x, v_y, v_z)$. *Then*

$$-\mathbf{v} = (-1)\mathbf{v} = (-v_x, -v_y, -v_z) \tag{2.10}$$

Figure 2.3: Illustration of vector addition. Top: the three vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w} = \mathbf{u} + \mathbf{v}$. Bottom: vector addition, and illustration of commutative law for vector addition: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.




**Vector addition**, illustrated in figure 2.3 proceeds as follows: Join the two vectors head to tail as show in the figure to the right and then draw the arrow from the tail to the head. Addition is commutative: the order in which the vectors are added does not matter. A vector sum can be calculated component-by-component; suppose that $\mathbf{v} = (v_x, v_y, v_z)$ and $\mathbf{u} = (u_x, u_y, u_z)$. Then

$$\mathbf{u} + \mathbf{v} = (u_x, u_y, u_z) + (v_x, v_y, v_z) = (u_x + v_x, u_y + v_y, u_z + v_z) \tag{2.11}$$

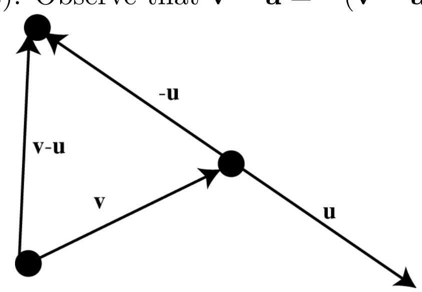
**Vector subtraction** is similar, and is induced by the idea that we want

$$\mathbf{v} - \mathbf{u} = \mathbf{v} + -\mathbf{u} \tag{2.12}$$

To obtain **u - v**, start (1) by placing the second vector at the head of the first and then (2) reflect it across its own tail to find $-\mathbf{u}$. Finally, (3) draw the arrow from the tail of u to the head of the reflection of v. Observe that this process is not reversible: **u - v** is not the same as **v - u**. In terms of components,

$$\mathbf{u}-\mathbf{v}=(u_x,u_y,u_z)-(v_x,v_y,v_z)=(u_x-v_x,u_y-v_y,u_z-v_z)=-(\mathbf{v}-\mathbf{u}) \tag{2.13}$$

Figure 2.4: Vector subtraction, demonstrating construction of $\mathbf{v}-\mathbf{u}$ (left) and $\mathbf{v}-\mathbf{u}$ (right). Observe that $\mathbf{v}-\mathbf{u}=-(\mathbf{v}-\mathbf{u})$




**Theorem 2.3** ***Properties of Vector Addition & Scalar Multiplication***. *Let* **v**, **u**, *and* **w** *be vectors, and* $a, b \in \mathbb{R}$ *be real numbers. Then the following properties hold:*

1. **Vector addition commutes:**

$$\mathbf{v}+\mathbf{u}=\mathbf{u}+\mathbf{v} \tag{2.14}$$

2. **Vector addition is associative:**

$$(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w}) \tag{2.15}$$

3. **Scalar multiplication is distributive across vector addition:**

$$(a+b)\mathbf{v} = a\mathbf{v}+b\mathbf{v} \tag{2.16}$$
$$a(\mathbf{v}+\mathbf{w}) = a\mathbf{v}+a\mathbf{w} \tag{2.17}$$

4. **Identity for scalar multiplication**

$$1\mathbf{v}=\mathbf{v}1=\mathbf{v} \tag{2.18}$$

5. **Properties of the zero vector**

$$0\mathbf{v}=\mathbf{0} \tag{2.19}$$
$$\mathbf{0}+\mathbf{v}=\mathbf{v}+\mathbf{0}=\mathbf{v} \tag{2.20}$$

**Definition 2.7** *The* **dot product** *or* **scalar product** $\mathbf{u}\cdot\mathbf{v}$ *between two vectors* $\mathbf{v} = (v_x, v_y, v_z)$ *and* $\mathbf{u} = (u_x, u_y, u_z)$ *is the scalar (number)*

$$\mathbf{u}\cdot\mathbf{v} = u_x v_x + u_y v_y + u_z v_z \tag{2.21}$$

Sometimes we will find it convenient to represent vectors as column matrices. The **column matrix** representation of $\mathbf{v}$ is given by the components $v_x$, $v_y$, and $v_z$ represented in a column, e.g.,

$$\mathbf{v} = \begin{pmatrix} v_x \\ v_y \\ vz \end{pmatrix}, \mathbf{u} = \begin{pmatrix} u_x \\ u_y \\ uz \end{pmatrix} \tag{2.22}$$

With this notation, the **transpose** of a vector is

$$\mathbf{u}^T = \begin{pmatrix} u_x & u_y & u_z \end{pmatrix} \tag{2.23}$$

and the dot product is

$$\mathbf{u}\cdot\mathbf{v} = \mathbf{u}^T\mathbf{v} = \begin{pmatrix} u_x & u_y & u_z \end{pmatrix} \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} = u_x v_x + u_y v_y + u_z v_z \tag{2.24}$$

which gives the same result as equation (2.21).

**Example 2.1** *Find the dot product of* $\vec{u} = (1, -3, 7)$ *and* $\vec{v} = (16, 4, 1)$

*Solution.* By equation (2.21),

$$\mathbf{u}\cdot\mathbf{v} = (1)(16) + (-3)(4) + (7)(1) = 16 - 12 + 7 = 11. \blacksquare$$

**Theorem 2.4** $\|\mathbf{v}\| = \sqrt{\mathbf{v}\cdot\mathbf{v}}$

*Proof.* From equation (2.21),

$$\mathbf{v}\cdot\mathbf{v} = v_x v_x + v_y v_y + v_z v_z = \|\mathbf{v}\|^2$$

where the last equality follows from equation (2.6).■

**Example 2.2** *Find the lengths of the vectors* $\vec{u} = (1, -3, 7)$ *and* $\vec{v} = (16, 4, 1)$

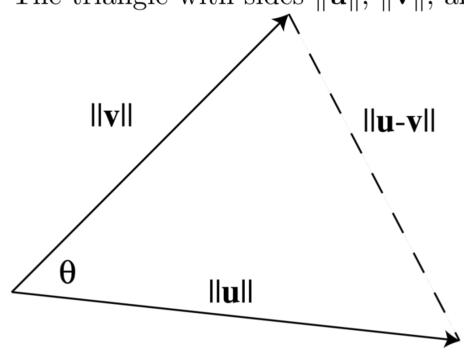
*Solutions.*

$$\begin{aligned} \|\mathbf{u}\| &= \sqrt{(1)^2 + (-3)^2 + (7)^2} = \sqrt{1 + 9 + 49} = \sqrt{59} \approx 7.681 \\ \|\mathbf{v}\| &= \sqrt{(16)^2 + (4)^2 + (1)^2} = \sqrt{256 + 16 + 1} = \sqrt{273} \approx 16.523 \blacksquare \end{aligned}$$

**Theorem 2.5** $\mathbf{u} \cdot \mathbf{v} = ||\mathbf{u}|| \|\mathbf{v}\| \cos\vartheta$, *where* $\vartheta$ *is the angle between the two vectors* $\mathbf{u}$ *and* $\mathbf{v}$.

*Proof.* We can use the law of cosines to determine the length $\|\mathbf{u} - \mathbf{v}\|$ (see figure 2.5).

Figure 2.5: The triangle with sides $\|\mathbf{u}\|$, $\|\mathbf{v}\|$, and $\|\mathbf{u} - \mathbf{v}\|$.




$$\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\|\|\mathbf{v}\| \cos\theta$$

Hence

$$\begin{aligned} 2\|\mathbf{u}\|\|\mathbf{v}\| \cos\theta &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - \|\mathbf{u} - \mathbf{v}\|^2 \\ &= \mathbf{u} \cdot \mathbf{u} + \mathbf{v} \cdot \mathbf{v} - (\mathbf{u} - \mathbf{v}) \cdot (\mathbf{u} - \mathbf{v}) \\ &= \mathbf{u} \cdot \mathbf{u} + \mathbf{v} \cdot \mathbf{v} - [\mathbf{u} \cdot (\mathbf{u} - \mathbf{v}) - \mathbf{v} \cdot (\mathbf{u} - \mathbf{v})] \\ &= \mathbf{u} \cdot \mathbf{u} + \mathbf{v} \cdot \mathbf{v} - [\mathbf{u} \cdot \mathbf{u} - \mathbf{u} \cdot \mathbf{v} - \mathbf{v} \cdot \mathbf{u} + \mathbf{v} \cdot \mathbf{v}] \\ &= 2\, \mathbf{u} \cdot \mathbf{v} \end{aligned}$$

Therefore the dot product can be written as

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\|\|\mathbf{v}\| \cos\theta \quad \blacksquare$$

**Corollary 2.1** . *Two vectors are perpendicular if and only if their dot product is zero.*

**Example 2.3** *Find the angle between the two vectors in* $\vec{u} = (1, -3, 7)$ *and* $\vec{v} = (16, 4, 1)$

*Solution.* We use the fact that $\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\|\|\mathbf{v}\| \cos\theta$. From the previous examples, we have

$$\begin{aligned} \|\mathbf{u}\,| &= \sqrt{59} \\ \|\mathbf{v}\| &= \sqrt{273} \\ \mathbf{u} \cdot \mathbf{v} &= (1)(16) + (-3)(4) + (7)(1) = 16 - 12 + 7 = 11. \end{aligned}$$

Hence

$$\begin{aligned}
\mathbf{u}\cdot\mathbf{v} &= \|\mathbf{u}\|\|\mathbf{v}\|\cos\theta \\
&\Rightarrow 11 = \sqrt{59}\sqrt{273}\cos\theta = \sqrt{16107}\cos\theta \\
&\Rightarrow \cos\theta = \frac{11}{\sqrt{16107}} \approx \frac{11}{127} \approx 0.087 \\
&\Rightarrow \theta \approx \arccos 0.087 \approx 29\,\text{deg}\ \blacksquare
\end{aligned}$$

**Example 2.4** *The basis vectors are all mutually orthogonal to one another:*

$$\begin{aligned}
\mathbf{i}\cdot\mathbf{i} = \mathbf{j}\cdot\mathbf{j} = \mathbf{k}\cdot\mathbf{k} = 1 \\
\mathbf{i}\cdot\mathbf{j} = \mathbf{j}\cdot\mathbf{k} = \mathbf{k}\cdot\mathbf{i} = 0
\end{aligned}$$

## Unit Vectors, Direction Angles and Direction Cosines

**Definition 2.8** *A* **unit vector** *is any vector with a magnitude of 1. If* $\mathbf{v}$ *is a unit vector, it is (optionally) denoted by* $\hat{\mathbf{v}}$.

**Theorem 2.6** *Let* $\mathbf{v} \neq \mathbf{0}$ *be any vector. Then a unit vector* $\hat{\mathbf{v}}$ *parallel to* $\mathbf{v}$ *is*

$$\hat{\mathbf{v}} = \frac{\mathbf{v}}{\|\mathbf{v}\|}$$

*Proof.* To verify that $\frac{\mathbf{v}}{\|\mathbf{v}\|}$ is a unit vector,

$$\begin{aligned}
\left\|\frac{\mathbf{v}}{\|\mathbf{v}\|}\right\| &= \left\|\left(\frac{v_x}{\|\mathbf{v}\|}, \frac{v_y}{\|\mathbf{v}\|}, \frac{v_z}{\|\mathbf{v}\|}\right)\right\| \\
&= \sqrt{\left(\frac{v_x}{\|\mathbf{v}\|}\right)^2 + \left(\frac{v_y}{\|\mathbf{v}\|}\right)^2 + \left(\frac{v_z}{\|\mathbf{v}\|}\right)^2} \\
&= \frac{1}{\|\mathbf{v}\|}\sqrt{v_x^2 + v_y^2 + v_z^2} \\
&= \frac{\|\mathbf{v}\|}{\|\mathbf{v}\|} = 1
\end{aligned}$$

To see that $\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|}$ and $\mathbf{v}$ are parallel,

$$\mathbf{u}\cdot\mathbf{v} = \frac{\mathbf{v}\cdot\mathbf{v}}{\|\mathbf{v}\|} = \frac{\|\mathbf{v}\|^2}{\|\mathbf{v}\|} = \|\mathbf{v}\|$$

but, letting $\theta$ be the angle between $\mathbf{u}$ and $\mathbf{v}$

$$\mathbf{u}\cdot\mathbf{v} = \|\mathbf{u}\|\|\mathbf{v}\|\cos\theta = \|\mathbf{v}\|\cos\theta$$

Equating the two expressions for $\mathbf{u}\cdot\mathbf{v}$ gives $\cos\theta = 1$ or $\theta = 1$. Hence the vectors are parallel. Since $\mathbf{u}$ has magnitude one and is parallel to $\mathbf{v}$ we conclude that $\hat{\mathbf{v}} = \mathbf{u} = \mathbf{v}/\|\mathbf{v}\|$. $\blacksquare$

**Example 2.5** *Find* $\hat{\mathbf{v}}$*, where* $\mathbf{v} = (3, 0, -4)$*.*

*Solution.* Since

$$\|\mathbf{v}\| = \sqrt{3^2 + (-4)^2} = 5$$

then

$$\hat{\mathbf{v}} = \frac{\mathbf{v}}{\|\mathbf{v}\|} = \left(\frac{3}{5}, 0, -\frac{4}{5}\right) \blacksquare$$

Three special unit vectors that are parallel to the $x$, $y$ and $z$ axes are often defined,

$$\mathbf{i} = (1, 0, 0) \tag{2.25}$$

$$\mathbf{j} = (0, 1, 0) \tag{2.26}$$

$$\mathbf{k} = (0, 0, 1) \tag{2.27}$$

In terms of **i**, **j** and **k**, any vector $\mathbf{v} = (v_x, v_y, v_z)$ can also be written as

$$\begin{aligned} v &= (v_x, v_y, v_z) \\ &= (v_x, 0, 0) + (0, v_y, 0) + (0, 0, v_z) \\ &= v_x(1, 0, 0) + v_y(0, 1, 0) + v_z(0, 0, 1) \\ &= v_x\mathbf{i} + v_y\mathbf{j} + v_z\mathbf{k} \end{aligned}$$

The unit vectors **i**, **j** and **k** are sometimes called the **basis vectors of Euclidean space**. Since **i**, **j** and **k** are always unit vectors, we will usually refer to them without the "hat."

If $\mathbf{v} = (v_x, v_y, v_z)$ then

$$\begin{aligned} \mathbf{v} \cdot \mathbf{i} = (v_x, v_y, v_z) \cdot (1, 0, 0) = v_x \\ \mathbf{v} \cdot \mathbf{j} = (v_x, v_y, v_z) \cdot (0, 1, 0) = v_y \\ \mathbf{v} \cdot \mathbf{k} = (v_x, v_y, v_z) \cdot (0, 0, 1) = v_z \end{aligned}$$

**Definition 2.9** *The* **direction angles** $\{\alpha, \beta, \gamma\}$ *of a vector are the angles between the vector and the three coordinate axes.*

**Definition 2.10** *The* **direction cosines** *of a vector are the cosines of its direction angles.*

Since the vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are parallel to the three coordinate axes, we have

$$\begin{aligned} \mathbf{v} \cdot \mathbf{i} &= \|\mathbf{v}\| \cos\alpha \\ \mathbf{v} \cdot \mathbf{j} &= \|\mathbf{v}\| \cos\beta \\ \mathbf{v} \cdot \mathbf{k} &= \|\mathbf{v}\| \cos\gamma \end{aligned}$$

Thus the three direction cosines are

$$\cos\alpha = \frac{\mathbf{v} \cdot \mathbf{i}}{\|\mathbf{v}\|}, \cos\beta = \frac{\mathbf{v} \cdot \mathbf{j}}{\|\mathbf{v}\|}, \cos\gamma = \frac{\mathbf{v} \cdot \mathbf{k}}{\|\mathbf{v}\|} \tag{2.28}$$

and the direction angles are

$$\alpha = \arccos\left(\frac{\mathbf{v}\cdot\mathbf{i}}{\|\mathbf{v}\|}\right), \beta = \arccos\left(\frac{\mathbf{v}\cdot\mathbf{j}}{\|\mathbf{v}\|}\right), \gamma = \arccos\left(\frac{\mathbf{v}\cdot\mathbf{k}}{\|\mathbf{v}\|}\right) \tag{2.29}$$

**Example 2.6** *Find the direction cosines and angles of the vector* $\mathbf{v} = (4, -2, -4)$

*Solution.* First, we calculate the magnitude of $\mathbf{v}$,

$$\|\mathbf{v}\| = \sqrt{(4)^2 + (-2)^2 + (-4)^2} = 6$$

Thus the three direction cosines are

$$\begin{aligned} \cos\alpha &= \frac{\mathbf{v}\cdot\mathbf{i}}{\|\mathbf{v}\|} = \frac{4}{6} = \frac{2}{3} \\ \cos\beta &= \frac{\mathbf{v}\cdot\mathbf{j}}{\|\mathbf{v}\|} = -\frac{2}{6} = -\frac{1}{3} \\ \cos\gamma &= \frac{\mathbf{v}\cdot\mathbf{k}}{\|\mathbf{v}\|} = -\frac{4}{6} = -\frac{2}{3} \end{aligned}$$

and the direction cosines are

$$\begin{aligned} \alpha = \arccos\left(\frac{2}{3}\right) &\approx 48.19\,\text{deg} \\ \beta = \arccos\left(-\frac{1}{3}\right) &\approx 109.47\,\text{deg} \\ \gamma = \arccos\left(\frac{2}{3}\right) &\approx 131.81\,\text{deg}\ \blacksquare \end{aligned}$$

## Projection of one vector on another vector.

Consider any two vectors $\mathbf{u}$ and $\mathbf{v}$, as illustrated in figure 2.6. We can always express $\mathbf{v}$ as the sum of two vectors:

$$\mathbf{v} = \mathbf{m} + \mathbf{n} \tag{2.30}$$

where $\mathbf{m}$ is parallel to $\mathbf{u}$ and $\mathbf{n}$ is perpendicular to $\mathbf{u}$. Let $\theta$ bet the angle between $\mathbf{u}$ and $\mathbf{v}$. Then

$$\|\mathbf{m}\| = \|\mathbf{v}\|\cos\theta \tag{2.31}$$

Since $\mathbf{m}$ and $\mathbf{u}$ are parallel, then they must have the same unit vectors, so that

$$\hat{\mathbf{m}} = \hat{\mathbf{u}}$$

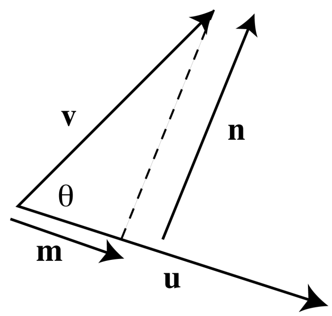
Therefore

$$\mathbf{m} = \|\mathbf{m}\|\hat{\mathbf{m}} = (\|\mathbf{v}\|\cos\theta)\,\hat{\mathbf{m}} = (\|\mathbf{v}\|\cos\theta)\,\hat{\mathbf{u}} = (\mathbf{v}\cdot\hat{\mathbf{u}})\,\hat{\mathbf{u}} \tag{2.32}$$

and since $\mathbf{v} = \mathbf{m} + \mathbf{n}$,

$$\mathbf{n} = \mathbf{v} - \mathbf{m} \tag{2.33}$$

Figure 2.6: A vector **u** is expressed as the sum of its components **m** parallel to **u** and **n** perpendictular to **u**.




**Definition 2.11** *The* **projection of a vector v on u** *is*

$$pr_{\mathbf{u}}((v)) = \mathbf{m} = (\mathbf{v}\cdot\hat{\mathbf{u}})\,\hat{\mathbf{u}} \tag{2.34}$$

**Example 2.7** *Express* $\mathbf{v} = (-3, 2, 1)$ *as the sum of vectors* **m** *parallel and* **n** *perpendicular to* $\mathbf{u} = (-3, 5, -3)$.

*Solution.*

$$\begin{aligned}
\|\mathbf{u}\| &= \sqrt{(-3)^2 + (5)^2 + (-3)^2} = \sqrt{43} \\
\hat{\mathbf{u}} &= \frac{\mathbf{u}}{\|\mathbf{u}\|} = \frac{1}{\sqrt{43}}\,(-3, 5, -3) \\
\mathbf{v}\cdot\hat{\mathbf{u}} &= \frac{1}{\sqrt{43}}\,((-3)(-3) + (5)(2) + (-3)(1)) = \frac{16}{\sqrt{43}} \\
\mathbf{m} &= (\mathbf{v}\cdot\hat{\mathbf{u}})\,\hat{\mathbf{u}} = \frac{16}{43}\,(-3, 5, -3) = \left(-\frac{48}{43}, \frac{80}{43}, -\frac{48}{43}\right) \\
\mathbf{n} &= \mathbf{v} - \mathbf{m} = (-3, 2, 1) - \left(-\frac{48}{43}, \frac{80}{43}, -\frac{48}{43}\right) = \frac{1}{7}\,(3, -26, 31)\,. \blacksquare
\end{aligned}$$

**Example 2.8** *Express* $\mathbf{v} = (2, -1, -2)$ *as the sum of vectors* **m** *parallel and* **n** *perpendicular to* $\mathbf{u} = (2, 4, 5)$.

*Solution.*

$$\begin{aligned}
\|\mathbf{u}\| &= \sqrt{(2)^2 + (4)^2 + (5)^2} = \sqrt{45} \\
\hat{\mathbf{u}} &= \frac{\mathbf{u}}{\|\mathbf{u}\|} = \frac{1}{\sqrt{45}}\,(2, 4, 5) \\
\mathbf{v}\cdot\hat{\mathbf{u}} &= \frac{1}{\sqrt{45}}\,((2)(2) + (-1)(4) + (-2)(5)) = -\frac{10}{\sqrt{45}} \\
\mathbf{m} &= (\mathbf{v}\cdot\hat{\mathbf{u}})\,\hat{\mathbf{u}} = -\frac{10}{45}\,(2, 4, 5) = -\frac{1}{9}\,(4, 8, 10) \\
\mathbf{n} &= \mathbf{v} - \mathbf{m} = (2, -1, -2) - \frac{1}{9}\,(4, 8, 10) = \left(-\frac{81}{43}, \frac{6}{43}, \frac{91}{43}\right) \blacksquare
\end{aligned}$$

## The Equation of a Plane

**To find the equation of a plane** through a point $\mathbf{P} = (x_1, y_1, z_1)$ that is perpendicular to some vector $\mathbf{n} = (A, B, C)$,

1. Let $\mathbf{Q} = (x, y, z)$ be any other point in the plane.
2. Let $\mathbf{v}$ be a vector in the plane pointing from $\mathbf{P}$ to $\mathbf{Q}$:

$$\mathbf{v} = \mathbf{Q} - \mathbf{P} = (x - x_1, y - y_1, z - z_1)$$

3. Since $\mathbf{v}$ must be perpendicular to $\mathbf{n}$, the dot product $\mathbf{v} \cdot \mathbf{n} = 0$, hence

$$A(x - x_1) + B(y - y_1) + (z - z_1) = 0$$

4. Rearrange to give

$$Ax + By + Cz = Ax_1 + By_1 + Cz_1 = \mathbf{n} \cdot \mathbf{P}$$

**Example 2.9** *Find the equation of a plane passing through (5, 1, -7) that is perpendicular to the vector (2,1,5)*

*Solution* By the construction described above, the equation is

$$2x + y + 5z = (2)(5) + (1)(1) + (5)(-7) = -24 \blacksquare$$

**Example 2.10** *Find the angle between the planes*

$$3x - 4y + 7z = 5$$

*and*

$$2x - 3y + 4z = 0$$

*Solution.* We find the normal vectors to the planes by reading off the coefficients of $x$, $y$, and $z$; the normal vectors are $\mathbf{n} = (3, -4, 7)$ and $\mathbf{m} = (2, -3, 4)$. Their angle of intersection $\theta$ is the same as the angle between the planes. To get this angle, we take the dot product, since $\mathbf{m} \cdot \mathbf{n} = \|\mathbf{m}\|\mathbf{n}\| \cos\theta$. We calculate that

$$\begin{aligned}
\|\mathbf{m}\| &= \sqrt{29} \\
\|\mathbf{n}\| &= \sqrt{74} \\
\mathbf{m} \cdot \mathbf{n} &= (2)(3) + (-3)(-4) + (4)(7) = 6 + 12 + 28 = 46 \\
\cos\theta &= \frac{\mathbf{m} \cdot \mathbf{n}}{\|\mathbf{m}\|\|\mathbf{n}\|} = \frac{46}{(29)(74)} = \frac{46}{2146} \approx 0.0214 \\
\theta &= \arccos .214 \approx 88.8 \deg \blacksquare
\end{aligned}$$

**Example 2.11** *Find the equation of a plane through* $(-1, 2, -3)$ *and parallel to the plane* $2x + 4y - z = 6$.

*Solution.* Since both planes are parallel, they must have the same normal vector. Hence

$$\mathbf{n} = (2, 4, -1)$$

Since in general the equation of any plane

$$\mathbf{n} \cdot (x, y, z) = \mathbf{n} \cdot \mathbf{P} \tag{2.35}$$

where $\mathbf{P}$ is a specific point on the plane and $\mathbf{n}$ is any vector normal to the plane, then the plane we are looking for is given by

$$2x + 4y - z = (2, 4, -1) \cdot (-1, 2, -3) = (2)(-1) + (4)(2) + (-1)(-3) = 9. \blacksquare$$

# Lecture 3

# The Cross Product

**Definition 3.1** *Let* $\mathbf{v}$ *and* $\mathbf{w}$ *be vectors. Then the* **cross product** $\mathbf{v} \times \mathbf{w}$ *is a vector with magnitude*

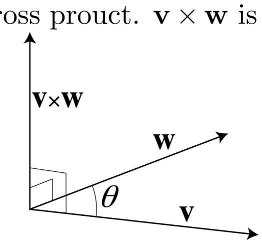
$$\|\mathbf{v} \times \mathbf{w}\| = \|\mathbf{v}\|\|\mathbf{w}\| \sin\theta \tag{3.1}$$

*and whose direction is perpendicular to the plane that contains* $\mathbf{v}$ *and* $\mathbf{w}$ *according to the* **right hand rule***:*

1. Place $\mathbf{v}$, $\mathbf{w}$ so that their tails of the vector coincide;
2. curl the fingers of your right hand from through the angle from $\mathbf{v}$ to $\mathbf{w}$.
3. Your thumb is pointing in the direction of $\mathbf{v} \times \mathbf{w}$

The construction of the cross product is illustrated in figure 3.1. Geometrically, the cross product gives the area of the parallelogram formed by the two vectors, as illustrated in figure 3.2.

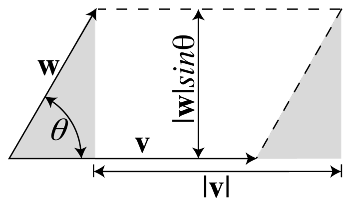
Figure 3.1: Geometry of the cross prouct. $\mathbf{v} \times \mathbf{w}$ is perpendicular to both $\mathbf{v}$ and $\mathbf{w}$.




There are a number of different ways to calculate the cross product by components. For example, suppose that $\mathbf{u} = (\alpha, \beta, \gamma)$ and $\mathbf{v} = (a, b, c)$. Then to calculate $\mathbf{u} \times \mathbf{v}$, define the matrix

$$\mathbf{U} = \begin{pmatrix} 0 & -\gamma & \beta \\ \gamma & 0 & -\alpha \\ -\beta & \alpha & 0 \end{pmatrix} \tag{3.2}$$

Figure 3.2: The magnitude of the cross product is equal to the area of the parallelogram formed by the two vectors. The areas of the two gray triangles are equal, hence the area of the parallelogram is equal to the area of the rectangle of width $\|\mathbf{v}\|$ and height $\|\mathbf{w}\| \sin\theta$.




The product of a square matrix $\mathbf{M}$ and a column vector $\mathbf{v}$ is a column vector $\mathbf{u} = \mathbf{M}\mathbf{v}$ that is defined as follows:

$$\begin{aligned}
\mathbf{M}\mathbf{v} &= \begin{pmatrix} m_{1,1} & m_{1,2} & m_{1,3} \\ m_{2,1} & m_{2,2} & m_{2,3} \\ m_{3,1} & m_{3,2} & m_{3,3} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \\
&= \begin{pmatrix} (m_{1,1} \quad m_{1,2} \quad m_{1,3}) \cdot (v_1, v_2, v_3) \\ (m_{2,1} \quad m_{2,2} \quad m_{2,3}) \cdot (v_1, v_2, v_3) \\ (m_{3,1} \quad m_{3,2} \quad m_{3,3}) \cdot (v_1, v_2, v_3) \end{pmatrix} \\
&= \begin{pmatrix} m_{1,1}v_1 + m_{1,2}v_2 + m_{1,3}v_3 \\ m_{2,1}v_1 + m_{2,2}v_2 + m_{2,3}v_3 \\ m_{3,1}v_1 + m_{3,2}v_2 + m_{3,3}v_3 \end{pmatrix}
\end{aligned}$$

Then the cross product is equivalent to the matrix product:

$$\mathbf{u} \times \mathbf{v} = \mathbf{U}\mathbf{v} = \begin{pmatrix} 0 & -\gamma & \beta \\ \gamma & 0 & -\alpha \\ -\beta & \alpha & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} c\beta - b\gamma \\ a\gamma - c\alpha \\ b\alpha - a\beta \end{pmatrix} \tag{3.3}$$

where now we are using column matrices to denote the vectors $\mathbf{v}$ and $\mathbf{u} \times \mathbf{v}$. The use of column vectors and matrices is particularly useful if we are computing a long string of vector calculations, such as $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$.

**Example 3.1** *Find* $\mathbf{v} \times \mathbf{w}$*, where* $\mathbf{v} = 2\mathbf{i} + \mathbf{j} - 2\mathbf{k}$ *and* $\mathbf{w} = 3\mathbf{i} + \mathbf{k}$*.*

*Solution.* From equation (3.3),

$$\mathbf{v} \times \mathbf{w} = \begin{pmatrix} 0 & 2 & 1 \\ -2 & 0 & -2 \\ -1 & 2 & 0 \end{pmatrix} \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -8 \\ -3 \end{pmatrix}$$

In our "standard" form, then, $\mathbf{v} \times \mathbf{w} = (1, -8, -3)$.■

*Proof of equation 3.3.*
To show that $\mathbf{Uv}$ is perpendicular to both $\mathbf{v}$ and $\mathbf{u}$ we compute the dot products:

$$\begin{aligned}
\mathbf{u}\cdot(\mathbf{Uv}) &= \mathbf{u}^T\mathbf{Uv} = \begin{pmatrix}\alpha & \beta & \gamma\end{pmatrix}\begin{pmatrix}c\beta - b\gamma \\ a\gamma - c\alpha \\ b\alpha - a\beta\end{pmatrix} \\
&= \alpha(c\beta - b\gamma) + \beta a\gamma - c\alpha) + \gamma(b\alpha - a\beta) = 0
\end{aligned}$$

$$\begin{aligned}
\mathbf{v}\cdot(\mathbf{Uv}) &= \mathbf{v}^T\mathbf{Uv} = \begin{pmatrix}a & b & c\end{pmatrix}\begin{pmatrix}c\beta - b\gamma \\ a\gamma - c\alpha \\ b\alpha - a\beta\end{pmatrix} \\
&= a(c\beta - b\gamma) + b(a\gamma - c\alpha) + c(b\alpha - a\beta) = 0
\end{aligned}$$

To show that the product given by equation (3.3) has length $\|\mathbf{u}\|\|\mathbf{v}\|\sin\theta$, observe that

$$\begin{aligned}
\|\mathbf{u}\|^2 &= (\|\mathbf{u}\|\sin\theta)^2 + (\|\mathbf{u}\|\cos\theta)^2 \\
\|\mathbf{u}\|^2\|\mathbf{v}\|^2 &= (\|\mathbf{u}\|\|\mathbf{v}\|\sin\theta)^2 + (\|\mathbf{u}\|\|\mathbf{v}\|\cos\theta)^2 \\
&= (\|\mathbf{u}\|\|\mathbf{v}\|\sin\theta)^2 + (\mathbf{u}\cdot\mathbf{v})^2 \\
(\|\mathbf{u}\|\|\mathbf{v}\|\sin\theta)^2 &= \|\mathbf{u}\|^2\|\mathbf{v}\|^2 - (\mathbf{u}\cdot\mathbf{v})^2 \\
&= \left(\alpha^2 + \beta^2 + \gamma^2\right)\left(a^2 + b^2 + c^2\right) - (\alpha a + \beta b + \gamma c)^2 \\
&= \alpha^2\left(a^2 + b^2 + c^2\right) + \beta^2\left(a^2 + b^2 + c^2\right) + \gamma^2\left(a^2 + b^2 + c^2\right) \\
&\quad -\alpha a\,(\alpha a + \beta b + \gamma c) - \beta b\,(\alpha a + \beta b + \gamma c) - \gamma c\,(\alpha a + \beta b + \gamma c) \\
&= \alpha^2 b^2 + \alpha^2 c^2 + \beta^2 a^2 + \beta^2 c^2 + \gamma^2 a^2 + \gamma^2 b^2 - 2\alpha\beta ab - 2\alpha\gamma ac - 2\beta\gamma bc
\end{aligned}$$

Now consider the magnitude $\|\mathbf{Uv}\|$,

$$\begin{aligned}
\|\mathbf{Uv}\|^2 &= \left\|\begin{pmatrix}c\beta - b\gamma \\ a\gamma - c\alpha \\ b\alpha - a\beta\end{pmatrix}\right\|^2 \\
&= (c\beta - b\gamma)^2 + (a\gamma - c\alpha\ )^2 + (b\alpha - a\beta)^2 \\
&= c^2\beta^2 - 2bc\beta\gamma + b^2\gamma^2 + a^2\gamma^2 - 2ac\alpha\gamma + c^2\alpha^2 + b^2\alpha^2 - 2ab\alpha\beta + a^2\beta^2 \\
&= (\|\mathbf{u}\|\|\mathbf{v}\|\sin\theta)^2
\end{aligned}$$

and therefore

$$\|\mathbf{Uv}\| = \|\mathbf{u}\|\|\mathbf{v}\|\sin\theta = \|\mathbf{u}\times\mathbf{v}\|$$

Hence the cross product's components are given by equation (3.3). ■

**Definition 3.2 (Determinant of a Square Matrix.)** *Let*

$$\mathbf{M} = \begin{pmatrix}a & b \\ c & d\end{pmatrix}$$

*be a* $2 \times 2$ *square matrix. Then its* **determinant** *is given by*

$$\det \mathbf{M} = |\mathbf{M}| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

*If* $\mathbf{M}$ *is a* $3 \times 3$ *matrix, then*

$$\det \mathbf{M} = |\mathbf{M}| = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix}$$

*and hence*

$$\det \mathbf{M} = a(ei - fh) - b(di - fg) + c(dh - eg) \tag{3.4}$$

From equation (3.4) we can calculate the following:

$$\det \mathbf{M} = \begin{pmatrix} a & b & c \end{pmatrix} \begin{pmatrix} ei - fh \\ fg - di \\ dh - eg \end{pmatrix} \tag{3.5}$$

Let $\mathbf{w} = (A, B, C)$, $\mathbf{u} = (\alpha, \beta, \gamma)$, and $\mathbf{v} = (a, b, c)$, as before. Then comparing equations (3.3) and (3.5)

$$\mathbf{w} \cdot (\mathbf{u} \times \mathbf{v}) = \begin{pmatrix} A & B & C \end{pmatrix} \begin{pmatrix} c\beta - b\gamma \\ a\gamma - c\alpha \\ b\alpha - a\beta \end{pmatrix} \tag{3.6}$$

$$= A(c\beta - b\gamma) + B(a\gamma - c\alpha) + C(b\alpha - a\beta) \tag{3.7}$$

$$= A \begin{vmatrix} \beta & \gamma \\ b & c \end{vmatrix} - B \begin{vmatrix} \alpha & \gamma \\ a & c \end{vmatrix} + C \begin{vmatrix} \alpha & \beta \\ a & b \end{vmatrix} \tag{3.8}$$

$$= \begin{pmatrix} A & B & C \\ \alpha & \beta & \gamma \\ a & b & c \end{pmatrix} \tag{3.9}$$

By analogy we can derive the following result.

**Theorem 3.1** *Let* $\mathbf{v} = (a, b, c)$, $\mathbf{u} = (\alpha, \beta, \gamma)$, *and let* $\mathbf{i}$, $\mathbf{j}$, *and* $\mathbf{k}$ *be the usual basis vectors. Then*

$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \alpha & \beta & \gamma \\ a & b & c \end{vmatrix} \tag{3.10}$$

*Proof.*

$$\begin{aligned} \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \alpha & \beta & \gamma \\ a & b & c \end{vmatrix} &= \mathbf{i} \begin{vmatrix} \beta & \gamma \\ b & c \end{vmatrix} - \mathbf{j} \begin{vmatrix} \alpha & \gamma \\ a & c \end{vmatrix} + \mathbf{k} \begin{vmatrix} \alpha & \beta \\ a & b \end{vmatrix} \\ &= \mathbf{i}(c\beta - b\gamma) - \mathbf{j}(c\alpha - a\gamma) + \mathbf{k}(b\alpha - a\beta) \\ &= \mathbf{u} \times \mathbf{v} \end{aligned}$$

where the last line follows from equation (3.3). ■

**Example 3.2** *Find* $\mathbf{v} \times \mathbf{w}$*, where* $\mathbf{v} = 2\mathbf{i} + \mathbf{j} - 2\mathbf{k}$ *and* $\mathbf{w} = 3\mathbf{i} + \mathbf{k}$*.*

*Solution.*

$$\begin{aligned}
\mathbf{v} \times \mathbf{w} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2 & 1 & -2 \\ 3 & 0 & 1 \end{vmatrix} = \mathbf{i}\begin{vmatrix} 1 & -2 \\ 0 & 1 \end{vmatrix} - \mathbf{j}\begin{vmatrix} 2 & -2 \\ 3 & 1 \end{vmatrix} + \mathbf{k}\begin{vmatrix} 2 & 1 \\ 3 & 0 \end{vmatrix} \\
&= \mathbf{i}\left((1)(1) - (-2)(0)\right) - \mathbf{j}\left((2)(1) - (3)(-2)\right) + \mathbf{k}\left((2)(0) - (1)(3)\right) \\
&= \mathbf{i} - 8\mathbf{j} - 3\mathbf{k}. \blacksquare
\end{aligned}$$

**Theorem 3.2 (Properties of the Cross Product)**. *Let* $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$ *be vectors, and let* $a$ *be any number. Then*

1. $\mathbf{w} \times \mathbf{v} = -\mathbf{v} \times \mathbf{w}$ – The cross product **anti-commutes**.
2. $(a\mathbf{v}) \times \mathbf{w} = a(\mathbf{v} \times \mathbf{w}) = \mathbf{v} \times (a\mathbf{w})$
3. $\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = \mathbf{u} \times \mathbf{v} + \mathbf{u} \times \mathbf{w}$ – The cross product is **left-distributive across vector addition**.
4. $\mathbf{u} \times \mathbf{v} = 0$ if and only if the vectors are parallel (assuming that $\mathbf{u}$, $\mathbf{v} \neq \mathbf{0}$.)
5. $\mathbf{i} \times \mathbf{j} = \mathbf{k}$, $\mathbf{j} \times \mathbf{k} = \mathbf{i}$, $\mathbf{k} \times \mathbf{i} = \mathbf{j}$
6. $\mathbf{j} \times \mathbf{i} = -\mathbf{k}$, $\mathbf{k} \times \mathbf{j} = -\mathbf{i}$, $\mathbf{i} \times \mathbf{k} = -\mathbf{j}$
7. $\mathbf{i} \times \mathbf{i} = \mathbf{j} \times \mathbf{j} = \mathbf{k} \times \mathbf{k} = 0$
8. $\mathbf{u} \cdot (\mathbf{u} \times \mathbf{v}) = \mathbf{v} \cdot (\mathbf{u} \times \mathbf{v}) = 0$
9. The vectors $\mathbf{u}, \mathbf{v}, \mathbf{u} \times \mathbf{v}$ form a **right handed triple**.
10. $(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$
11. $\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) = (\mathbf{u} \cdot \mathbf{w})\mathbf{v} - (\mathbf{u} \cdot \mathbf{v})\mathbf{w}$

**Example 3.3** *Prove that the cross product is* ***right-distributive*** *across vector additions using the properties in theorem 3.2, i.e., show that* $(\mathbf{v}+\mathbf{w})\times\mathbf{u} = \mathbf{v}\times\mathbf{u}+\mathbf{w}\times\mathbf{u}$.

*Solution.* By property (1)

$$(\mathbf{v} + \mathbf{w}) \times \mathbf{u} = -\mathbf{u} \times (\mathbf{v} + \mathbf{w})$$

By the left distributive property,

$$-\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = -\mathbf{u} \times \mathbf{v} - \mathbf{u} \times \mathbf{w}$$

Applying anti-commutivity a second time,

$$-\mathbf{u} \times \mathbf{v} - \mathbf{u} \times \mathbf{w} = \mathbf{v} \times \mathbf{u} + \mathbf{w} \times \mathbf{u}$$

Hence

$$(\mathbf{v} + \mathbf{w}) \times \mathbf{u} = \mathbf{v} \times \mathbf{u} + \mathbf{w} \times \mathbf{u}. \blacksquare$$

**Example 3.4** *Let* $\mathbf{a} = \langle 3,3,1\rangle, \mathbf{b} = \langle -2,-1,0\rangle, \mathbf{c} = \langle -2,-3,-1\rangle$. *Find* $\mathbf{a}\times(\mathbf{b}\times\mathbf{c})$

*Solution.* By property (11) of theorem 3.2,

$$\begin{aligned}
\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) &= \mathbf{b}(\mathbf{a}\cdot\mathbf{c}) - \mathbf{c}(\mathbf{a}\cdot\mathbf{b}) \\
&= (-2,-1,0)\left((3,3,1)\cdot(-2,-3,-1)\right) - (-2,-3,-1)\left((3,3,1)\cdot(-2,-1,0)\right) \\
&= -16(-2,-1,0) + 9(-2,-3,-1) \\
&= (32,16,0) + (-18,-27,-9) \\
&= (14,-11,-9)\blacksquare
\end{aligned}$$

### Finding the equation of a plane through three points P, Q, R

Suppose we are given the coordinates of three (non-collinear) points, **P**, **Q**, and **R**, and want to find the equation of the plain to contains all three points. The following procedure will give you this equation.

1. Use the points to define two vectors

$$\overrightarrow{\mathbf{PQ}} = \mathbf{Q} - \mathbf{P}$$
$$\overrightarrow{\mathbf{QR}} = \mathbf{R} - \mathbf{Q}$$

It does not matter which pair you use, as long as you take any two non-parallel vectors from among the six possible vectors $\overrightarrow{\mathbf{PQ}}$, $\overrightarrow{\mathbf{QR}}$, $\overrightarrow{\mathbf{RP}}$, $\overrightarrow{\mathbf{QP}}$, $\overrightarrow{\mathbf{RQ}}$ and $\overrightarrow{\mathbf{PR}}$. (Note that if you chose $\overrightarrow{\mathbf{PQ}}$ as your first vector, the second can be any of the other vectors except for $\overrightarrow{\mathbf{QP}}$.

2. The vectors you chose in step (1) define the plane. Both vectors lie in the plane that contains the three points. Calculate the cross product

$$\mathbf{n} = \overrightarrow{\mathbf{PQ}} \times \overrightarrow{\mathbf{QR}}$$

Since **n** is perpendicular to both vectors, it must be normal to the plane they are contained in.

3. Pick any one of the three points **P**, **Q**, **R**, say $\mathbf{P} = (p_x, p_y, p_z)$, and let $\mathbf{X} = (x, y, z)$ be *any point in the plane.* Then the vector

$$\overrightarrow{\mathbf{PX}} = \mathbf{X} - \mathbf{P} = (x - p_x, y - p_y, z - p_z)$$

lies in the plane.

4. Since $\overrightarrow{\mathbf{PX}}$ lies in the plane, and the vector **n** is normal to the plane, they must be perpendicular to one another, i.e.,

$$\mathbf{n}\cdot\overrightarrow{\mathbf{PX}} = 0$$

5. In fact, every point $\mathbf{X}$ in the plane can be represented by some vector $\overrightarrow{\mathbf{PX}}$ pointing from $\mathbf{P}$ to $\mathbf{X}$, and hence we can say the plane is the locus of all points satisfying $\mathbf{n} \cdot \overrightarrow{\mathbf{PX}} = 0$ and hence
$$\mathbf{n} \cdot (x - p_x, y - p_y, z - p_z) = 0$$
or equivalently,
$$\mathbf{n} \cdot (x, y, z) = \mathbf{n} \cdot \mathbf{P}$$
since $\mathbf{P} = (p_x, p_y, p_z)$.

It is usually easier to reconstruct this procedure each time than it is to memorize a set of formulas, as in the following example.

**Example 3.5** *Find the equation of the plane that contains the three points* $\mathbf{P} = (1, 3, 0)$, $\mathbf{Q} = (3, 4, -3)$ *and* $\mathbf{R} = (3, 6, 2)$.

*Solution.* We start by finding two vectors in the plane, for example,

$$\overrightarrow{\mathbf{QP}} = \mathbf{P} - \mathbf{Q} = (1, 3, 0) - (3, 4, -3) = (-2, -1, 3)$$
$$\overrightarrow{\mathbf{QR}} = \mathbf{R} - \mathbf{Q} = (3, 6, 2) - (3, 4, -3) = (0, 2, 5)$$

Then a normal vector to the plane is given by

$$\begin{aligned}
\mathbf{n} &= \overrightarrow{\mathbf{QP}} \times \overrightarrow{\mathbf{QR}} \\
&= (-2, -1, -3) \times (0, 2, 5) \\
&= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -2 & -1 & -3 \\ 0 & 2 & 5 \end{vmatrix} \\
&= \mathbf{i} \begin{vmatrix} 3 & -1 \\ -5 & 6 \end{vmatrix} - \mathbf{j} \begin{vmatrix} -4 & -1 \\ 2 & 6 \end{vmatrix} + \mathbf{k} \begin{vmatrix} 4 & 3 \\ 2 & -5 \end{vmatrix} \\
&= -11\mathbf{i} + 10\mathbf{j} - 4\mathbf{k} \\
&= (-11, 10, -4)
\end{aligned}$$

Then for any point $(x, y, z)$ in the plane, the vector from $\mathbf{R}$ (where $\mathbf{R}$ was chosen completely randomly from the three points) to $(x, y, z)$ is given by

$$\mathbf{v} = (x - 3, y - 6, z - 2)$$

Since $\mathbf{v}$ and $\mathbf{n}$ must be perpendicular, then their dot product is zero:

$$\begin{aligned}
0 &= \mathbf{n} \cdot \mathbf{v} = (-11, 10, -4) \cdot (x - 3, y - 6, z - 2) \\
&= -11(x - 3) + 10(y - 6) - 4(z - 2) \\
&= -11x + 33 + 10y - 60 - 4z + 8 \\
&= -11x + 10y - 4z - 19
\end{aligned}$$

Hence the equation of the plane that contains the three points is

$$11x - 10y + 4z = -19 \blacksquare$$

**Example 3.6** *Find the equation of the plane through* $(2,-3,2)$ *and parallel to the plane containing the vectors* $\mathbf{v} = 4\mathbf{i} + 3\mathbf{j} - \mathbf{k}, \mathbf{w} = 2\mathbf{i} - 5\mathbf{j} + 6\mathbf{k}$.

*Solution.*The normal to the desired plane will also be perpendicular to the plane containing **v** and **w**. Since the cross product of any two vectors is by definition perpendicular to the plane containing both vectors, then one such normal vector is

$$\begin{aligned}
\mathbf{n} &= \mathbf{v} \times \mathbf{w} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 4 & 3 & -1 \\ 2 & -5 & 6 \end{vmatrix} \\
&= \mathbf{i} \begin{vmatrix} 3 & -1 \\ -5 & 6 \end{vmatrix} - \mathbf{j} \begin{vmatrix} 4 & -1 \\ 2 & 6 \end{vmatrix} + \mathbf{k} \begin{vmatrix} 4 & 3 \\ 2 & -5 \end{vmatrix} \\
&= 13\mathbf{i} - 26\mathbf{j} - 26\mathbf{k}
\end{aligned}$$

Since a point on the plane is (2,-3,2), the equation of the plane is

$$\mathbf{n} \cdot (x, y, z) = \mathbf{n} \cdot \mathbf{P}$$
$$(13, -26, -26) \cdot (x, y, z) = (13, -26, -26) \cdot (2, -3, 2)$$
$$13x - 26y - 26z = 52$$

Dividing the last equation through by 13 gives $x - 2y - 2z = 4$ as the equation of the desired plane. ■

# Lecture 4

# Lines and Curves in 3D

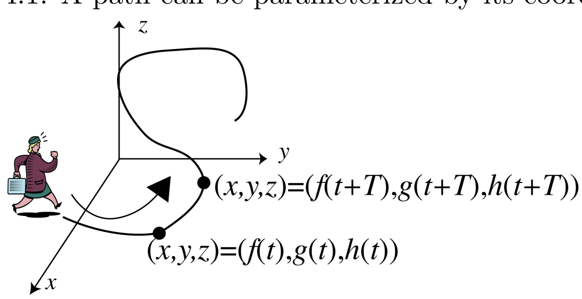
We can parameterize a curve in three dimensions in the same way that we did in two dimensions, associating a function with each of the three coordinates,

$$x = f(t), y = g(t), z = h(t) \tag{4.1}$$

where $t$ is allowed to vary over some interval that we will call $I$. It is useful to envision yourself as moving along the curve from one end to the other. At any time $t$, you are at a point $(x, y, z)$ on the curve. We think of the parameter $t$ as time and then

$x = f(t)$ is your x-coordinate at time t
$y = g(t)$is your y-coordinate at time t
$z = h(t)$ is your z-coordinate at time t.

Figure 4.1: A path can be parameterized by its coordinates.

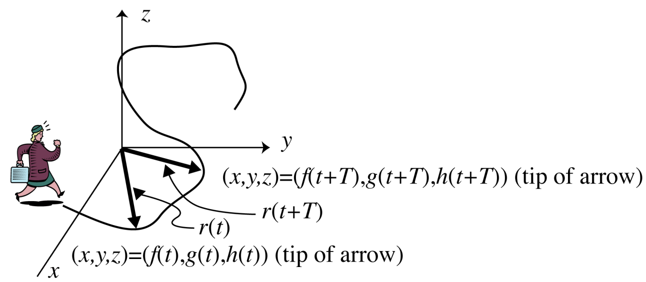



We define our **position vector** $\mathbf{r}(t)$ at any time $t$ as the vector pointing from the origin to our position $(x, y, z)$ at the time $t$. This vector is then a function of $t$ and is given by

$$\mathbf{r}(t) = (x(t), y(t), z(t)) = f(t)\mathbf{i} + g(t)\mathbf{j} + h(t)\mathbf{k} \tag{4.2}$$

Suppose that we walk along a straight line, as illustrated in figure 4.3, starting at the point $\mathbf{P}_0$, at a time $t = 0$, and arrive at the point $\mathbf{P}$ at time $t$, moving with a

Figure 4.2: The vector $\mathbf{r}(t) = (f(t), g(t), h(t))$ describes the position as a function of time.




constant speed $v$, the whole time, then the distance between $\mathbf{P}_0$ and $\mathbf{P}$ is the the length of the vector $\overrightarrow{\mathbf{P}_0\mathbf{P}}$. Since speed is *distance/time*,

$$v = \frac{\left\|\overrightarrow{\mathbf{P}_0\mathbf{P}}\right\|}{t} \tag{4.3}$$

or

$$\left\|\overrightarrow{\mathbf{P}_0\mathbf{P}}\right\| = vt \tag{4.4}$$

If we define the **veclocity vector** as a vector of length $v$ pointing in our direction of motion,

$$\mathbf{v} = v\hat{\mathbf{v}} \tag{4.5}$$

then it must be true that $\mathbf{v}$ is parallel to $\overrightarrow{\mathbf{P}_0\mathbf{P}}$, so that

$$\mathbf{P} - \mathbf{P}_0 = \overrightarrow{\mathbf{P}_0\mathbf{P}} = (v\hat{\mathbf{v}})\,t \tag{4.6}$$

Since the coordinates of our position vector $\mathbf{r}$, are, by definition, the same as the coordinates of the point $\mathbf{P}$ (see equation (4.2)),

$$\boxed{\mathbf{r} = \mathbf{r}_0 + \mathbf{v}t} \tag{4.7}$$

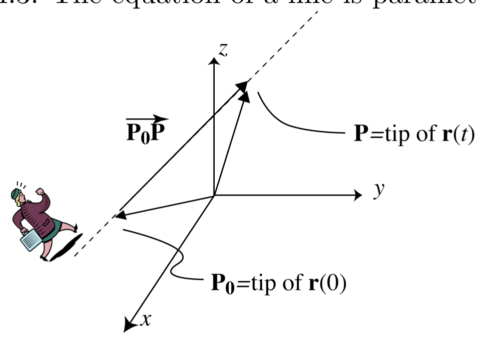
where we have defined $\mathbf{r}_0$ to be our position vector at time $t = 0$. Equation (4.7) is the equation of a line through $\mathbf{r}_0$ in the direction $\mathbf{v}$.

**Theorem 4.1** *The equation of a line through the point* $\mathbf{r}_0$ *and parallel to the vector* $\mathbf{v} \neq \mathbf{0}$ *is given by* $\mathbf{r} = \mathbf{r}_0 + \mathbf{v}t$*. If* $\mathbf{u} \neq \mathbf{0}$ *is* **any vector parallel (or anti-parallel) to v** *then* $\mathbf{r} = \mathbf{r}_0 + \mathbf{u}t$ *gives an equation for the same line.*

If we denote our coordinates at any time $t$ by $(x, y, z)$, and our velocity vector by $(v_x, v_y, v_z)$, then

$$\begin{aligned} (x, y, z) &= (x_0, y_0, z_0) + (v_x, v_y, v_z)t \\ &= (x_0, y_0, z_0) + (v_xt, v_yt, v_zt) \\ &= (x_0 + v_xt, y_0 + v_yt, z_0 + v_zt) \end{aligned} \tag{4.8}$$

Figure 4.3: The equation of a line is parametrized by.




If two vectors are equal then each of their components must be equal,

$$\boxed{x = x_0 + v_x t, y = y_0 + v_y t, z = z_0 + v_z t} \tag{4.9}$$

Equation (4.9) gives the **parametric equation of a line.** The numbers $v_x$, $v_y$, and $v_z$ are called the **direction numbers** of the line.[1]

**Example 4.1** *Find the parametric equation of a line through the points* $(2, -1, 5)$ *and* $(7, -2, 3)$*.*

*Solution.* We can define

$$\begin{aligned} \mathbf{r}_0 &= (2, -1, 5) \\ \mathbf{r}_1 &= (7, -2, 3) \end{aligned}$$

so that

$$\mathbf{v} = \mathbf{r}_1 - \mathbf{r}_0 = (7, -2, 3) - (2, -1, 5) = (5, -1, -2)$$

is a vector parallel to the line. Hence the equation of the line is

$$\mathbf{r} = (2, -1, 5) + (5, -1, -2)t = (2 + 5t, -1 - t, 5 - 2t)$$

The parametric equations of the line are

$$x = 2 + 5t, y = -1 - t, z = 5 - 2t. \blacksquare$$

If $v_x \neq 0, v_y \neq 0$, and $v_z \neq 0$, then we can solve for the variable $t$ in each of equations (4.9),

$$t = (x - x_0)/v_x \tag{4.10}$$
$$t = (y - y_0)/v_y \tag{4.11}$$
$$t = (z - z_0)/v_z \tag{4.12}$$

[1]The term "direction numbers" is rarely used and should probably be avoided, even though it is mentioned prominently in the text.

Since $t$ must be the same in all three equations, we obtain the **symmetric equations of a line**,

$$\boxed{\frac{x - x_0}{v_x} = \frac{y - y_0}{v_y} = \frac{z - z_0}{v_z}} \tag{4.13}$$

**Example 4.2** *Find the symmetric equations of the line we found in example 4.1.*

*Solution.* In the earlier example we calculated that

$$\mathbf{v} = (v_x, v_y, v_z) = (5, -1, -2)$$

is parallel to the line, and that the line goes through the point

$$(x_0, y_0, z_0) = (2, -1, 5)$$

Hence

$$\begin{array}{ccc} x_0 = 2, & y_0 = -1, & z_0 = 5 \\ v_x = 5, & v_y = -1, & v_z = -2 \end{array}$$

and thus the symmetric equations of the line are

$$\frac{x - 2}{5} = \frac{y + 1}{-1} = \frac{z - 5}{-2} \blacksquare$$

**Example 4.3** *Find the equation of the plane that contains the two lines*

$$x = -2 + 2t, y = 1 + 4t, z = 2 + t$$

*and*

$$x = 2 + 4t, y = 3 + 2t, z = 1 - t$$

*Solution.* The equation of a plane through a point $\mathbf{P}$ that has a normal vector $\mathbf{n}$ is

$$\mathbf{n} \cdot (x, y, z) = \mathbf{P} \cdot \mathbf{n}$$

Thus we need to find (a) a point on the plane; and (b) a vector that is perpendicular to the plane.

One way to find the normal vector is to find two non-parallel vectors in the plane and take their cross product. But we can read off two such vectors from the equation of the line:

$$\mathbf{u} = (2, 4, 1)$$

$$\mathbf{v} = (4, 2, -1)$$

Hence a normal vectors is

$$\mathbf{n} = \mathbf{u} \times \mathbf{v} = \begin{pmatrix} 0 & -1 & 4 \\ 1 & 0 & -2 \\ -4 & 2 & 0 \end{pmatrix} \begin{pmatrix} 4 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} -6 \\ 6 \\ -12 \end{pmatrix}$$

Note: alternatively, we could have calculated

$$\mathbf{n} = \mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2 & 4 & 1 \\ 4 & 2 & -1 \end{vmatrix} = -6\mathbf{i} + 6\mathbf{j} - 12\mathbf{k}$$

We also need to find a point on the plane. We do this by picking any time, say $t = 0$, and substituting that into either of the lines. The first line, for example, gives $\mathbf{P} = (-2, 1, 2)$ at $t = 0$. The equation of the plane through $\mathbf{P}$ with normal vector $\mathbf{n}$ is then

$$(x, y, z) \cdot (-6, 6, -12) = (-2, 1, 2) \cdot (-6, 6, -12)$$

or

$$-6x + 6y + 12z = 12 + 6 - 24 = -6$$

Suppose we had picked a different point $\mathbf{P}$, say, for example, the second line at $t = 2$, $\mathbf{P} = (10, 7, -1)$. Then we would calculate

$$(x, y, z) \cdot (-6, 6, -12) = (-6, 6, -12) \cdot (10, 7, -1)$$

$$-6x + 6y + 12z = -60 + 42 + 12 = -6$$

In other words ***we get the same equation regardless of which point we pick on the plane***. ■

**Example 4.4** *Find the equation of the plane containing the line*

$$x = 1 + 2t, y = -1 + 3t, z = 4 + t$$

*and the point* $\mathbf{P} = (1, -1, 5)$ (text section 14.4, exercise 19).

*Solution.* We already know one point on the plane; we also need a normal vector to get the equation of the plane. We can get a second point on the plane from the equation of a line; at $t = 0$ we have

$$\mathbf{Q} = (x(0), y(0), z(0)) = (1, -1, 4)$$

The vector

$$\mathbf{u} = \overrightarrow{\mathbf{QP}} = \mathbf{P} - \mathbf{Q} = (1, -1, 5) - (1, -1, 4) = (0, 0, 1)$$

We can get a second vector that lies in the plane from the equation of the line, $\mathbf{v} = (2, 3, 1)$ which is the velocity vector of the line. Hence a normal vector to the plane is

$$\mathbf{u} = \mathbf{v} \times \mathbf{v} = (0, 0, 1) \times (2, 3, 1) = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} = \begin{pmatrix} -3 \\ 2 \\ 0 \end{pmatrix}$$

The equation of the plane is

$$(x, y, z) \cdot (-3, 2, 0) = (1, -1, 5) \cdot (-3, 2, 0)$$

$$-3x + 2y = -5. ■$$

**Example 4.5** *Find the equation of the plane containing the line*

$$x = 3t, y = 1 + t, z = 2t$$

*and parallel to the intersection of the planes*

$$2x - y + z = 0, y + z + 1 = 0$$

(text section 14.4, exercise 20)

*Solution.* As usual, we need a point **P** in the plane, and a normal vector **n** to the plane. The normal vectors of the two planes are

$$\mathbf{u} = (2, -1, 1), \mathbf{v} = (0, 1, 1)$$

hence

$$\mathbf{w} = \mathbf{u} \times \mathbf{v} = \begin{pmatrix} 0 & -1 & -1 \\ 1 & 0 & -2 \\ 1 & 2 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ -2 \\ 2 \end{pmatrix}$$

must be parallel to both planes because it is perpendicular to both normal vectors. Since **w** is parallel to both planes it must be parallel to their intersection.

A second vector in the solution plane is direction vector of the line, $\mathbf{r} = (3, 1, 2)$.

A normal vector to the solution plane is the cross product

$$\begin{aligned} \mathbf{n} &= \mathbf{r} \times \mathbf{w} = (3, 1, 2) \times (-2, -2, 2) \\ &= \begin{pmatrix} 0 & -2 & 1 \\ 2 & 0 & -3 \\ -1 & 3 & 0 \end{pmatrix} \begin{pmatrix} -2 \\ -2 \\ 2 \end{pmatrix} = \begin{pmatrix} 6 \\ -10 \\ -4 \end{pmatrix} \end{aligned}$$

Taking the equation of the given line at $t = 0$ we see that the solution plane pass through the point $(0, 1, 0)$ and has normal vector $(6, -3, 2)$. Hence the equation of the solution plane is

$$(x, y, z) \cdot (6, -10, -4) = (0, 1, 0) \cdot (6, -10, 4)$$

$$6x - 10y - 4z = -10$$

$$3x - 5y - 2z = -5 \blacksquare$$

## Tangent Line to a Parametrized Curve in 3-Space

Suppose we know the parameterization of some curve

$$\mathbf{r}(t) = (f(t), g(t), h(t)) \tag{4.14}$$

and would like to find a line that is tangent to this curve at $t$. The derivative of our position is

$$\begin{aligned}
\mathbf{r}'(t) &= \frac{d\mathbf{r}(t)}{dt} = \lim_{\epsilon\to 0} \frac{\mathbf{r}(t+\epsilon) - \mathbf{r}(t)}{\epsilon} && (4.15)\\
&= \lim_{\epsilon\to 0} \frac{(f(t+\epsilon), g(t+\epsilon), h(t+\epsilon)) - (f(t), g(t), h(t))}{\epsilon} && (4.16)\\
&= \lim_{\epsilon\to 0} \left( \frac{f(t+\epsilon) - f(t)}{\epsilon}, \frac{g(t+\epsilon) - g(t)}{\epsilon}, \frac{h(t+\epsilon) - h(t)}{\epsilon} \right) && (4.17)\\
&= \left( \lim_{\epsilon\to 0} \frac{f(t+\epsilon) - f(t)}{\epsilon}, \lim_{\epsilon\to 0} \frac{g(t+\epsilon) - g(t)}{\epsilon}, \lim_{\epsilon\to 0} \frac{h(t+\epsilon) - h(t)}{\epsilon} \right) && (4.18)\\
&= \left(f'(t), g'(t), h'(t)\right) && (4.19)\\
& && (4.20)
\end{aligned}$$

In other words, you can differentiate term by term.

The derivative $\mathbf{r}'(t)$ gives a vector that is tangent to the curve $\mathbf{r}(t)$,

**Example 4.6** *Find the symmetric equations of the tangent line to the curve parameterized by*

$$\mathbf{r}(t) = \left(3t, -4t^2, \frac{1}{\pi}\sin \pi t\right)$$

*at* $t = 1$.

*Solution.* A tangent vector to the curve at any time $t$ is given by the derivative,

$$\mathbf{r}'(t) = (3, -8t, \cos \pi t)$$

At $t = 1$ we have the tangent vector

$$\mathbf{r}'(1) = (3, -8, -1)$$

and a point on the curve

$$\mathbf{r}(1) = (3, -4, 0)$$

The symmetric equations of the line are then

$$\frac{x-3}{3} = -\frac{y+4}{8} = -z$$

# Lecture 5

# Velocity, Acceleration, and Curvature

## Length of a Curve

Our definition of the length of a curve follows our intuition. Lay a string along the path of the curve, pick up the string, straighten it out, and measure the length of the line.

**Theorem 5.1** *Suppose that a curve is parameterized as*

$$\mathbf{r}(t) = (f(t), g(t), h(t)) \tag{5.1}$$

*on some interval at* $< b$. *Then the* **arc length from *a* to *b*** *or the* **length of the curve from *a* to *b*** *is given by the integral*

$$\boxed{s = \int_a^b \left|\mathbf{r}'(t)\right| dt = \int_a^b \sqrt{(f'(t))^2 + (g'(t))^2 + (h'(t))^2} dt} \tag{5.2}$$

*Proof.* Let $n$ be some large integer and define $\epsilon = (b-a)/n$. Divide $[a, b]$ into $n$ intervals

$$[a, a+\epsilon], [a+\epsilon, a+2\epsilon], ..., [b-\epsilon, b]$$

and approximate the curve by straight line segments

$$\overline{\mathbf{r}(a)\mathbf{r}(a+\epsilon)}, \overline{\mathbf{r}(a+\epsilon)\mathbf{r}(a+2\epsilon)}, ..., \overline{\mathbf{r}(b-2\epsilon)\mathbf{r}(b-\epsilon)}, \overline{\mathbf{r}(b-\epsilon)\mathbf{r}(a-\epsilon)},$$

The end points of the path during the time interval from $t \in [a+i\epsilon, a+(i+1)\epsilon]$ are given by

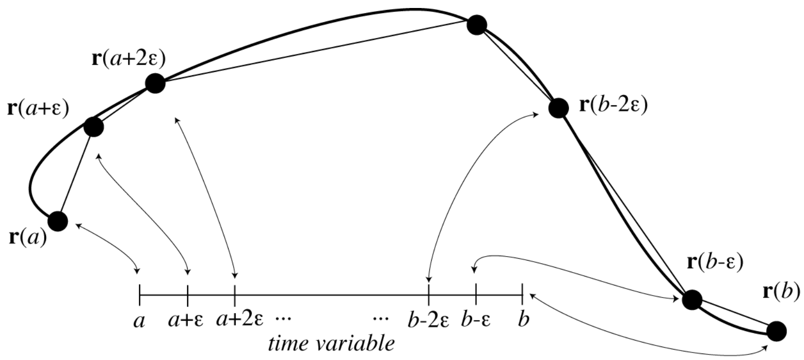
$$\mathbf{r}(a+i\epsilon) = (f(a+i\epsilon), g(a+i\epsilon), h(a+i\epsilon))$$

and

$$\mathbf{r}(a+(i+1)\epsilon) = (f(a+(i+1)\epsilon), g(a+(i+1)\epsilon), h(a+(i+1)\epsilon))$$

Hence the length of the line segment from $\mathbf{r}(a+i\epsilon)$ to $\mathbf{r}(a+(i+1)\epsilon)$

Figure 5.1: The arc length can be determined by approximating the arc by line segments, summing the length of all the line segments, and taking the limit as the number of segments becomes infinite.




$$s_i = \sqrt{[\Delta f_i]^2 + [\Delta g_i]^2 + [\Delta h_i]^2}$$

where

$$\Delta f_i = f(a + (i+1)\epsilon) - f(a + i\epsilon) \tag{5.3}$$
$$\Delta g_i = g(a + (i+1)\epsilon) - g(a + i\epsilon) \tag{5.4}$$
$$\Delta h_i = h(a + (i+1)\epsilon) - h(a + i\epsilon) \tag{5.5}$$

The total length of all of the line segments added together is then

$$s = \sum_{i=1}^{n} s_i = \sum_{i=1}^{n} \sqrt{[\Delta f_i]^2 + [\Delta g_i]^2 + [\Delta h_i]^2}$$

If we let $\Delta t = \epsilon$ and take the limit as $n \to \infty$, or equivalently, as $\epsilon \to 0$,

$$s = \lim_{\epsilon\to 0} \sum_{i=1}^{n} \frac{\Delta t}{\Delta t} \sqrt{\Delta f_i^2 + \Delta g_i^2 + \Delta h_i^2} \tag{5.6}$$
$$= \lim_{\epsilon\to 0} \sum_{i=1}^{n} \Delta t \sqrt{\left(\frac{\Delta f_i}{\Delta t}\right)^2 + \left(\frac{\Delta g_i}{\Delta t}\right)^2 + \left(\frac{\Delta h_i}{\Delta t}\right)^2} \tag{5.7}$$
$$= \int_a^b \sqrt{(f'(t))^2 + (g'(t))^2 + (h'(t))^2} dt. \blacksquare \tag{5.8}$$

**Definition 5.1** *Suppose that a curve is parameterized as in equation* (5.1). *Then we say that the curve is* **smooth** *if* $\mathbf{r}'(t)$ *exists and is continuous on* $[a, b]$ *and* $\mathbf{r}'(t) \neq 0$ *for all* $x$ *in* $[a, b]$.

**Definition 5.2** *Let* $\mathbf{r}(t) = (f(t), g(t), h(t))$. *Then if the derivatives* $f'(t)$, $g'(t)$ *and* $h'(t)$ *exist, we define the* **velocity** *by*

$$\boxed{\mathbf{v}(t) = \frac{d\mathbf{r}}{dt} = \mathbf{r}'(t) = \big(f'(t), g'(t), h'(t)\big)} \tag{5.9}$$

*and the* **speed** *as*

$$\boxed{\frac{ds}{dt} = \big\|\mathbf{r}'(t)\big\| = \|\mathbf{v}(t)\|} \tag{5.10}$$

*Furthermore, if the second derivatives* $f''(t)$, $g''(t)$ *and* $h''(t)$ *also exist we define the* **acceleration** *as*

$$\boxed{\mathbf{a}(t) = \frac{d\mathbf{v}(t)}{dt} = \mathbf{r}''(t) = \big(f''(t), g''(t), h''(t)\big)} \tag{5.11}$$

**Example 5.1** *Find the velocity and acceleration of the motion described by the curve* $\mathbf{r} = (\cos t, \sin 3t, t)$ *at* $t = \pi/2$ .

*Solution.* The time-dependent velocity and acceleration are found by differentiation:

$$\begin{aligned} \mathbf{v}(t) &= \mathbf{r}'(t) = (-\sin t, 3\cos 3t, 1) \\ \mathbf{a}(t) &= \mathbf{r}''(t) = (-\cos t, -9\sin 3t, 0) \end{aligned}$$

Therefore at $t = \pi/2$

$$\begin{aligned} \mathbf{r}'(\pi/2) &= (-1, 0, 1) \\ \mathbf{r}''(\pi/2) &= (0, 9, 0). \blacksquare \end{aligned}$$

**Example 5.2** *Find the length of the curve parametrized by* $\mathbf{r}(t) = \big(t^3, -2t^3, 6t^3\big)$ *over the interval* $(0, 1)$.

*Solution.* We use the formula for arc length, equation (5.2) with

$$f(t) = t^3, g(t) = 2t^3, h(t) = 6t^3$$

Since

$$f'(t) = 3t^2, g'(t) = 6t^2, h'(t) = 18t^2$$

and therefore the arc length is given by

$$\begin{aligned} s &= \int_0^1 \sqrt{(f'(t))^2 + (g'(t))^2 + (h'(t))^2} \; dt \\ &= \int_0^1 \sqrt{(3t^2)^2 + (6t^2)^2 + (18t^2)^2} \; dt \\ &= \int_0^1 t^2\sqrt{9 + 36 + 324} \; dt \\ &= \sqrt{369} \int_0^1 t^2 \; dt \\ &= \frac{\sqrt{369}}{3} t^3 \bigg|_0^1 = \sqrt{41} \quad \blacksquare \end{aligned}$$

**Example 5.3** *Find the length of the curve* $\mathbf{r}(t) = (t\cos t, t\sin t, t)$ *over the interval* $(3, 4)$.

*Solution.* Differentiating,

$$\mathbf{r}'(t) = (-t\sin t + \cos t, t\cos t + \sin t, 1)$$

hence

$$\begin{aligned}
s &= \int_3^4 \sqrt{(-t\sin t + \cos t)^2 + (t\cos t + \sin t)^2 + (1)^2}\, dt \\
&= \int_3^4 \sqrt{t^2\sin^2 t - 2t\sin t\cos t + \cos^2 t + t^2\cos^2 t + 2\sin t\cos t + \sin^2 t + 1}\, dt \\
&= \int_3^4 \sqrt{t^2\sin^2 t + \cos^2 t + t^2\cos^2 t + \sin^2 t + 1}\, dt \\
&= \int_3^4 \sqrt{t^2(\sin^2 t + \cos^2 t) + (\cos^2 t + \sin^2 t) + 1}\, dt \\
&= \int_3^4 \sqrt{t^2 + 1 + 1}\, dt = \int_3^4 \sqrt{t^2 + 2} dt
\end{aligned}$$

By formula 44 on the inside back cover of the text

$$\int \sqrt{x^2 \pm a^2} dx = \frac{x}{2}\sqrt{x \pm a^2} \pm \frac{a^2}{2} \ln\left|x + \sqrt{x^2 \pm a^2}\right|$$

so that

$$\begin{aligned}
s &= \int_3^4 \sqrt{t^2 + 2}\, dt = \left(\frac{t}{2}\sqrt{t^2 + 2} + \ln\left|t + \sqrt{t^2 + 2}\right|\right)\Bigg|_3^4 \\
&= \left(2\sqrt{4^2 + 2} + \ln\left|4 + \sqrt{4^2 + 2}\right|\right) - \left(\frac{3}{2}\sqrt{3^2 + 2} + \ln|3 + \sqrt{3^2 + 2}|\right) \\
&= \left(2\sqrt{18} + \ln\left|4 + \sqrt{18}\right|\right) - \left(\frac{3}{2}\sqrt{11} + \ln\left|3 + \sqrt{11}\right|\right) \\
&\approx 8.485 + 2.109 - 4.975 - 1.843 \approx 3.776. \blacksquare
\end{aligned}$$

## Tangent Vectors

Our definition of the velocity is the same as the definition of the tangent vector we in the previous section. Therefore a ***unit tangent vector*** is

$$\boxed{\hat{\mathbf{T}} = \frac{\mathbf{v}}{\|\mathbf{v}\|}} \tag{5.12}$$

**Example 5.4** *Find a unit tangent vector to the curve* $\mathbf{r}(t) = \left(\frac{1}{2}t^2, \frac{1}{3}t^3, -18t\right)$ *at* $t = 1$.

*Solution.* The velocity vector is

$$\mathbf{v}(t) = \frac{d\mathbf{r}(t)}{dt} = \left(t, t^2, -18\right)$$

At $t = 1$,

$$\mathbf{v}(1) = (1, 1, -18)$$

The magnitude of the velocity is

$$\|\mathbf{v}(1)\| = \sqrt{(1)^2 + (1)^2 + (18)^2} = \sqrt{326}$$

Therefore

$$\hat{\mathbf{T}} = \frac{1}{\sqrt{326}}(1, 1, -18)\,. \quad \blacksquare$$

**Definition 5.3** *The* **Curvature** $(t)$ *is the magnitude of the rate of change of the direction of the unit tangent vector measured with respect to distance,*

$$\boxed{\kappa(t) = \left\|\frac{d\mathbf{T}}{ds}\right\|} \tag{5.13}$$

*The* **Radius of Curvature** ***R*** *is the reciprocal of the curvature,*

$$\boxed{R = \frac{1}{\kappa}} \tag{5.14}$$

By the chain rule,

$$\kappa(t) = \left\|\frac{d\hat{\mathbf{T}}}{ds}\right\| = \left\|\frac{d\hat{\mathbf{T}}}{dt}\frac{dt}{ds}\right\| = \left\|\frac{d\hat{\mathbf{T}}}{dt}\Big/\frac{ds}{dt}\right\| = \left\|\frac{d\hat{\mathbf{T}}}{dt}\Big/\|\mathbf{v}(t)\|\right\| \tag{5.15}$$

This gives us a more useful formula for calculating the curvature directly from the parameterization:

$$\boxed{\kappa(t) = \frac{1}{\|v(t)\|}\left\|\frac{d\hat{\mathbf{T}}}{dt}\right\|} \tag{5.16}$$

**Example 5.5** *Find the curvature and radius of curvature of the curve* $\mathbf{r} = (5\cos t, 5\sin t, 6t)$.

*Solution.* The velocity is

$$\mathbf{v}(t) = (-5\sin t, 5\cos t, 6)$$

hence the speed is

$$\begin{aligned} \|\mathbf{v}(t)\| &= \sqrt{(-5\sin t)^2 + (5\cos t)^2 + 6^2} \\ &= \sqrt{25(\sin^2 t + \cos^2 t) + 36} \\ &= \sqrt{25+36} = \sqrt{61} \end{aligned}$$

The unit tangent vector is then

$$\hat{\mathbf{T}}(t) = \frac{\mathbf{v}(t)}{\|\mathbf{v}(t)\|} = \frac{1}{\sqrt{61}}(-5\sin t, 5\cos t, 6)$$

The curvature is

$$\begin{aligned} \kappa(t) &= \frac{1}{\|v(t)\|}\left\|\frac{d\mathbf{T}}{dt}\right\| \\ &= \frac{1}{61}\|(-5\cos t, -5\sin t, 0)\| \\ &= \frac{1}{61}\sqrt{(-5\cos t)^2 + (-5\sin t)^2} \\ &= \frac{1}{61}\sqrt{25(\sin^2 t + \cos^2 t)} = \frac{5}{61} \approx 0.08197 \end{aligned}$$

The corresponding radius of curvature is $R = 1/\kappa = 61/5$. ■

## The Acceleration Vector

**Theorem 5.2** *The vector* $d\mathbf{T}/ds$ *is normal to the curve.*

*Proof.* By the product rule for derivatives,

$$\frac{d}{ds}(\mathbf{T}\cdot\mathbf{T}) = \mathbf{T}\cdot\frac{d\mathbf{T}}{ds} + \left(\frac{d\mathbf{T}}{ds}\right)\cdot\mathbf{T} = 2\mathbf{T}\cdot\frac{d\mathbf{T}}{ds}$$

But since $\mathbf{T} = \mathbf{v}/\|\mathbf{v}\|$ (see equation (5.12)),

$$\mathbf{T}\cdot\mathbf{T} = \frac{\mathbf{v}\cdot\mathbf{v}}{\|\mathbf{v}\|^2} = \frac{\|\mathbf{v}\|^2}{\|\mathbf{v}\|^2} = 1$$

$$\frac{d}{ds}(\mathbf{T}\cdot\mathbf{T}) = 0$$

Hence

$$\mathbf{T}\cdot\frac{d\mathbf{T}}{ds} = 0$$

which means that $d\mathbf{T}/ds$ is perpendicular to the tangent vector $\mathbf{T}$. Thus $d\mathbf{T}/ds$ is also normal to the curve.■

**Definition 5.4** *The* **Principal unit normal vector N** *is*

$$\boxed{\mathbf{N} = \frac{d\mathbf{T}/ds}{\|d\mathbf{T}/ds\|} = \frac{1}{\kappa}\frac{d\mathbf{T}}{ds}} \tag{5.17}$$

**N** is not the only unit normal vector to the curve at **P**; in fact, on could define an infinite number of unit normal vectors to a curve at any given point. To do so, merely find the plane perpendicular to the tangent vector. An vector in this plane is normal to the curve. One such vector that is often used is the following, which is perpendicular to **both T** and **N**.

**Definition 5.5** *The* **binormal vector** *is* $\boxed{\mathbf{B} = \mathbf{T} \times \mathbf{N}}$.

**Definition 5.6** *The triple of normal vectors* $\{\mathbf{T}, \mathbf{N}, \mathbf{B}\}$ *is called the* **trihedral at P**.

**Definition 5.7** *The plane formed by* **T** *and* **N** *is called the* **osculating plane at P**.

Since $\mathbf{T} = \mathbf{v}/\|\mathbf{v}\|$ we can write

$$\mathbf{v} = \mathbf{T}\|\mathbf{v}\| = \mathbf{T}\frac{ds}{dt} \tag{5.18}$$

(see equation (5.10).) The acceleration vector (equation (5.11)) is

$$\begin{aligned} \mathbf{a} &= \frac{d\mathbf{v}}{dt} = \frac{d}{dt}\left(\mathbf{T}\frac{ds}{dt}\right) = \mathbf{T}\frac{d^2s}{dt} + \frac{d\mathbf{T}}{dt}\frac{ds}{dt} \\ &= \mathbf{T}\frac{d^2s}{dt} + \left(\frac{d\mathbf{T}}{ds}\frac{ds}{dt}\right)\frac{ds}{dt} \\ &= \mathbf{T}\frac{d^2s}{dt} + \frac{d\mathbf{T}}{ds}\left(\frac{ds}{dt}\right)^2 \end{aligned}$$

From equation (5.17),

$$\boxed{\mathbf{a} = \mathbf{T}\frac{d^2s}{dt} + \kappa\mathbf{N}\left(\frac{ds}{dt}\right)^2} \tag{5.19}$$

Equation (5.19) breaks the acceleration into two perpendicular components, one that is tangent to the curve:

$$\boxed{a_{\|} = \frac{d^2s}{dt}}$$

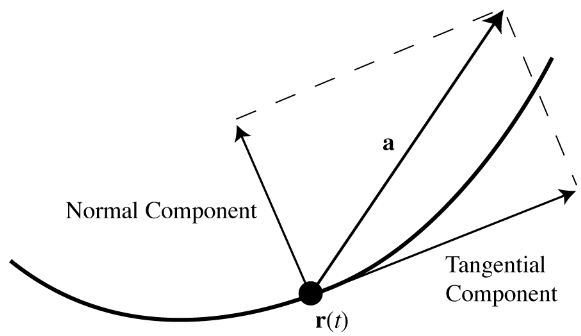
and one that is perpendicular to the curve:

$$\boxed{a_{\perp} = \kappa\left(\frac{ds}{dt}\right)^2}$$

so that

$$\boxed{\mathbf{a} = a_{\|}\mathbf{T} + a_{\perp}\mathbf{N}}$$

Figure 5.2: Components of the acceleration.




**Example 5.6** *Find the normal tangent and perpendicular components of the acceleration of the curve* $\mathbf{r}(t) = (7 + 21t, 14 - 42t, 28\sin t)$ *at* $t = \pi/3$.

*Solution.*Differentiating $\mathbf{r}$ gives the velocity

$$\mathbf{v}(t) = \mathbf{r}'(t) = (21, -42, 28\cos t) = 7(3, -6, 4\cos t)$$

and the speed as

$$\frac{ds}{dt} = \|\mathbf{v}\| = 7\sqrt{45 + 16\cos^2 t}$$

Thus the tangent component of the acceleration is

$$a_{\|}(t) = \frac{d^2 s}{dt} = 7\frac{d}{dt}\sqrt{45 + 16\cos^2 t} = -\frac{112\cos t \sin t}{\sqrt{45 + 16\cos^2 t}}$$

hence

$$a_{\|}(\pi/3) = -4\sqrt{3}$$

The acceleration vector is

$$\mathbf{a}(t) = \frac{d\mathbf{v}}{dt} = (0, 0, -28\sin t)$$

hence

$$\|\mathbf{a}(\pi/3)\| = 28\sin(\pi/3) = 14\sqrt{3}$$

Since $\|\mathbf{a}\|^2 = a_{\|}^2 + a_{\perp}^2$, the square of the normal component at $t = \pi/3$ is

$$a_{\perp}^2 = \left(14\sqrt{3}\right)^2 - \left(4\sqrt{3}\right)^2 = 540$$

and consequently

$$a_{\perp} = \sqrt{540} = 6\sqrt{15}. \quad \blacksquare$$

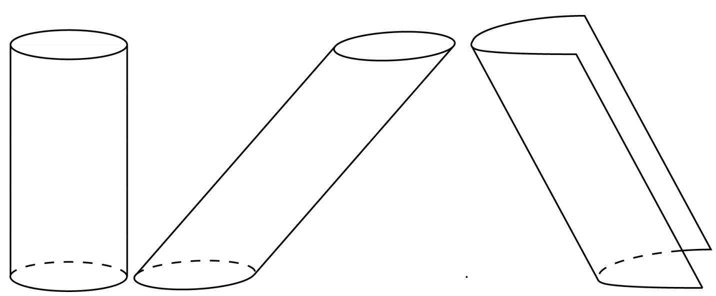
# Lecture 6

# Surfaces in 3D

The text for section 14.6 is no more than a catalog of formulas for different shapes in 3D. You should be able to recognize these shapes from their equations but you will not be expected to sketch them during an exam.

**Definition 6.1** *Let $C$ be any curve that lines in a single plane $R$, and let $L$ be any line that intersects $C$ but does not line in $R$. Then the set of all points on lines parallel to $L$ that intersect $C$ is called a* **cylinder**. *The curve $C$ is called the* **generating curve of the cylinder**.

Figure 6.1: Cylinders. Left: a right circular cylinder generated by a circle and a line perpendicular to the cylinder. Center: a circular cylinder generated by a circle and a line that is not perpendicular to the circle. Both circular cylinders extend to infinity on the top and bottom of the figure. Right: a parabolic cylinder, generated by a parabola and a line that is not in the plane of the parabola. The sheets of the parabola extend to infinity to the top, bottom, and right of the figure.



**Definition 6.2** *A simpler definition of a cylinder (actually, this is a* **right cylinder oriented parallel to one of the coordinate axes***) is to consider any curve in a plane, such as the $xy$ plane. This is an equation in two-variables, $x$ and $y$. Then consider the same formula as describing a surface in 3D. This is the cylinder generated by the curve.*

**Definition 6.3** *A* **quadric surface** *is any surface described by a second degree equation, i.e., by any equation of the form*

$$Ax^2 + By^2 + Cz^2 + Dxy + Exz + Fyz + Gx + Hy + Iz + J = 0 \tag{6.1}$$

*where* $A, B, C, D, E, F, G, H, I, J$ *are any constants and at least one of* $A, B, C, D, E, F$ *are non-zero.*

**Theorem 6.1** *Any quadric surface can be transformed, by a combination of rotation and translation, to one of the two following forms:*

$$Ax^2 + By^2 + Cz^2 + J = 0 \quad \textit{or} \tag{6.2}$$

$$Ax^2 + By^2 + Iz = 0 \tag{6.3}$$

*Quadric surfaces of the form given by equation* (6.2) *are called* **central quadrics** *because they are symmetric with respect to the coordinate planes and the origin.*

Table of Standard Quadric Surfaces.

| | |
|---|---|
| Ellipsoid | $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ |
| Hyperboloid of one Sheet | $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$ |
| Hyperboloid of two Sheets | $\frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$ |
| Elliptic Paraboloid | $z = \frac{x^2}{a^2} + \frac{y^2}{b^2}$ |
| Hyperbolic Paraboloid (saddle) | $z = \frac{y^2}{b^2} - \frac{x^2}{a^2}$ |
| Elliptic Cone | $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0$ |

# Lecture 7

# Cylindrical and Spherical Coordinates

## Cylindrical Coordinates

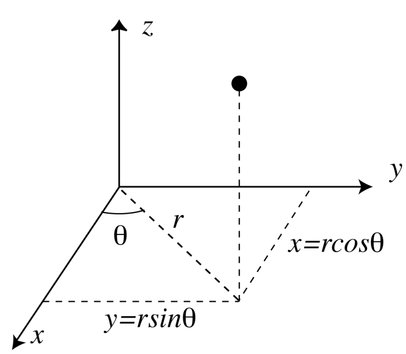
**Cylindrical coordinates** are polar coordinates with $z$ added to give the distance of a point above the $xy$ plane. A point $\mathbf{P} = (x, y, z)$ is described by the cylindrical coordinates $(r, \theta, z)$ where:

$\boldsymbol{r}$ is the distance, measured in the $xy$-plane, from the origin to the projection of $\mathbf{P}$ into the $xy$ plane.

$\theta$ is the polar angle, measured in the $xy$-plane, of the projection of $\mathbf{P}$ into the $xy$-plane.

$\boldsymbol{z}$ is the same is both cylindrical and cartesian coordinates.

Figure 7.1: Cylindrical coordinates $(r, \theta, z)$ shown in terms of the usual Cartesian coordinate frame.




**To find $(x, y, z)$ given $(\mathbf{r}, \theta, \mathbf{z})$:**

$$x = r\cos\theta$$

$$y = r\sin\theta$$
$$z = z$$

**To find** $(\mathbf{r}, \theta, \mathbf{z})$**given** $(\boldsymbol{x},\ \boldsymbol{y},\ \boldsymbol{z})$**:**

$$r = \sqrt{x^2 + y^2}$$
$$\tan\theta = y/x \quad \text{or} \quad \theta = \tan^{-1}(y/x)$$
$$z = z$$

**Example 7.1** *Convert the equation*

$$x^2 - y^2 + 2yz = 25$$

*to cylindrical coordinates*

*Solution.*We have $x = r\cos\theta$, $y = r\sin\theta$ and $z$ remains unchanged. Therefore the equation can be written as

$$25 = x^2 - y^2 + 2yz = (r\cos\theta)^2 - (r\sin\theta)^2 + 2zr\cos\theta$$

With some factoring and application of a trigonometric identity:

$$25 = r^2(\cos^2\theta - \sin^2\theta) + 2zr\cos\theta$$
$$25 = r^2\cos 2\theta + 2zr\cos\theta. \quad \blacksquare$$

**Example 7.2** *Convert the equation*

$$r^2\cos 2\theta = z$$

*from Cylindrical to Cartesian coordinates.*

*Solution.* With some algebra,

$$\begin{aligned} z &= r^2\cos 2\theta = r^2(\cos^2\theta - \sin^2\theta) \\ &= r^2\cos^2\theta - r^2\sin^2\theta \\ &= (r\cos\theta)^2 - (r\sin\theta)^2 \\ &= x^2 - y^2 \quad \blacksquare \end{aligned}$$

**Example 7.3** *Convert the expression*

$$r = 2z\sin\theta$$

*from Cylindrical coordinates to Cartesian coordinates.*

*Solution.* Multiply through by $r$ to give

$$r^2 = 2zr\sin\theta$$

Then use the identities $r^2 = x^2 + y^2$ and $y = r\sin\theta$ to get

$$x^2 + y^2 = 2zy. \quad \blacksquare$$

## Spherical Coordinates

In **spherical coordinates** each point $\mathbf{P} = (x, y, z)$ in Cartesian coordinates is represented by a triple $(\rho,\ \theta,\ \phi)$ where:

$\rho$ (the Greek letter "rho") is the distance from the origin to **P**.

$\theta$ (the Greek letter "theta") is the same as in cylindrical and polar coordinates.

$\phi$ (the Greek letter "phi')' is the angle between the $z$ axis and the line from the origin to **P**.

**Given the spherical coordinates** $(\rho, \theta, \phi)$, to find the Cartesian coordinates $(x, y, z)$,

$$x = \rho \sin\phi \cos\theta$$

$$y = \rho \sin\phi \sin\theta$$

$$z = \rho \cos\phi$$

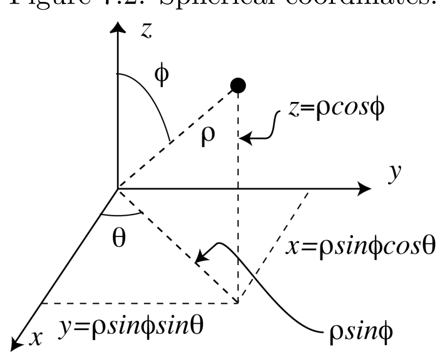
**Given the Cartesian coordinates** $(x, y, z)$, to find the spherical coordinates $(\rho, \theta, \phi)$,

$$\rho^2 = x^2 + y^2 + z^2$$

$$\tan\theta = y/x$$

$$\cos\phi = z/\sqrt{x^2 + y^2 + z^2}$$

Figure 7.2: Spherical coordinates.




**Example 7.4** *Convert the equation* $2x^2+2y^2-4z^2 = 0$ *from Cartesian Coordinates to Spherical coordinates.*

*Solution.*

$$\begin{aligned}
0 &= 2x^2 + 2y^2 - 4z^2 \\
&= 2(\rho \sin\phi \cos\theta)^2 + 2(\rho \sin\phi \sin\theta)^2 - 4(\rho\cos\phi)^2 \\
&= 2\rho^2 \sin^2\phi \cos^2\theta + 2\rho^2 \sin^2\phi \sin^2\theta - 4\rho^2 \cos^2\phi \\
&= 2\rho^2 \sin^2\phi (\cos^2\theta + \sin^2\theta) - 4\rho^2\cos^2\phi \\
&= 2\rho^2 \sin^2\phi - 4\rho^2\cos^2\phi \\
&= 2\rho^2 (\sin^2\phi - 2\cos^2\phi) \\
&= 2\rho^2 (1 - \cos^2\phi - 2\cos^2\phi) \\
&= 2\rho^2 (1 - 3\cos^2\phi)
\end{aligned}$$

Therefore we have two possible solutions:

$$\rho = 0 \qquad \text{(the single point at the origin)}$$

or

$$\begin{aligned}
1 &= 3\cos^2\phi \\
&\Rightarrow \cos^2\phi = 1/3 \\
&\Rightarrow \cos\phi = \pm 1/\sqrt{3} \\
&\Rightarrow \phi = \cos^{-1}(1/\sqrt{3})
\end{aligned}$$

which is the equation of a cone passing through the origin (and hence actually includes the first solution at the origin). ■

**Example 7.5** *Convert the equation*

$$\rho \sin\phi = 1$$

*from spherical coordinates to Cartesian coordinates.*

*Solution.* Squaring both sides of the equation gives

$$\begin{aligned}
\rho^2 \sin^2\phi &= 1 \\
&\rightarrow \rho^2(1 - \cos^2\phi) = 1 \\
&\rightarrow (x^2+y^2+z^2)\left(1 - \frac{z^2}{x^2+y^2+z^2}\right) = 1 \\
&\rightarrow x^2 + y^2 + z^2 - z^2 = 1 \\
&\rightarrow x^2 + y^2 = 1
\end{aligned}$$

which is the equation of a cylinder of radius 1 centered on the z-axis. ■

# Lecture 8

# Functions of Two Variables

In this section we will extend our definition of a function to allow for multiple variables in the argument. Before we formally define a multivariate function, we recall a few facts about real numbers and functions on real numbers. Recall our original, general definition of a function:

> A **function** $f$ is a rule that associates two sets, in the sense that each object $x$ in the first set $D$ is associated with a single object $y$ in the second set $R$, and we write
>
> $$f : D \mapsto R$$
>
> and
>
> $$y = f(x).$$
>
> We call $D$ the **domain of the function** and $R$ the **range of the function**.

In terms of functions of a real variable, we made some notational conventions:

- The symbol $\mathbb{R}$ represents the set of real numbers.
- Any line, such as any of the coordinate axes, is equivalent to $\mathbb{R}$ because there is a one-to-one relationship between the real numbers and the points on a line.
- For real valued functions, the both the domain $D$ and range $R$ are subsets of the real numbers, and we write

  $$D \subset \mathbb{R}$$

  $$R \subset \mathbb{R}$$

  so that

  $$f : (D \subset \mathbb{R}) \mapsto (R \subset \mathbb{R})$$

  If, in fact, $D = R = \mathbf{R}$, we write

  $$f : \mathbb{R} \mapsto \mathbb{R}$$

**Example 8.1** *The function $y = 3x + 7$ has a domain $D = \mathbb{R}$ and range $R = \mathbb{R}$. We can write $f : \mathbb{R} \mapsto \mathbb{R}$.* ■

**Example 8.2** *The function $y = x^2$ only takes on positive values, regardless of the value of $x$. Its domain is still $D = \mathbb{R}$ but its range is just the nonnegative real numbers, which we write as the union of the positive reals and $\{0\}$, namely, $R = \mathbb{R} \cup \{0\}$, and we can write $f : \mathbb{R} \mapsto \mathbb{R} \cup \{0\}$.* ■

**Example 8.3** *Find the natural domain and range of the function*

$$y = \frac{\sqrt{4 - x^2}}{1 - x}.$$

*Solution.* We make two observations: (1) the denominator becomes zero (and hence the function is undefined) when $x = 1$, and (2) the square root is undefined unless its argument is non-negative. Hence we require $x \neq 1$ and

$$4 - x^2 \geq 0 \Rightarrow 4 \geq x^2$$

Of course $x^2$ can not be negative, so we have

$$0 \leq x^2 \leq 4 \Rightarrow -2 \leq x \leq 2$$

If we include the earlier restriction $x \neq 1$ this becomes

$$D = \{x : -2 \leq x < 1 \text{ or } 1 < x \leq 2\} = [-2, 1) \cup (1, 2]$$

As the argument $x$ gets closer to 1 the function can take on arbitrarily large (for $x < 1$) or large negative (for $x > 1$) values and hence the range of the function is $\mathbb{R}$, so that

$$f : [2, 1) \cup (1, 2] \mapsto \mathbb{R} \ ■$$

Our goal now is to extend our definition of a function to include two variables. Such functions will have the form

$$z = f(x, y)$$

and will associate a point a on the $z$-axis with points in the $xy$-plane. Such functions will represent surfaces, and we will use the following notational observations:

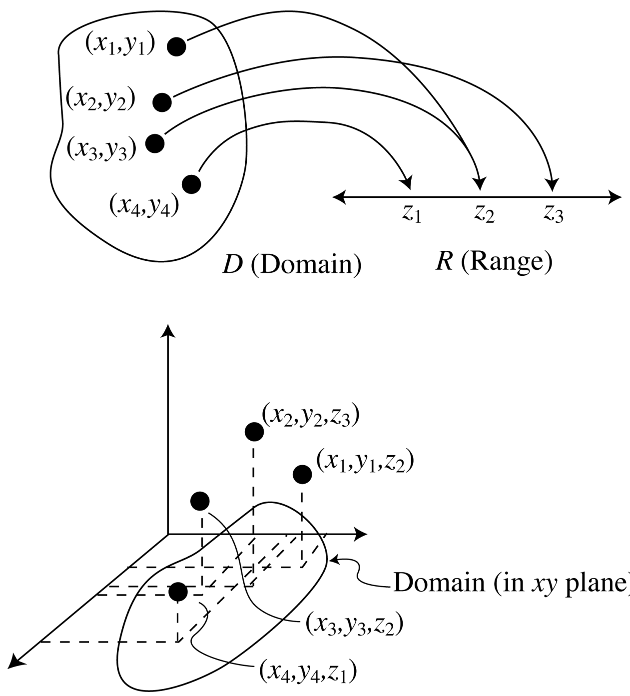
- $\mathbb{R}^2$ represents the set of ordered pairs $(x, y)$, where $x, y \in \mathbb{R}$.
- Any plane, such as any of the coordinate planes, is equivalent to $\mathbb{R}^2$ because there is a one-to-one relationship between the points on a plane and the set of all real-valued ordered pairs $(x, y)$.

so that we can write

$$f : (D \subset \mathbb{R}^2) \mapsto (R \subset \mathbb{R})$$

The ***natural domain*** of a function of two variables is the set in the $xy$ plane for which the function definition makes sense. The rules for determining the natural domain are the same as they are form functions of a single variable:

Figure 8.1: Visualization of a multivariate function. In both figures, the domain represents a set in the $xy$-plane, and the range is a subset of the $z$-axis. In the top, the function is visualized as a mapping between sets. The arrows indicate the mapping for four points in the domain that are mapped to three points in the range. The bottom figure visualizes this as a mapping embedded in 3D space, with each point $z = f(x, y)$ represented as a 3-tuple $(x, y, z)$ in space.




(a) Don't divide by zero

(b) Don't take the square root of a negative number.

Anything that remains is part of the natural domain of the function

> "Exclude the impossible and what remains, however improbable, is the solution." [Sherlock Holmes]

**Example 8.4** *Find the natural domain of the function*

$$f(x, y) = \sqrt{16 - x^2 - y^2}$$

*Solution.* Since we can't take the square root of a negative number, we exclude points where

$$16 - x^2 - y^2 < 0$$

Thus the natural domain is the set of all $(x, y)$ where

$$16 - x^2 - y^2 \geq 0,$$

or equivalently, the set where

$$x^2 + y^2 \leq 16$$

This is a disk (including its boundary) of radius 4 centered at the origin. ■

**Example 8.5** *Find the natural domain of the function*

$$f(x, y) = \frac{x^2 + y}{x + y^2}.$$

*Solution.* Since we can't divide by zero, we exclude all points where

$$x + y^2 = 0.$$

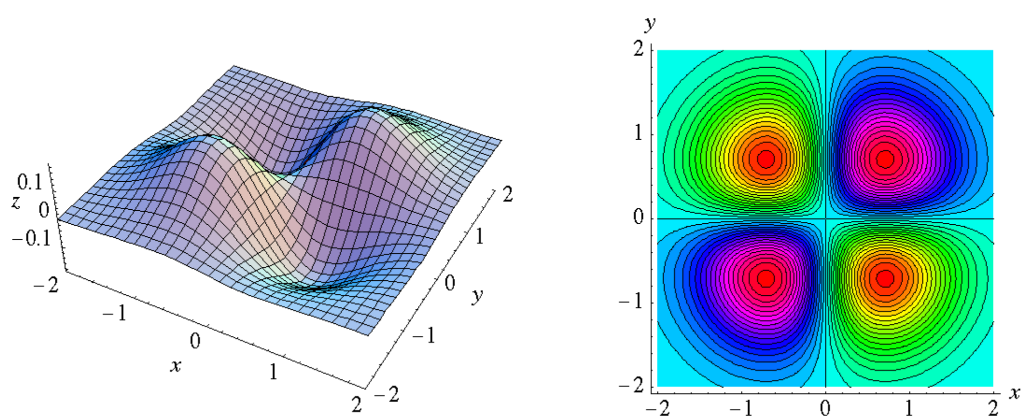
This is the parabola

$$x = -y^2.$$

Since there are not square roots, this is all we have to exclude, so the domain is the entire real plane except for points on the parabola $x = -y^2$, which we can write as

$$D = \left\{(x, y) : x \neq -y^2\right\}. \blacksquare$$

Figure 8.2: Left: A visualization of the function $y = xye^{-(x^2+y^2)}$ as a surface in 3D space. Right: Contour plot for the same function with the contours marked at values of $z = -0.18, -0.17, ..., 0.17, 0.18$. Colors(yell/green, negative; blue, pink: positive) are used to emphasize the values, but this is not always a feature of contour plots. It is more common to label each contour with a small number indicating the $z$-value it represents.




Functions of two variables are generally surfaces when plotted in 3-space; we have already seen some of them in plots of planes and quartic surfaces. Consider, for example, the function

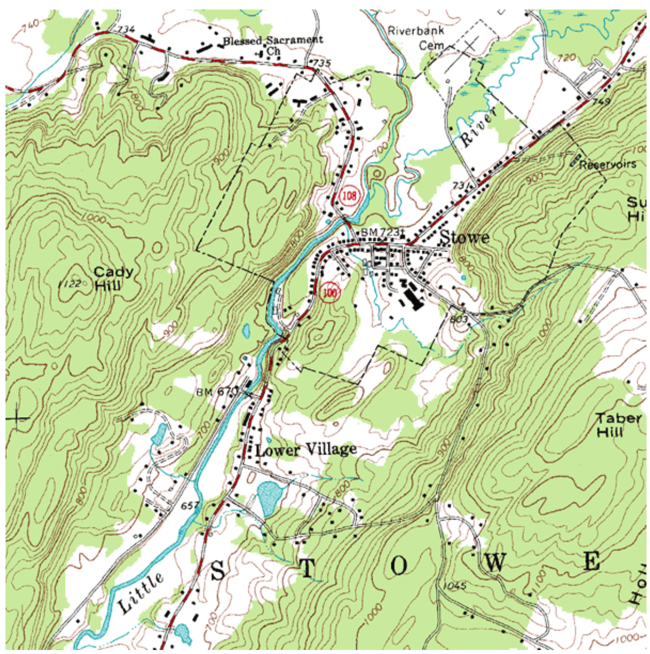
$$f : \mathbb{R}^2 \mapsto \mathbb{R}$$

described by the equation

$$z = xye^{-(x^2+y^2)}$$

It can be represented by a surface something like the one illustrated in figure 8.2. In addition to the surface plant, another useful tool is the **contour plot**, illustrated in the fight-hand illustration of figure 8.2. If we take a surface plot and slice it along planes parallel to the $xy$ plane we will obtain a series of sections that can be represented by curves in the $xy$ plane. Each curve, called a countour, represents a particular slice, or height, above the $xy$ plane. Countour plots are commonly used by geographers, for example, to draw topographic (e.g., figure 8.3) or weather maps ( see figure 8.4), among other things. Each curve in a contour plot is called a **level curve**.

Figure 8.3: Example of a contour plot used to illustrate altitude on a topographic map. Contours are drawn for every 20 feet of altitude; small numbers next to contours annotate the contours at 100 foot intervals. [Taken from USGS Digital Raster file 044072d6.tif for the Stowe, VT, USA quadrangle, as published at http://en.wikipedia.org/ article "Topographic Maps." This image is in the public domain.]

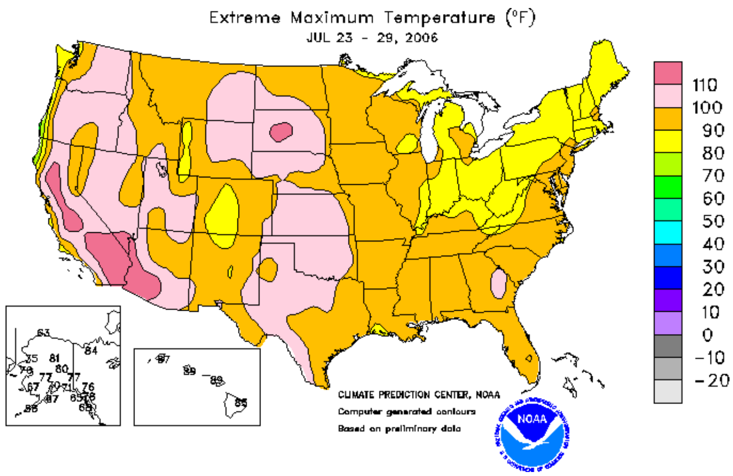



**Example 8.6** *Find the level curves of the function*

$$z = 2 - x - y^2$$

*for* $z = 1, 0, -1$.

Figure 8.4: Example of a contour plot illustrating the maximum temperature during the week of July 23, 2006. Rather than annotating the contours the space between the contours is colored, and the temperature is read by comparison with the legend on the figure. [Taken from the US NOAA/National Weather Service Climate Prediction center, as posted at http://www.cpc.ncep.noaa.gov.]




*Solution.* The level curves for $z$=0 are the curves

$$\begin{aligned} 0 &= 2 - x - y^2 \\ &\Rightarrow y^2 = 2 - x \\ &\Rightarrow y = \pm\sqrt{2 - x} \end{aligned}$$

The level curves for $z$=1 are the curves

$$\begin{aligned} 1 &= 2 - x - y^2 \\ &\Rightarrow y^2 = 2 - x - 1 = 1 - x \\ &\Rightarrow y = \pm\sqrt{1 - x} \end{aligned}$$

The level curves for $z$=-1 are the curves

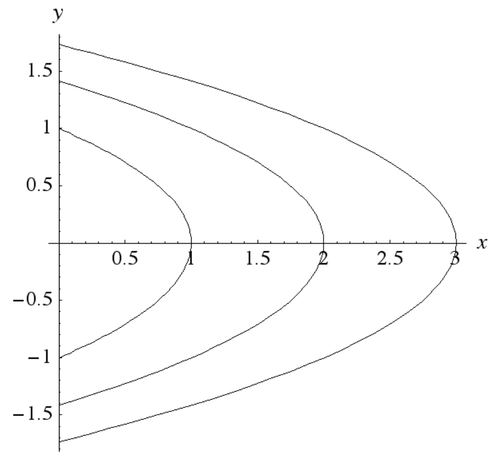
$$\begin{aligned} -1 &= 2 - x - y^2 \\ &\Rightarrow y^2 = 2 - x + 1 = 3 - x \\ &\Rightarrow y = \pm\sqrt{3 - x} \end{aligned}$$

The level curves are illustrated in figure 8.5. ■

**Example 8.7** *Find a general form for the level curves $z = k$ for the function*

$$z = \frac{x^2 + y}{x + y^2}$$

Figure 8.5: Level curves for the function $z = 2 - x - y^2$ at $z = -1, 0, 1$.




*Solution.* Substituting $z = k$ and cross-multiplying, we find

$$\begin{aligned} x^2 + y &= k(x + y^2) = kx + ky^2 \\ x^2 - kx &= ky^2 - y \end{aligned}$$

Completing the squares on the left hand side of the equation,

$$\begin{aligned} x^2 - kx &= x^2 - kx + (k/2)^2 - (k/2)^2 \\ &= (x - k/2)^2 - k^2/4 \end{aligned}$$

Doing a similar manipulation on the right hand side of the equation,

$$\begin{aligned} ky^2 - y &= k(y^2 - y/k) \\ &= k\left[y^2 - (y/k) + \left(\frac{1}{2k}\right)^2 - \left(\frac{1}{2k}\right)^2\right] \\ &= k\left[\left(y - \frac{1}{2k}\right)^2 - \frac{1}{4k^2}\right] \\ &= k(y - 1/(2k))^2 - 1/(4k) \end{aligned}$$

Equating the two expressions,

$$\begin{aligned} (x - k/2)^2 - k^2/4 &= k(y - (1/2k))^2 - 1/4k \\ (x - k/2)^2 - k(y - (1/2k))^2 &= \frac{k^2}{4} - \frac{1}{4k} = \frac{k^3 - 1}{4k} \\ \frac{(x - k/2)^2}{(k^3 - 1)/4k} - \frac{(y - (1/2k))^2}{(k^3 - 1)/(4k^2)} &= 1 \end{aligned}$$

which is a hyperbola. ■

# Lecture 9

# The Partial Derivative

**Definition 9.1** *Let $f(x, y)$ be a function of two variables. Then* **the partial derivative of $f$ with respect to $x$** *is defined as*

$$\frac{\partial f}{\partial x} = f_x(x, y) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x} \tag{9.1}$$

*if the limit exists, and the* **the partial derivative of $f$ with respect to $y$** *is similarly defined as*

$$\frac{\partial f}{\partial y} = f_y(x, y) = \lim_{\Delta y \to 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y} \tag{9.2}$$

*if that limit exists.*

Calculation of partial derivatives is similar to calculation of ordinary derivatives. To calculate $\partial f/\partial x$, for example, differentiate with respect to $x$ while treating $y$ as a constant; to calculate $\partial f/\partial y$, differentiate with respect to $y$ while treating $x$ as a constant.

**Example 9.1** *Find the partial derivatives of $f(x, y) = x$.*

*Solution.* Applying equations (9.1) and (9.2), we calculate

$$\begin{aligned} \frac{\partial f}{\partial x} &= f_x(x, y) = \lim_{h \to 0} \frac{f(x + h, y) - f(x, y)}{h} \\ &= \lim_{h \to 0} \frac{(x + h) - x}{h} = 1 \\ \frac{\partial f}{\partial y} &= f_y(x, y) = \lim_{h \to 0} \frac{f(x, y + h) - f(x, y)}{h} \\ &= \lim_{h \to 0} \frac{x - x}{h} = 0. \blacksquare \end{aligned}$$

All of the usual rules of differentiation apply to partial derivatives, including things like the product rule, the quotient rule, and the derivatives of basic functions.

**Example 9.2** *Find the partial derivatives with respect to $x$ and $y$ of $f(x, y) = e^y \sin x$.*

*Solution.*

$$\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}(e^y \sin x) = e^y \frac{\partial}{\partial x}(\sin x) = e^y \cos x \tag{9.3}$$

$$\frac{\partial f}{\partial y} = \frac{\partial}{\partial y}(e^y \sin x) = \sin x \frac{\partial}{\partial y}(e^y) = e^y \sin x \blacksquare \tag{9.4}$$

**Example 9.3** *Find the partial derivatives of $f(x, y) = x^3y^2 - 5x + 7y^3$.*

*Solution.*

$$\begin{aligned}
\frac{\partial f}{\partial x} &= \frac{\partial}{\partial x}(x^3y^2 - 5x + 7y^3) \\
&= \frac{\partial}{\partial x}(x^3y^2) - \frac{\partial}{\partial x}(5x) + \frac{\partial}{\partial x}(7y^3) \\
&= y^2 \frac{\partial}{\partial x}(x^3) - \frac{\partial}{\partial x}(5x) + \frac{\partial}{\partial x}(7y^3) \\
&= 3x^2y^2 - 5 \\
\frac{\partial f}{\partial y} &= \frac{\partial}{\partial y}(x^3y^2 - 5x + 7y^3) \\
&= \frac{\partial}{\partial y}(x^3y^2) - \frac{\partial}{\partial y}(5x) + \frac{\partial}{\partial y}(7y^3) \\
&= x^3 \frac{\partial}{\partial y}(y^2) - \frac{\partial}{\partial y}(5x) + \frac{\partial}{\partial y}(7y^3) \\
&= 2x^3y + 21y^2 \blacksquare
\end{aligned}$$

## Physical Meaning of Partials.

Partial derivatives represent rates of change, just as ordinary derivatives. If $T(x,t)$ is the temperature of an object as a function of position and time then

$$\frac{\partial T}{\partial x}(x, t)$$

represents the change in temperature at time $t$ with respect to position, i.e., the slope of the temperature vs. position curve, and

$$\frac{\partial T}{\partial t}(x, t)$$

gives the change in temperature at a fixed position $x$, with respect to time.

If, on the other hand, $T(x,y)$ gives the temperature of an object as a function of position in the $xy$ plane then

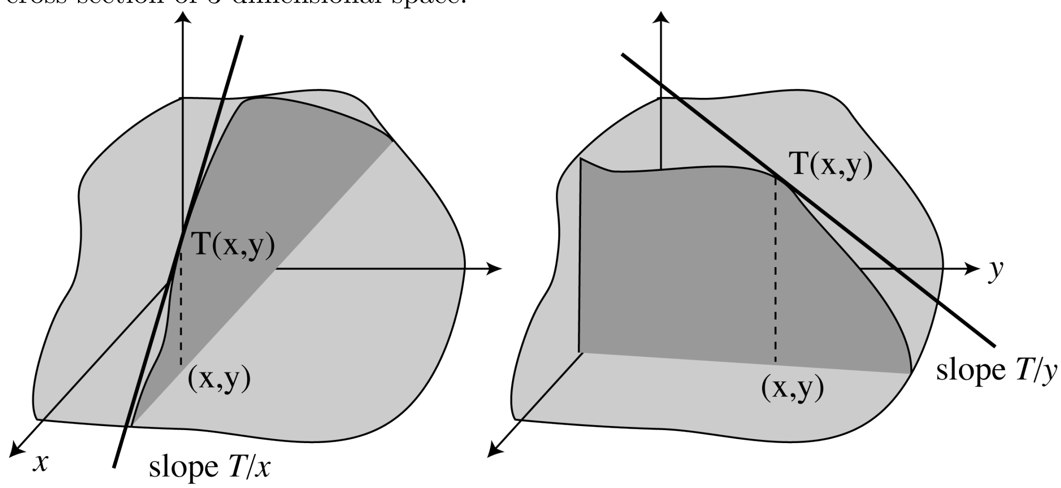
$$\frac{\partial T}{\partial x}(x, y)$$

represents the slope of the temperature vs position curve in a plane perpendicular to the y-axis, while

$$\frac{\partial T}{\partial y}(x, y)$$

represents the slope of the temperature vs position curve in a plane perpendicular to the y-axis.

Figure 9.1: The partial derivatives represent the ordinary slope of a function in a cross-section of 3-dimensional space.




**Example 9.4** *Suppose that the temperature in degrees Celsius on a metal plate in the xy plane is given by*

$$T(x, y) = 4 + 2x^2 + y^3$$

*where $x$ and $y$ are measured in feet. What is the change of temperature with respect to distance, measured in feet, if we start moving from the point $(3, 2)$ in the direction of the positive y-axis, as illustrated in figure 9.2 ?*

*Solution.* If we move in the direction of the positive y-axis we are moving perpendicular to the $x$-axis (with $x$ fixed) in the $xy$ plane, so we are interesting in finding the partial derivative $\partial T/\partial y$.

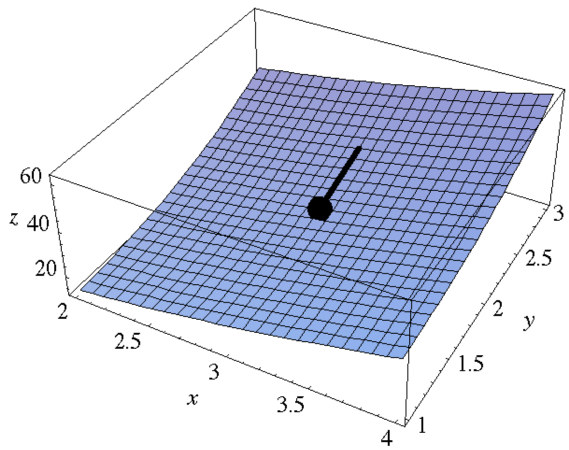
$$\frac{\partial T(x, y)}{\partial y} = \frac{\partial}{\partial y}(4 + 2x^2 + y^3) = 0 + 0 + 3y^2 = 3y^2$$

At the point $(3, 2)$ we have $x = 3$ and $y = 2$. Thus

$$\frac{\partial T(3, 2)}{\partial y} = 3(2)^2 = 12 \deg/\text{foot}$$

In other words, the temperature increases by 12 degrees for every foot we move in the $y$ direction.■

Figure 9.2: The function $T(x, y)$ discused in example 9.4.




## Higher order partial derivatives

The second and higher order partial derivatives

$$f_{xx} = \frac{\partial}{\partial x}\frac{\partial f}{\partial x} = \frac{\partial^2 f}{\partial x^2}$$

$$f_{yy} = \frac{\partial}{\partial y}\frac{\partial f}{\partial y} = \frac{\partial^2 f}{\partial y^2}$$

In addition, there are mixed partial derivatives for higher orders.

$$f_{xy} = (f_x)_y = \frac{\partial}{\partial y}\frac{\partial f}{\partial x} = \frac{\partial^2 f}{\partial y \partial x}$$

$$f_{yx} = (f_y)_x = \frac{\partial}{\partial x}\frac{\partial f}{\partial y} = \frac{\partial^2 f}{\partial x \partial y}$$

**Example 9.5** *Find all the second order partial derivatives of*

$$f(x, y) = x^3y^2 - 5x + 7y^3$$

From equations (9.3) and (9.4),

$$\frac{\partial f}{\partial x} = 3x^2y^2 - 5, \qquad \frac{\partial f}{\partial y} = 2x^3y + 21y^2$$

Hence

$$f_{xx} = \frac{\partial}{\partial x}\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}(3x^2y^2 - 5) = 6xy^2$$

$$f_{xy} = \frac{\partial}{\partial y}\frac{\partial f}{\partial x} = \frac{\partial}{\partial y}(3x^2y^2 - 5) = 6x^2y$$

$$f_{yy} = \frac{\partial}{\partial y}\frac{\partial f}{\partial y} = \frac{\partial}{\partial y}(2x^3y + 21y^2) = 2x^3 + 42y$$
$$f_{yx} = \frac{\partial}{\partial x}\frac{\partial f}{\partial y} = \frac{\partial}{\partial x}(2x^3y + 21y^2) = 6x^2y.$$

We observe in passing that $f_{xy} = f_{yx}$. ■

**Example 9.6** *The **Heat Equation** is*

$$\frac{\partial u}{\partial t} = c\frac{\partial^2 u}{\partial x^2}$$

*where $u(x,t)$ gives the temperature of an object as a function of position and time. Show that*

$$u = \frac{1}{\sqrt{t}}e^{-x^2/(4ct)}$$

*satisfies the heat equation (i.e., that it is a solution of the heat equation).*

*Solution.* By the product rule:

$$\begin{aligned}\frac{\partial u}{\partial t} &= \frac{\partial}{\partial t}\left(\frac{1}{\sqrt{t}}e^{-x^2/(4ct)}\right)\\ &= \frac{1}{\sqrt{t}}\frac{\partial}{\partial t}e^{-x^2/(4ct)} + e^{-x^2/(4ct)}\frac{\partial}{\partial t}\frac{1}{\sqrt{t}}\end{aligned}$$

By the chain rule,

$$\begin{aligned}\frac{\partial}{\partial t}e^{-x^2/(4ct)} &= e^{-x^2/(4ct)}\frac{\partial}{\partial t}\left(\frac{-x^2}{4ct}\right)\\ &= e^{-x^2/(4ct)}\left(\frac{-x^2}{4c}\right)\frac{\partial}{\partial t}(t^{-1})\\ &= e^{-x^2/(4ct)}\left(\frac{-x^2}{4c}\right)(-t^{-2})\\ &= \frac{x^2}{4ct^2}e^{-x^2/(4ct)}\end{aligned}$$

Furthermore

$$\frac{\partial}{\partial t}\frac{1}{\sqrt{t}} = \frac{\partial}{\partial t}(t^{-1/2}) = (-1/2)t^{-3/2}$$

Hence

$$\begin{aligned}\frac{\partial u}{\partial t} &= \frac{1}{\sqrt{t}}\left(\frac{x^2}{4ct^2}\right)e^{-x^2/4ct} + e^{-x^2/4ct}(-1/2)t^{-3/2}\\ &= e^{-x^2/4ct}\left(\frac{x^2}{4ct^{5/2}} - \frac{1}{2t^{3/2}}\right)\\ &= \frac{e^{-x^2/4ct}}{\sqrt{t}}\left(\frac{x^2}{4ct^2} - \frac{1}{2t}\right)\end{aligned}$$

Similarly, by the chain rule

$$\begin{aligned}\frac{\partial u}{\partial x} &= \frac{\partial}{\partial x}\left(\frac{1}{\sqrt{t}}e^{-x^2/(4ct)}\right) = \frac{1}{\sqrt{t}}e^{-x^2/(4ct)}\frac{\partial}{\partial x}\left(\frac{-x^2}{4ct}\right)\\ &= \frac{1}{\sqrt{t}}e^{-x^2/(4ct)}\left(\frac{-2x}{4ct}\right) = -\frac{x}{2ct\sqrt{t}}e^{-x^2/(4ct)}\end{aligned}$$

By the product rule,

$$\begin{aligned}\frac{\partial^2 u}{\partial x^2} &= -\frac{\partial}{\partial x}\left(\frac{x}{2ct\sqrt{t}}e^{-x^2/(4ct)}\right)\\ &= \frac{-x}{2ct\sqrt{t}}\frac{\partial}{\partial x}e^{-x^2/(4ct)} - e^{-x^2/(4ct)}\frac{\partial}{\partial x}\frac{x}{2ct\sqrt{t}}\\ &= \frac{-x}{2ct\sqrt{t}}\left(e^{-x^2/(4ct)}\right)\left(\frac{-2x}{4ct}\right) + \left(-e^{-x^2/(4ct)}\right)\left(\frac{1}{2ct\sqrt{t}}\right)\\ &= \frac{1}{\sqrt{t}}e^{-x^2/(4ct)}\left(\frac{x^2}{4c^2t^2} - \frac{1}{2ct}\right)\end{aligned}$$

Multiplying the last equation through by $c$,

$$c\frac{\partial^2 u}{\partial x^2} = \frac{1}{\sqrt{t}}e^{-x^2/(4ct)}\left(\frac{x^2}{4ct^2} - \frac{1}{2t}\right) = \frac{\partial u}{\partial t} \blacksquare$$

## Equivalence of Mixed Partials

We observed at the end of example (9.5) that $f_{xy} = f_{yx}$. This property is true in general, although it is not stated formally in the book until section 15.3 theorem B. To see why it is true we observe the following:

$$\begin{aligned}f_{xy}(x,y) &= \frac{\partial}{\partial y}\frac{\partial f(x,y)}{\partial x} = \frac{\partial}{\partial y}f_x(x,y)\\ &= \lim_{k\to 0}\frac{f_x(x,y+k) - f_x(x,y)}{k}\\ &= \lim_{k\to 0}\frac{\lim_{h\to 0}\dfrac{f(x+h,y+k) - f(x,y+k)}{h} - \lim_{h\to 0}\dfrac{f(x+h,y) - f(x,y)}{h}}{k}\\ &= \lim_{k\to 0}\lim_{h\to 0}\frac{[f(x+h,y+k) - f(x,y+k)] - [f(x+h,y) - f(x,y)]}{hk}\\ &= \lim_{k\to 0}\lim_{h\to 0}\frac{[f(x+h,y+k) - f(x+h,y)] - [f(x,y+k) - f(x,y)]}{hk}\\ &= \lim_{h\to 0}\frac{\lim_{k\to 0}\dfrac{f(x+h,y+k) - f(x+h,y)}{k} - \lim_{k\to 0}\dfrac{f(x,y+k) - f(x,y)}{k}}{h}\\ &= \lim_{h\to 0}\frac{f_y(x+h,y) - f_y(x,y)}{h}\\ &= \frac{\partial}{\partial x}f_y(x,y) = f_{yx}(x,y).\end{aligned}$$

Hence in general it is safe to assume that $f_{xy} = f_{yx}$.

## Harmonic Functions

**Definition 9.2** *A function $f(x, y, z)$ is said to be harmonic if it satisfies* **Laplace's Equation**,

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = 0$$

Harmonic functions are discussed in problems 33 and 34 of the text. To see why they called Harmonic Functions, we look for solutions that satisfy

$$f(x, y, z) = X(x)Y(y)Z(z)$$

where $X$ is only a function of $x$, $Y$ is only a function $y$, and $Z$ is only a function of $z$. Taking partial derivatives,

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial^2}{\partial x^2}X(x)Y(y)Z(z) = \frac{d^2 X(x)}{dx^2}Y(y)Z(z) = X''(x)Y(y)Z(z)$$

$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial^2}{\partial y^2}X(x)Y(y)Z(z) = X(x)\frac{d^2 Y(y)}{dy^2}Z(z) = X(x)Y''(y)(y)Z(z)$$

$$\frac{\partial^2 f}{\partial z^2} = \frac{\partial^2}{\partial z^2}X(x)Y(y)Z(z) = X(x)Y(y)\frac{d^2 Z(z)}{dz^2} = X(x)Y(y)Z''(z)(z)$$

Therefore

$$X''(x)Y(y)Z(z) + X(x)Y''(y)Z(z) + X(x)Y(y)Z''(z) = 0$$

Dividing through by $f(x, y, z) = X(x)Y(y)Z(z)$ gives

$$\frac{X''(x)}{X(x)} + \frac{Y''(y)}{Y(y)} + \frac{Z''(z)}{Z(z)} = 0$$

Rearranging

$$\frac{X''(x)}{X(x)} = -\frac{Y''(y)}{Y(y)} - \frac{Z''(z)}{Z(z)}$$

The left hand side of the equation depends only on $x$ and not on $y$ or $z$, while the right hand side of the equation depends on $y$ and $z$ and not on $x$. The only way this can be true is if they are both equal to a constant, call it $K_1$. Then

$$\frac{X''(x)}{X(x)} = K_1 = -\frac{Y''(y)}{Y(y)} - \frac{Z''(z)}{Z(z)}$$

and hence

$$X''(x) = K_1 X(x)$$

$$\frac{Y''(y)}{Y(y)} + K_1 = -\frac{Z''(z)}{Z(z)}$$

Repeating the process, the last equation must equal a constant that we call $K_2$ so that

$$Z''(z) = -K_2 Z z$$

$$Y''(y) = (K_2 - K_1)Y(y) = K_3 Y(y)$$

where$K_3 = K_2 - K_1$. So each of the three functions $X(x)$, $Y(y)$, $Z(z)$ satisfy second order differential equations of the form

$$X''(x) = -kX(x)$$

which you might recognize as the equation of a ”spring” or ”oscillating string.”

# Lecture 10

# Limits and Continuity

In this section we will generalize the definition of a limit to functions of more than one variable. In particular, we will give a meaning to the expression

$$\lim_{(x,y)\to(a,b)} f(x,y) = L$$

which is read as

"The limit of $f(x,y)$ as the point $(x,y)$ approaches the point $(a,b)$ is $L$."

The complication arises from the fact that we can approach the point $(a,b)$ from any direction. In one dimension, we could either approach from the left or from the right, and we defined a variety of notations to take this into account:

$$L^+ = \lim_{x\to a^+} f(x)$$

$$L^- = \lim_{x\to a^-} f(x)$$

and then the limit

$$\lim_{x\to a} f(x) = L$$

is defined only when

$$L = L^+ = L^-,$$

i.e., the limit is only defined when the limits from the left and the right both exist and are equal to one another. In two dimensions, we can approach from any direction, not just from the left or the right (figure 10.1).

Let us return to the single-dimensional case, as illustrated in figure 10.2. Whichever direction we approach $a$ from, we must get the same value. If the function approaches the same limit from both directions (the left and the right) then we say the limit exists. Formally, we say that

$$\lim_{x\to a} f(x) = L$$

exists if and only if for any $\epsilon > 0$, no matter how small, there exists some $\delta > 0$ (that is allowed to depend functionally on $\epsilon$) such that whenever $0 < |x-a| < \delta$

Figure 10.1: In one dimension we can only approach a point along the $x$-axis from two directions: the left and the right. In two dimensions we can approach a point in the $xy$-plane from an infinite number of directions.

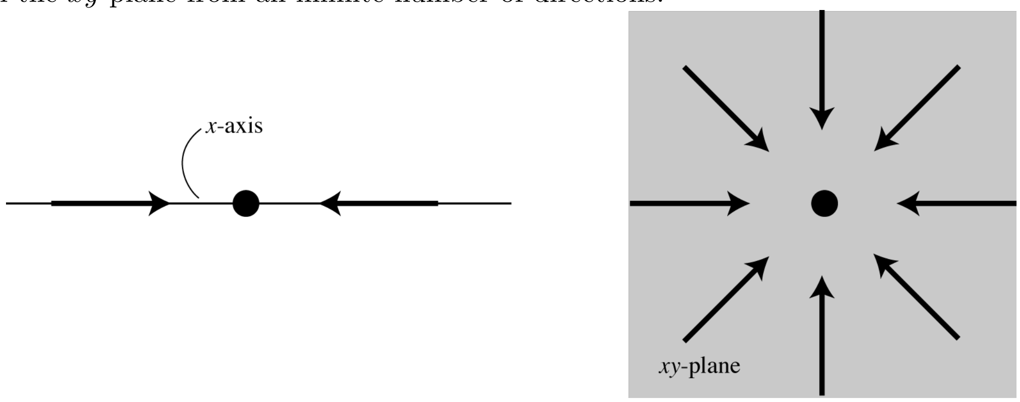
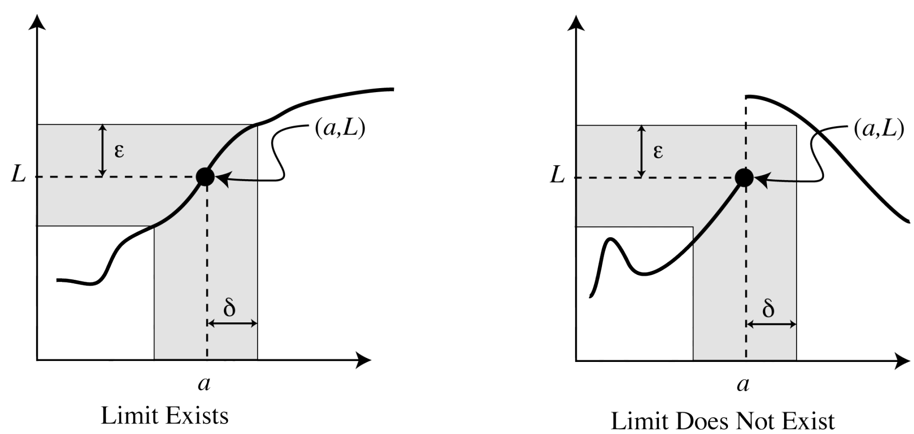
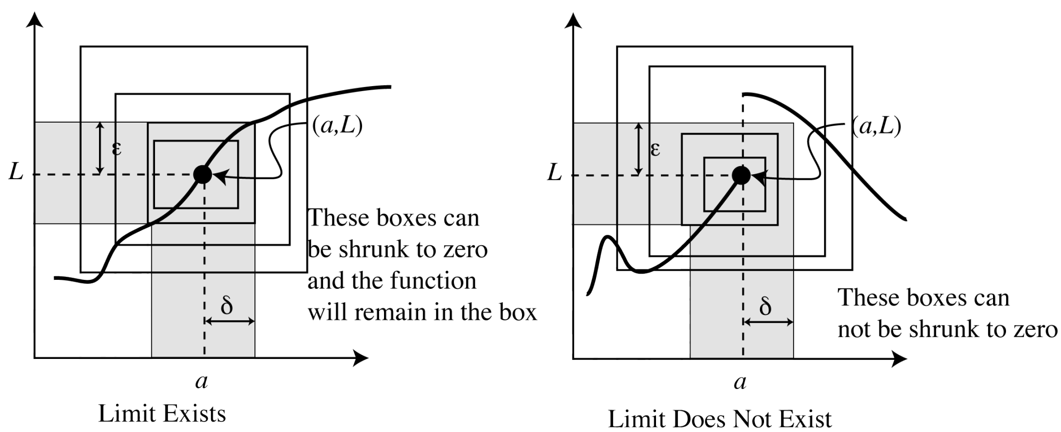



Figure 10.2: The limit of a function of one variable. Left: the function $f(x) \to L$ from both directions, and we say that $\lim_{x\to a} = L$ exists. Right: The function $f(x)$ approaches $L$ from the left, but not from the right.



then $|f(x) - L| < \epsilon$. No matter how close we get to $L$ along the $y$-axis (within a distance $\epsilon$), there is some interval about $a$ along the $x$-axis (of width we call $2\delta$ such that, if we draw a box about the point $(a, L)$ of size $\delta \times \epsilon$, and then slowly shrink the box to zero, the function will always remain within the box. If the limit does not exist, then as we shrink the box, at some point there will not be a representation of the function from left to right across the box (see figure 10.3).

Figure 10.3: Limits of a function of a single variable. On the left, the limit exists, and we can shrink a box about $(a, L)$ however small we like, and the function remains entirely within the box. On the right, we eventually get to a point where the function from the left is not in the box, and the limit does not exist.




Now consider the case of a limit in three dimensions. Rather than approaching the point to within an interval of width $2\delta$ and letting the size $\delta \to 0$ we can approach in the $xy$ plane from any direction. So the interval becomes a disk of radius $\delta$, and the box becomes a cylinder of radius $\delta$ and height $\epsilon$.

**Definition 10.1** *A* **neighborhood** *of radius $\delta$ of a point* $\mathbf{P}$ *is the set of all points* $\mathbf{Q}$ *satisfying* $\|\mathbf{P} - \mathbf{Q}\| < \delta$. *In 1D, a neighborhood is called an* **interval**. *In 2D, a neighborhood is called a* **disk**. *In 3D (and higher) a neighborhood of* $\mathbf{P}$ *is called a* **ball**.

Mathematically, the definition of a limit is essentially the same in all dimensions.

**Definition 10.2** *Let $f(x, y)$ be a function of two variables. Then the* **limit of** $f(x, y)$ **as** $(x, y)$ **approaches** $(a, b)$, *which we write as*

$$\lim_{(x,y)\to(a,b)} f(x, y) = L$$

*exists if and only if for any $\epsilon > 0$ there exists some $\delta > 0$ such that whenever $0 < \|f(x, y) - (a, b)\| < \delta$ (i.e., $(x, y)$ is in a neighborhood of radius $\delta$ of $(a, b)$) then $|f(x, y) - L| < \epsilon$ (i.e., $f(x, y)$ is in some neighborhood of radius $\epsilon$ of $L$.*

We can get a general definition of a limit (in any dimension) by merely omitting the language specific to 2D.

**Definition 10.3** *Let* $\mathbf{f}(\mathbf{x}) : \mathbb{R}^m \mapsto \mathbb{R}^n$*, let* $\mathbf{P}$ *be a point in* $\mathbb{R}^m$*, and* $\mathbf{L}$ *a point in* $\mathbb{R}^n$*. Then the* **limit of $\mathbf{f}(\mathbf{x})$ as $\mathbf{x}$ approaches $\mathbf{P}$**, *which we write as*

$$\lim_{\mathbf{x}\to\mathbf{P}} \mathbf{f}(\mathbf{x}) = \mathbf{L}$$

*exists if and only if for any* $\epsilon > 0$ *there exists some* $\delta > 0$ *such that whenever* $0 < \|\mathbf{f}(\mathbf{x}) - (\mathbf{P})\| < \delta$ *(i.e.,* $\mathbf{x}$ *is in a neighborhood of radius* $\delta$ *of* $\mathbf{P}$*) then* $\|\mathbf{f}(\mathbf{x}) - \mathbf{L}\| < \epsilon$ *(i.e.,* $\mathbf{f}(\mathbf{x})$ *is in some neighborhood of radius* $\epsilon$ *of* $\mathbf{L}$*).*

**Example 10.1** *Calculate*

$$\lim_{(x,y)\to(0,0)} \frac{x^2 - y^2}{x^2 + y^2}$$

*as you approach the origin*

(a) *along the x-axis;*

(b) *along the line y=3x;*

(c) *along the parabola* $y = 5x^2$.

*Solution.*

(a) along the x-axis we have y=0, and we can approach the origin by letting $x \to 0$. Therefore

$$\lim_{(x,y)\to(0,0),y=0} \frac{x^2 - y^2}{x^2 + y^2} = \lim_{x\to 0} \frac{x^2 - 0^2}{x^2 + 0^2} = \lim_{x\to 0} \frac{x^2}{x^2} = \lim_{x\to 0} 1 = 1$$

(b) Along the line $y = 3x$ we have

$$\lim_{(x,y)\to(0,0),y=3x} \frac{x^2 - y^2}{x^2 + y^2} = \lim_{x\to 0} \frac{x^2 - (3x)^2}{x^2 + (3x)^2} = \lim_{x\to 0} \frac{-8x^2}{10x^2} = -\frac{4}{5}$$

(c) Along the parabola $y = 5x^2$ we have

$$\begin{aligned} \lim_{(x,y)\to(0,0),y=x^2} \frac{x^2 - (5x^2)^2}{x^2 + (5x^2)^2} &= \lim_{x\to 0} \frac{x^2 - 25x^4}{x^2 + 25x^4} = \lim_{x\to 0} \frac{x^2(1 - 25x^2)}{x^2(1 + 25x^2)} \\ &= \lim_{x\to 0} \frac{1 - 25x^2}{1 + 25x^2} = 1 \end{aligned}$$

The first and third limits are the same, but the third limit is different. Therefore the limit does not exist. ■

As the above example shows, to show that a limit does not exist we need to calculate the limit along different approach paths and show that different numbers result. Showing that limit does exist is considerably more difficult than showing that it does not exist.

**Example 10.2** *Show that*

$$\lim_{(x,y)\to(0,0)} \frac{x^4 - y^4}{x^4 + y^4}$$

*does not exist.*

*Solution.* We calculate the limit in two directions: along the x-axis and along the y-axis.

Along the x-axis, y=0, so that

$$\lim_{(x,y)\to(0,0),x-axis} \frac{x^4 - y^4}{x^4 + y^4} = \lim_{x\to 0} \frac{x^4 - 0^4}{x^4 + 0^4} = 1$$

Along the y-axis, x=0, so that

$$\lim_{(x,y)\to(0,0),y-axis} \frac{x^4 - y^4}{x^4 + y^4} = \lim_{y\to 0} \frac{0^4 - y^4}{0^4 + y^4} = -1$$

Since the limits along the different paths are unequal we may conclude that the limit does not exist. ■

**Definition 10.4** *A function is said to be* **continuous at a point** $(a, b)$ *if*

$$\lim_{(x,y)\to(a,b)} f(x, y) = f(a, b)$$

*A function is said to be* **continuous on a set** $S$ *if it is continuous at every point in* $S$*.*

**Theorem 10.1** *The following classes of functions are continuous:*

(a) *Lines*

(b) *Polynomials*

(c) *Composite functions of continuous functions, e.g.,* $f(g(x, y))$

(d) *Rational functions except where the denominator equals zero.*

**Example 10.3** *Determine where the function*

$$f(x, y) = \frac{x^4 - y^4}{x^4 + y^4}$$

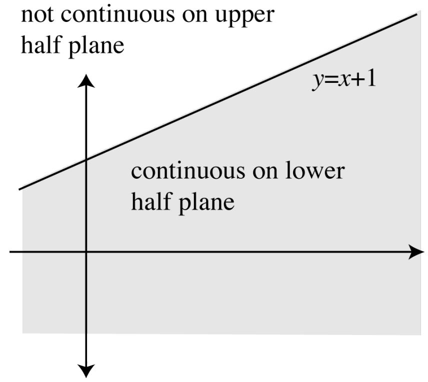
*is continuous.*

*Solution.* $f(x, y)$ is a rational function. Therefore it is continuous except where the denominator is zero. The function is not continuous when

$$x^4 + y^4 = 0$$

which only occurs at the origin. Therefore the function is continuous every except for the origin.■

Figure 10.4: The function $z = \sqrt{x - y + 1}$ is continuous on the lower halp plane beneath the line $y = x + 1$.




**Example 10.4** *Determine where the function*

$$f(x, y) = \sqrt{x - y + 1}$$

*is continuous.*

This is a function of the form $f(x, y) = g(h(x, y))$ where $g(z) = \sqrt{z}$ is a function of a single argument that is defined and continuous on $z > 0$, and $h(x, y) = x - y + 1$ is also a continuous function. Since this is a composite function of a continuous function, then it is continuous everywhere the argument of $g(z)$ is positive. This requires $x - y + 1 > 0$, or equivalently, $y < x + 1$ which is the half-plane underneath the line $y = x + 1$. In other words, the function is continuous everywhere in the half plane below $y = x + 1$ (see figure 10.4). ■

# Lecture 11

# Gradients and the Directional Derivative

We will learn in this section that the concepts of **differentiability** and **local linearity** are equivalent, and we will generalize our definitions of the derivative and and partial derivative from one-dimensional objects or slopes to generalized slopes in higher dimensional spaces.

**Definition 11.1** *A function $f(x) : \mathbb{R} \mapsto \mathbb{R}$ is said to be* **locally linear at** $x = a$ *if it can be approximated by a line in some neighborhood of $a$, i.e., if there is some constant $m$ such that*

$$f(a + h) = f(a) + hm + h\epsilon(h) \tag{11.1}$$

*where*

$$\lim_{h\to 0} \epsilon(h) = 0$$

*The number $m$ is called the* **slope** *of the line.*

Suppose that a function $f(x)$ is locally linear at $a$. Solving equation 11.1 for $\epsilon(h)$,

$$\epsilon(h) = \frac{f(a + h) - f(a)}{h} - m$$

and therefore

$$0 = \lim_{h\to 0} \epsilon(h) = \lim_{h\to 0} \left( \frac{f(a + h) - f(a)}{h} - m \right) = f'(a) - m$$

The last equality only makes sense if the limit exists, which we know it does because we have assumed that $f(x)$ is locally linear, and therefore $\lim_{h\to 0} \epsilon(h)$ exists and equals zero. But the first term in the final limit is the derivative $f'(a)$, and we say the $f(x)$ is **differentiable** if and only of the derivative $f'(a)$, defined by this limit, exists. Therefore we have proven that **a function is locally linear if and only if it is differentiable**. Our immediate goal is to extend this result to multivariate functions.

**Definition 11.2** *A function $f(x,y):\mathbb{R}^2\mapsto\mathbb{R}$ is said to be* **locally linear at the point** $(a,b)$ *if there exist numbers $h$ and $k$ such that*

$$f(a+h,b+k)=f(a,b)+hf_x(a,b)+kf_y(a,b)+h\epsilon(h,k)+k\delta(h,k) \tag{11.2}$$

*where*

$$\lim_{h\to 0}\epsilon(h)=0$$
$$\lim_{k\to 0}\delta(k)=0$$

***Definition 11.3*** *A function is said to be* **differentiable at P** *if it is locally linear at* **P**.

***Definition 11.4*** *A function is said to be* **differentiable on an open set** $R$ *if it is differentiable at every point in $R$.*

If we define the vectors

$$\mathbf{P}=(a,b)$$
$$\mathbf{h}=(h,k)$$
$$\mathbf{e}(h)=(\epsilon,\delta)$$

Then equation 11.2 becomes

$$f(\mathbf{P}+\mathbf{h})=f(\mathbf{P})+\mathbf{h}\cdot(f_x(\mathbf{P}),f_y(\mathbf{P}))+\mathbf{h}\cdot\mathbf{e}$$

Rearranging terms, dividing by $h=\|h\|$, and taking the limit as $\|\mathbf{h}\|\to 0$

$$\lim_{\|\mathbf{h}\|\to 0}\frac{f(\mathbf{P}+\mathbf{h})-f(\mathbf{P})}{\|\mathbf{h}\|}=\lim_{\|\mathbf{h}\|\to 0}\frac{\mathbf{h}\cdot(f_x(\mathbf{P}),f_y(\mathbf{P}))}{\|\mathbf{h}\|}+\lim_{\|\mathbf{h}\|\to 0}\frac{\mathbf{h}\cdot\mathbf{e}}{\|\mathbf{h}\|}$$

Defining the unit vector $\hat{\mathbf{h}}=\mathbf{h}/\|\mathbf{h}\|$,

$$\lim_{\|\mathbf{h}\|\to 0}\frac{f(\mathbf{P}+\mathbf{h})-f(\mathbf{P})}{\|\mathbf{h}\|}=\lim_{\|\mathbf{h}\|\to 0}\hat{\mathbf{h}}\cdot(f_x(\mathbf{P}),f_y(\mathbf{P}))+\lim_{\|\mathbf{h}\|\to 0}\hat{\mathbf{h}}\cdot\mathbf{e}$$

The first limit on the right does not depend on the length $\|\mathbf{h}\|$ because $\mathbf{h}$ only appears as a unit vector, which has length 1, so that

$$\lim_{\|\mathbf{h}\|\to 0}\frac{f(\mathbf{P}+\mathbf{h})-f(\mathbf{P})}{\|\mathbf{h}\|}=\hat{\mathbf{h}}\cdot(f_x(\mathbf{P}),f_y(\mathbf{P}))+\lim_{\|\mathbf{h}\|\to 0}\hat{\mathbf{h}}\cdot\mathbf{e}$$

The second limit on the right depends on $\mathbf{h}$ through the vector $\mathbf{e}$; but the components of the vector $\mathbf{e}$ approach 0 as the components of $\mathbf{h}$ approach 0. Since the dot product of a vector of length 1 with a vector length 0 is zero,

$$\lim_{\|\mathbf{h}\|\to 0}\frac{f(\mathbf{P}+\mathbf{h})-f(\mathbf{P})}{\|\mathbf{h}\|}=\hat{\mathbf{h}}\cdot(f_x(\mathbf{P}),f_y(\mathbf{P})) \tag{11.3}$$

This gives us a generalized definition of the derivative in the direction of any vector $\mathbf{h}$. We first define the gradient vector of functions of two and three variables, which we will use heavily in the remainder of this course.

**Definition 11.5** *The* **gradient** *of a function $f(x,y): \mathbb{R}^2 \mapsto \mathbb{R}$ of two variables is given by*

$$\boxed{\nabla f(x,y) = \text{grad} f(x,y) = \mathbf{i}\frac{\partial f}{\partial x} + \mathbf{j}\frac{\partial f}{\partial y}} \tag{11.4}$$

*The* **gradient** *of a function $f(x,y,z): \mathbb{R}^3 \mapsto \mathbb{R}$ of three variables is*

$$\boxed{\nabla f(x,y) = \text{grad} f(x,y) = \mathbf{i}\frac{\partial f}{\partial x} + \mathbf{j}\frac{\partial f}{\partial y} + \mathbf{k}\frac{\partial f}{\partial z}} \tag{11.5}$$

Note that if we apply the definition in three-dimensions to a function of two variables, we obtain the same result as the first definition, because the partial derivative of $f(x,y)$ with respect to $z$ is zero ($z$ does not appear in the equation).

**Example 11.1** *Find the gradient of $f(x,y) = \sin^3(x^2y)$*

*Solution.*

$$\begin{aligned}
\nabla f &= \mathbf{i}\frac{\partial f}{\partial x} + \mathbf{j}\frac{\partial f}{\partial y} = \mathbf{i}\frac{\partial}{\partial x}\sin^3(x^2y) + \mathbf{j}\frac{\partial}{\partial y}\sin^3(x^2y) \\
&= \mathbf{i}\left[3\sin^2(x^2y))\frac{\partial}{\partial x}\sin(x^2y)\right] + \mathbf{j}\left[3\sin^2(x^2y))\frac{\partial}{\partial y}\sin(x^2y)\right] \\
&= \mathbf{i}\left[3\sin^2(x^2y))\cos(x^2y)(2xy)\right] + \mathbf{j}\left[3\sin^2(x^2y))\cos(x^2y)(x^2)\right] \\
&= 3x\sin^2\left(x^2y\right) cos\left(x^2y\right)(2y\mathbf{i} + x\mathbf{j}) \blacksquare
\end{aligned}$$

**Example 11.2** *Find the gradient of $f(x,y) = x^2y + y^2z + z^2x$*

*Solution.*

$$\begin{aligned}
\nabla f &= \mathbf{i}\frac{\partial f}{\partial x} + \mathbf{j}\frac{\partial f}{\partial y} + \mathbf{k}\frac{\partial f}{\partial z} \\
&= \mathbf{i}\frac{\partial}{\partial x}\left(x^2y + y^2z + z^2x\right) + \mathbf{j}\frac{\partial}{\partial y}\left(x^2y + y^2z + z^2x\right) + \\
&\qquad \mathbf{k}\frac{\partial}{\partial z}\left(x^2y + y^2z + z^2x\right) \\
&= \mathbf{i}(2xy + z^2) + \mathbf{j}(x^2 + 2yz) + \mathbf{k}(y^2 + 2zx) \blacksquare
\end{aligned}$$

**Theorem 11.1 Properties of the Gradient Vector** *Suppose that $f$ and $g$ are functions and $c$ is a constant. Then the following are true:*

$$\nabla[f+g] = \nabla f + \nabla g \tag{11.6}$$

$$\nabla(cf) = c\nabla f \tag{11.7}$$

$$\nabla(fg) = f\nabla g + g\nabla f \tag{11.8}$$

Returning to equation 11.3, let $\mathbf{u} = u\mathbf{h}/h$ where $h = \|\mathbf{h}\|$,i.e., any vector that is parallel to $\mathbf{h}$ but has magnitude $u$. Then

$$\lim_{h\to 0} \frac{f(\mathbf{P} + (h/u)\mathbf{u}) - f(\mathbf{P})}{h} = \frac{\mathbf{u}\cdot\nabla f(\mathbf{P})}{u} \tag{11.9}$$

Since $u$ is finite (but fixed), then $h/u \to 0$ as $h \to 0$, so that

$$\mathbf{u}\cdot\nabla f(\mathbf{P}) = \lim_{h/u\to 0} \frac{f(\mathbf{P} + (h/u)\mathbf{u}) - f(\mathbf{P})}{h/u} \tag{11.10}$$

Finally, if we define a new $k = h/u$ then we can make the following generalization of the derivative.

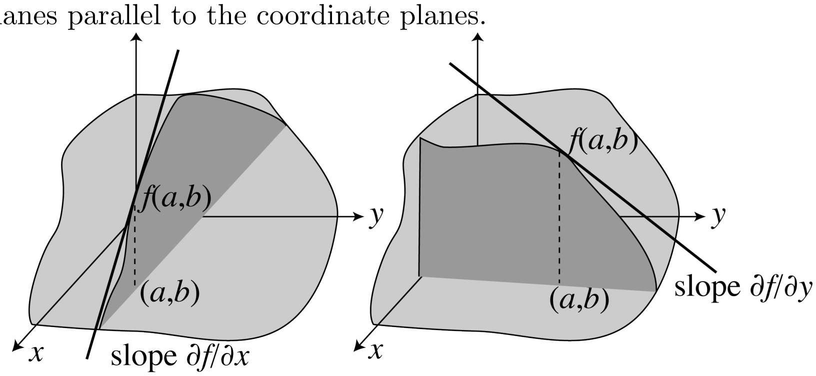
**Definition 11.6** *The* **directional derivative of $f$ in the direction of $\mathbf{u}$ at $\mathbf{P}$** *is given by*

$$\boxed{\mathbf{D_u} f(\mathbf{P}) = \mathbf{u}\cdot\nabla f(\mathbf{P}) = \lim_{k\to 0} \frac{f(\mathbf{P} + k\mathbf{u}) - f(\mathbf{P})}{k}} \tag{11.11}$$

**Theorem 11.2** *The following are equivalent:*

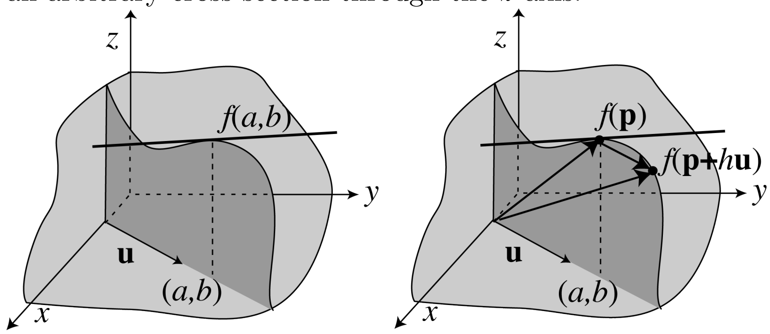
1. *$f(x, y)$ is differentiable at* $\mathbf{P}$
2. *$f(x, y)$ is locally linear at* $\mathbf{P}$
3. *$\mathbf{D_u} f(\mathbf{P})$ is defined at* $\mathbf{P}$
4. *The partial derivatives $\partial f/\partial x$ and $\partial f/\partial y$ exists and are continuous in some neighborhood of* $\mathbf{P}$.

Figure 11.1: Geometric interpretation of partial derivatives as the slopes of a function in planes parallel to the coordinate planes.




**Example 11.3** *Find the directional derivative of $f(x, y) = 2x^2 + xy - y^2$ at $\mathbf{p} = (3, -2)$ in the direction $\mathbf{a} = \mathbf{i} - \mathbf{j}$.*

Figure 11.2: Geometric interpretation of directional derivatives as the slope of a function in an arbitrary cross-section through the $z$-axis.




*Solution.* We need to calculate $D_{\mathbf{a}}f(p) = \mathbf{a}\cdot\nabla f$. But

$$\begin{aligned}\nabla f(x,y) &= \nabla(2x^2+xy-y^2)\\ &= \mathbf{i}\frac{\partial}{\partial x}(2x^2+xy-y^2)+\mathbf{j}\frac{\partial}{\partial y}(2x^2+xy-y^2)\\ &= \mathbf{i}(4x+y)+\mathbf{j}(x-2y)\end{aligned}$$

Hence

$$\begin{aligned}D_{\mathbf{a}}f(p) &= \mathbf{a}\cdot\nabla f = (\mathbf{i}-\mathbf{j})\cdot(\mathbf{i}(4x+y)+\mathbf{j}(x-2y))\\ &= [(1)(4x+y)+(-1)(x-2y)]\\ &= [4x+y-x+2y] = 3(x+y)\end{aligned}$$

At the point $\mathbf{p}$=(3,-2) $D_{\mathbf{a}}f(\mathbf{p}) = 3(3-2) = 3.$■

**Example 11.4** *Find the directional derivative of*$f(x,y,z) = x^3y - y^2z^2$ *at* $\mathbf{p} = (-2,1,3)$ *in the direction* $\mathbf{a} = \mathbf{i} - 2\mathbf{j} + 2\mathbf{k}$.

*Solution.*

$$\begin{aligned}\nabla f(x,y,z) &= \nabla(x^3y-y^2z^2)\\ &= \mathbf{i}\frac{\partial}{\partial x}(x^3y-y^2z^2)+\mathbf{j}\frac{\partial}{\partial y}(x^3y-y^2z^2)+\mathbf{k}\frac{\partial}{\partial z}(x^3y-y^2z^2)\\ &= \mathbf{i}(3x^2y)+\mathbf{j}(x^3-2yz^2)+\mathbf{k}(-2y^2z)\end{aligned}$$

and therefore

$$\mathbf{a}\cdot\nabla f = 3x^y - 2x^3 + 4yz^2 - 4y^z$$

The directional derivative at $\mathbf{p} = (-2,1,3)$ is

$$D_{\mathbf{a}}f(\mathbf{p}) = 3(-2)^2(1) - 2(-2)^3 + 4(1)(3)^2 - 4(1)^2(3) = 52 \blacksquare$$

**Note**: Some books use the notation

$$\frac{\partial f}{\partial \mathbf{u}} = D_{\mathbf{u}}f(\mathbf{p})$$

for the directional derivative. The following example justifies this odd notation.

**Example 11.5** *Simplify the expression*

$$\frac{\partial f}{\partial \mathbf{u}} = D_{\mathbf{u}} f(p)$$

*when* $\mathbf{u} = \mathbf{k}$, *i.e., a unit vector in the z-direction.*

*Solution.*

$$\begin{aligned}
D_{\mathbf{u}} f(p) &= \mathbf{u} \cdot \nabla f = \mathbf{k} \cdot \nabla f = \mathbf{k} \cdot \left( \mathbf{i}\frac{\partial f}{\partial x} + \mathbf{j}\frac{\partial f}{\partial y} + \mathbf{k}\frac{\partial f}{\partial z} \right) \\
&= \mathbf{k} \cdot \mathbf{i}\frac{\partial f}{\partial x} + \mathbf{k} \cdot \mathbf{j}\frac{\partial f}{\partial y} + \mathbf{k} \cdot \mathbf{k}\frac{\partial f}{\partial z} \\
&= (0)\frac{\partial f}{\partial x} + (0)\frac{\partial f}{\partial y} + (1)\frac{\partial f}{\partial z} \\
&= \frac{\partial f}{\partial z} = D_z f \quad \blacksquare
\end{aligned}$$

**Theorem 11.3** *The directional derivatives along a coordinate axis is the partial derivative with respect to that axis:*

$$D_{\mathbf{i}} f = \frac{\partial f}{\partial x} = f_x$$

$$D_{\mathbf{j}} f = \frac{\partial f}{\partial y} = f_y$$

$$D_{\mathbf{k}} f = \frac{\partial f}{\partial z} = f_z$$

**Theorem 11.4** *A function* $f(x, y)$ *increases the most rapidly (as you move away from* $\mathbf{p}$=(x,y)*) in the direction of the gradient. The magnitude of the rate of change is given by* $\|\nabla f(\mathbf{p})\|$. *The function decreases most rapidly in the direction of* $-\nabla f(\mathbf{p})$, *with magnitude* $-\|\nabla f(\mathbf{p})\|$.

*Proof.* The directional derivative, which gives the rate of change in *any* direction $\mathbf{u}$ at $\mathbf{p}$ is

$$D_{\mathbf{u}} f = \mathbf{u} \cdot \nabla f = |\mathbf{u}|\, |\nabla f| \cos\theta$$

Consider the set of all possible unit vectors $\mathbf{u}$ emanating from $\mathbf{p}$. Then the maximum of the directional derivative occurs when $\cos\theta = 1$,i.e., when $\mathbf{u}$ is parallel to $\nabla f$, and

$$\max(D_{\mathbf{u}} f(\mathbf{p})) = \|\nabla f(\mathbf{p})\|$$

Similarly the rate of maximum decrease occurs when $\cos\theta = -1$ $\blacksquare$

**Example 11.6** *Find a unit vector in the direction in which* $f(x, y) = e^y \sin x$ *increases most rapidly at* $\mathbf{p} = (5\pi/6, 0)$

*Solution.* The gradient vector is

$$\begin{aligned}\nabla f &= \nabla(e^y \sin x) = \left(\frac{\partial}{\partial x} e^y \sin x, \frac{\partial}{\partial y} e^y \sin x\right) = (e^y \cos x, e^y \sin x) \\ &= e^y (\cos x, \sin x)\end{aligned}$$

At $\mathbf{p} = (5\pi/6, 0)$, we find that the direction of steepest increase is

$$\nabla f(\mathbf{p}) = e^0 (\cos(5\pi/6), \sin(5\pi/6)) = \left(-\sqrt{3}/2, 1/2\right)$$

and the magnitude of the increase is

$$\|\nabla f(\mathbf{p})\| = \sqrt{(\sqrt{3}/2)^2 + (1/2)^2} = \sqrt{3/4 + 1/4} = 1$$

Since $\mathbf{u} = \nabla f(\mathbf{p})$ has a magnitude of 1, it is a unit vector in the direction that $f$ increases most rapidly.■

**Example 11.7** *Suppose that the temperature $T(x, y, z)$ of a ball of some material centered at the origin is given by the function*

$$T(x, y, z) = \frac{100}{10 + x^2 + y^2 + z^2}$$

*Starting at the point (1, 1, 1), in what direction must you move to obtain the greatest increase in temperature?*

*Solution.* We need to move in the direction of the gradient.

$$\begin{aligned}\nabla T(x, y, z) &= \mathbf{i}\frac{\partial T}{\partial x} + \mathbf{j}\frac{\partial T}{\partial y} + \mathbf{k}\frac{\partial T}{\partial z} \\ &= \mathbf{i}\frac{\partial}{\partial x}\frac{100}{10 + x^2 + y^2 + z^2} + \mathbf{j}\frac{\partial}{\partial y}\frac{100}{10 + x^2 + y^2 + z^2} \\ &\quad + \mathbf{k}\frac{\partial}{\partial z}\frac{100}{10 + x^2 + y^2 + z^2}\end{aligned}$$

The partial derivative with respect to $x$ is

$$\begin{aligned}\frac{\partial}{\partial x}\frac{100}{10 + x^2 + y^2 + z^2} &= 100\frac{\partial}{\partial x}(10 + x^2 + y^2 + z^2)^{-1} \\ &= 100(-1)(10 + x^2 + y^2 + z^2)^{-2}\frac{\partial}{\partial x}(10 + x^2 + y^2 + z^2) \\ &= \frac{-200x}{(10 + x^2 + y^2 + z^2)^2}\end{aligned}$$

By symmetry,

$$\begin{aligned}\frac{\partial}{\partial y}\frac{100}{10 + x^2 + y^2 + z^2} &= \frac{-200y}{(10 + x^2 + y^2 + z^2)^2} \\ \frac{\partial}{\partial z}\frac{100}{10 + x^2 + y^2 + z^2} &= \frac{-200z}{(10 + x^2 + y^2 + z^2)^2}\end{aligned}$$

Hence

$$\nabla T(x,y,z) \quad = \quad -\frac{200(x\mathbf{i}+y\mathbf{j}+z\mathbf{k})}{(10+x^2+y^2+z^2)^2}$$

At the point (1,1,1) we have

$$\frac{\partial}{\partial x}\left.\frac{-200x}{10+x^2+y^2+z^2}\right|_{(1,1,1)} = \frac{-200}{169}$$

and similarly for the other two partial derivatives. Hence the direction of steepest increase in temperature is

$$\nabla T(x,y,z)|_{(1,1,1)} = -\frac{200}{169}(\mathbf{i}+\mathbf{j}+\mathbf{k})\blacksquare$$

**Theorem 11.5** *The gradient vector is perpendicular to the level curves of a function.*

To see why this theorem is true, recall the definition of a level curve: it is a curve along which the value of $z = f(x,y)$ is a constant. Thus if you move along a tangent vector to a level curve, the function will not change, and the directional derivative is zero. Since the directional derivative is the dot product of the gradient and the direction of motion, which in this case is the tangent vector to the level curve, this dot product is zero. But a dot product of two non-zero vectors can only be zero if the two vectors are perpendicular to one another. Thus the gradient must be perpendicular to the level curve.

**Corollary 11.1** *Let $f(x,y,z) = 0$ describe a surface in 3D space. Then the three-dimensional gradient vector $\nabla f$ is a normal vector to the surface.*

This result follows because a surface $f(x,y,z) = 0$ is a level surface of a function $w = f(x,y,z)$. The value of $f(x,y,z)$ does not change. So if you consider any tangent vector, the directional derivative along the tangent vector is zero. Consequently the tangent vector is perpendicular to the gradient.

# Lecture 12

# The Chain Rule

Recall the chain rule from Calculus I: To find the derivative of some function $u = f(x)$ with respect to a new variable $t$, we calculate

$$\frac{du}{dt} = \frac{du}{dx}\frac{dx}{dt}$$

We can think of this as a sequential process, as illustrated in figure 12.1.

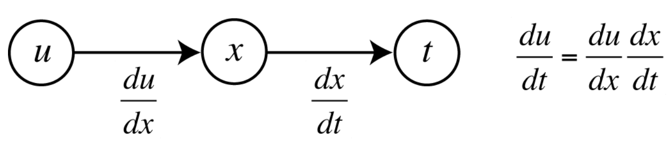
1. Draw a labeled node for each variable. The original function should be the furthest to the left, and the desired final variable the node furthest to the right.

2. Draw arrows connecting the nodes.

3. Label each arrow with a derivative. The variable on the top of the derivative corresponds to the variable the arrow is coming *from* and the variable on the bottom of the derivative is the variable the arrow is going *to*.

4. follow the path described by the labeled arrows. The derivative of the variable at the start of the path with respect to the variable at the end of the path is the product of the derivatives you meet along the way.

Figure 12.1: Visualization of the chain rule for a function of a single variable

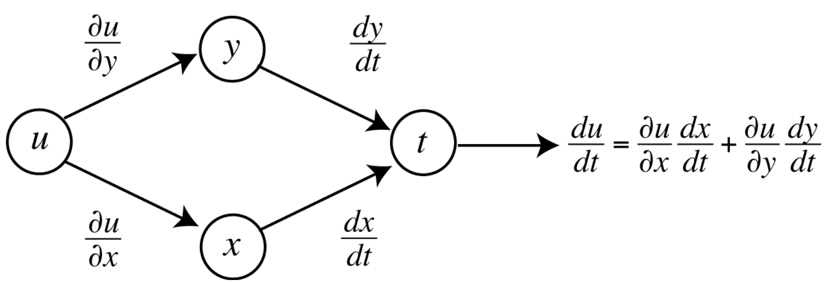



Now suppose $u$ is a function of two variables $x$ and $y$ rather than just one. Then we need to have two arrows emanating from $u$, one to each variable. This is illustrated in figure 12.2. The procedure is modified as follows:

1. Draw a node for the original function $u$, each variable it depends on $x, y, ...$, and the final variable that we want to find the derivative with respect to $t$.

2. Draw an arrow from the $u$-node to each of its variable nodes, and then draw an arrow from each of the variable nodes to the final node. Observe that by following the arrows there are now two possible paths we can follow.
3. Label the arrows emanating from the original function with partial derivatives. Label the arrows emanating from the variables themselves to the final variable with ordinary derivatives.
4. Consider each possible pat from $u$ to $t$, and from the product of derivatives as you follow the path. The top path in figure 12.1, for example, gives
$$\frac{\partial u}{\partial y}\frac{dy}{dt}$$
and the bottom path gives
$$\frac{\partial u}{\partial x}\frac{dx}{dt}$$
5. Add together the products for each possible path. This gives
$$\frac{d}{dt}u(x,y) = \frac{\partial u}{\partial y}\frac{dy}{dt} + \frac{\partial u}{\partial x}\frac{dx}{dt}$$

Figure 12.2: Chain rule for a function of two variables




This method has an obvious generalization if there are more variables; for example, if $u = u(x, y, z)$, we add a third node labeled $z$ and another path from $u$ to $t$ through $z$. The result is

$$\frac{d}{dt}u(x,y,z) = \frac{\partial u}{\partial y}\frac{dy}{dt} + \frac{\partial u}{\partial x}\frac{dx}{dt} + \frac{\partial u}{\partial z}\frac{dz}{dt}$$

**Example 12.1** *Let $w = x^4y^2, x = \sin t, y = t^2$. Find $dw/dt$ using the chain rule.*

*Solution.*

$$\begin{aligned}\frac{dw}{dt} &= \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt} \\ &= \frac{\partial}{\partial x}(x^4y^2)\frac{d}{dt}(\sin t) + \frac{\partial}{\partial y}(x^4y^2)\frac{d}{dt}(t^2) \\ &= (4x^3y^2)(\cos t) + x^4(2y)(2t)\end{aligned}$$

The problem is not finished because the answer we have derived depends on $x$, $y$, and $t$, and should only depend on $t$. To finish the problem, we need to substitute the expressions we are given for $x$ and $y$ as a function of $t$, namely $x = \sin t$ and $y = t^2$. This leads to

$$\begin{aligned}
\frac{dw}{dt} &= 4x^3y^2\cos t + 4x^4yt \\
&= 4(\sin t)^3(t^2)^2\cos t + 4(\sin t)^4(t^2)t \\
&= 4t^4\sin^3 t\cos t + 4t^2\sin^4 t \\
&= 4t^2\sin^3 t(t^2\cos t + \sin t) \quad \blacksquare
\end{aligned}$$

**Example 12.2** *Find $du/dr$, as a function of $r$, using the chain rule for $u = z\sqrt{x+y}$, $x = e^{3r}$, $y = 12r$ and $z = \ln r$.*

*Solution.*

$$\begin{aligned}
\frac{du}{dr} &= \frac{\partial u}{\partial x}\frac{dx}{dr} + \frac{\partial u}{\partial y}\frac{dy}{dr} + \frac{\partial u}{\partial z}\frac{dy}{dr} \\
&= \frac{\partial}{\partial x}\left(z\sqrt{x+y}\right)\frac{d}{dr}(e^{3r}) + \frac{\partial}{\partial y}\left(z\sqrt{x+y}\right)\frac{dy}{dr}(12r) \\
&\qquad + \frac{\partial}{\partial z}\left(z\sqrt{x+y}\right)\frac{d}{dr}(\ln r) \\
&= \left[z(1/2)(x+y)^{-1/2}\right]\left[e^{3r}(3)\right] + \left[z(1/2)(x+y)^{-1/2}\right](12) \\
&\qquad + \left(\sqrt{x+y}\right)(1/r) \\
&= \frac{3ze^{3r}}{2\sqrt{x+y}} + \frac{12z}{2\sqrt{x+y}} + \frac{\sqrt{x+y}}{r} \\
&= \frac{3ze^{3r}}{2\sqrt{x+y}} + \frac{6z}{\sqrt{x+y}} + \frac{\sqrt{x+y}}{r} \\
&= \frac{3\ln(r)e^{3r}}{2\sqrt{e^{3r}+12r}} + \frac{6\ln r}{\sqrt{e^{3r}+12r}} + \frac{\sqrt{e^{3r}+12r}}{r} \quad \blacksquare
\end{aligned}$$

**Example 12.3** *Find $dw/dt$, as a function of $t$, using the chain rule, for $w = xy + yz + xz$, $x = t^2$, $y = 1 - t^2$, and $z = 1 - t$.*

*Solution.*

$$\begin{aligned}
\frac{dw}{dt} &= \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt} + \frac{\partial w}{\partial z}\frac{dy}{dt} \\
&= \frac{\partial(xy+yz+xz)}{\partial x}\frac{d(t^2)}{dt} + \frac{\partial(xy+yz+xz)}{\partial y}\frac{d(1-t^2)}{dt} \\
&\quad + \frac{\partial(xy+yz+xz)}{\partial z}\frac{d(1-t)}{dt} \\
&= (y+z)(2t) + (x+z)(-2t) + (y+x)(-1) \\
&= 2t(y+z-x-z) - y - x \\
&= 2t(y-x) - y - x \\
&= 2t(1-t^2-t^2) - (1-t^2) - t^2 \\
&= 2t(1-2t^2) - 1 \\
&= 2t - 4t^3 - 1 \quad \blacksquare
\end{aligned}$$

**Example 12.4** *Find $\partial w/\partial t$ and $\partial w/\partial s$ using the chain rule and express the result as a function of $s$ and $t$, for $w = 5x^2 - y\ln x$, $x = s+t$, and $y = s^3 t$.*

*Solution.* Here $w$ only depends on two variables $x$ and $y$, so the chain rules become

$$\frac{\partial w}{\partial t} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial t}$$

$$\frac{\partial w}{\partial s} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial s}$$

The second factor in each term is a partial derivative because $x$ and $y$ are functions of two variables, $s$ and $t$, and not just functions of a single variable. From the first equation,

$$\begin{aligned}
\frac{\partial w}{\partial t} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial t} \\
&= \frac{\partial(5x^2 - y\ln x)}{\partial x}\frac{\partial(s+t)}{\partial t} + \frac{\partial(5x^2 - y\ln x)}{\partial y}\frac{\partial(s^3 t)}{\partial t} \\
&= (10x - y/x)(1) + (-\ln x)(s^3) \\
&= 10(s+t) - \frac{s^3 t}{s+t} - [\ln(s+t)]s^3
\end{aligned}$$

Similarly,

$$\begin{aligned}
\frac{\partial w}{\partial s} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial s} \\
&= \frac{\partial(5x^2 - y\ln x)}{\partial x}\frac{\partial(s+t)}{\partial s} + \frac{\partial(5x^2 - y\ln x)}{\partial y}\frac{\partial(s^3 t)}{\partial s} \\
&= (10x - y/x)(1) + (-\ln x)(3s^2 t) \\
&= 10(s+t) - \frac{s^3 t}{s+t} - 3s^2 t\ln(s+t) \quad \blacksquare
\end{aligned}$$

**Example 12.5** *Find $\partial w/\partial t$ and $\partial w/\partial s$ using the chain rule and express the result as a function of $s$ and $t$, where*

$$w = x^3 + y + z^2 + xy$$

$$x = st$$

$$y = s + t$$

$$z = s + 4t$$

*Solution.* Now $w$ is a function of 3 variables: $x$, $y$ and $z$, so there are three terms in each of the chain rule formulas. All of the derivatives are partial because we are asked to find the partial derivative:

$$\frac{\partial w}{\partial t} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial t}$$

$$\frac{\partial w}{\partial s} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial s}$$

To find $\partial w/\partial t$ we calculate

$$\begin{aligned}
\frac{\partial w}{\partial t} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial t} \\
&= \frac{\partial(x^3 + y + z^2 + xy)}{\partial x}\frac{\partial(st)}{\partial t} + \frac{\partial(x^3 + y + z^2 + xy)}{\partial y}\frac{\partial(s+t)}{\partial t} \\
&\quad + \frac{\partial(x^3 + y + z^2 + xy)}{\partial z}\frac{\partial(s+4t)}{\partial t} \\
&= (3x^2 + y)(s) + (1 + x)(1) + (2z)(4) \\
&= s(3x^2 + y) + 1 + x + 8z
\end{aligned}$$

Substituting for $x$, $y$, and $z$ gives

$$\begin{aligned}
\frac{\partial w}{\partial t} &= s(3x^2 + y) + 1 + x + 8z \\
&= s\left(3(st)^2 + (s+t)\right) + 1 + st + 8(s + 4t) \\
&= 3s^3t^2 + 3s + 3t + 1 + st + 8s + 32t \\
&= 3s^3t^2 + 3s + 35t + 1 + st + 8s
\end{aligned}$$

To find $\partial w/\partial s$, we similarly calculate

$$\begin{aligned}
\frac{\partial w}{\partial s} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial s} \\
&= \frac{\partial(x^3 + y + z^2 + xy)}{\partial x}\frac{\partial(st)}{\partial s} + \frac{\partial(x^3 + y + z^2 + xy)}{\partial y}\frac{\partial(s+t)}{\partial s} \\
&\quad + \frac{\partial(x^3 + y + z^2 + xy)}{\partial z}\frac{\partial(s+4t)}{\partial s} \\
&= = (3x^2 + y)(t) + (1 + x)(1) + 2z(1) \\
&= = (3x^2 + y)t + 1 + x + 2z
\end{aligned}$$

Substituting for $x$, $y$, and $z$,

$$\begin{aligned}
\frac{\partial w}{\partial s} &= \left(3(st)^2 + s + t\right) t + 1 + st + 2(s+4t) \\
&= 3s^2t^3 + st + t^2 + st + 2s + 8t \\
&= 3s^2t^3 + 2st + t^2 + 2s + 8t \quad \blacksquare
\end{aligned}$$

## Generalization to Higher Dimensions

If $f$ is a function of multiple variables,

$$f(u, v, w, x, y, z, ...) = f(u(t), v(t), w(t), x(t), ...)$$

The figures are generalized in the obvious way; in between the node for $f$ and the node for $t$ we put nodes for each of $u$, $v$, $w$, $x$,... and so forth, and draw arrows as before. There are still only two derivatives in each path, but there are a whole lot more paths, and we have to add up the products over each path. The result is

$$\frac{df}{dt} = \frac{\partial f}{\partial u}\frac{du}{dt} + \frac{\partial f}{\partial v}\frac{dv}{dt} + \frac{\partial f}{\partial w}\frac{dw}{dt} + \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt} + \cdots$$

Usually mathematicians use an indexed variable when the number of variables becomes large, and write a function of $n$-variables as

$$f(x_1, x_2, x_3, ..., x_n)$$

This is considered a function in $n$-dimensional space, sometimes called $\mathbb{R}^n$. The chain rule for a function in $\mathbb{R}^n$ is

$$\frac{df}{dt} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\frac{dx_i}{dt}$$

## Relationship of Chain Rule to the Gradient Vector and the Directional Derivative

If $x$, $y$, and $z$ are functions of a parameter $t$, then the position vector

$$r(t) = (x(t), y(t), z(t))$$

traces out a curve in three dimensions. The rate of change of any function $f(x,y,z)$ as a particle moves along this curve is $df/dt$. Using the chain rule, this derivative is

$$\begin{aligned}
\frac{df}{dt} &= \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt} \\
&= \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right) \cdot \left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}\right) \\
&= \nabla f \cdot \frac{d}{dt}\mathbf{r}(t) \\
&= \mathbf{v} \cdot \nabla f
\end{aligned}$$

because the velocity vector has been defined as $\mathbf{v} = \mathbf{r}'(t)$. Thus

$$\frac{df}{dt} = \mathbf{v}(t) \cdot \nabla f = D_{\mathbf{v}} f(t)$$

where $v = \|\mathbf{v}(t)\|$ is the speed at time $t$. Thus the total rate of change of a function as you move along a curve $\mathbf{r}(t)$ is the directional derivative in the direction of the tangent (or velocity) vector.

## Proof of the Chain Rule

We outline the proof for a function $f(x(t),\, y(t))$ of two variables; the generalization to higher dimensions is the same.The derivative of $f$ with respect to $t$ is

$$\frac{df}{dt} = \lim_{\Delta t \to 0} \frac{f(t+\Delta t) - f(t)}{\Delta t}$$

But

$$f(t+\Delta t) = f(x(t+\Delta t), y(t+\Delta t))$$

so that

$$\begin{aligned} f(t+\Delta t) - f(t) &= f(x(t+\Delta t), y(t+\Delta t)) - f(x(t), y(t)) \\ &= f(x(t+\Delta t), y(t+\Delta t)) - f(x(t), y(t+\Delta t)) \\ &\qquad + f(x(t), y(t+\Delta t)) - f(x(t), y(t)) \end{aligned}$$

This is allowed because the two middle terms add to zero. Let

$$\Delta x = x(t+\Delta t) - x(t), \Delta y = y(t+\Delta t) - y(t)$$

Then

$$\begin{aligned} f(t+\Delta t) - f(t) &= f(x+\Delta x, y+\Delta y) - f(x, y+\Delta y) \\ &\qquad + f(x, y+\Delta y) - f(x(t), y(t)) \end{aligned}$$

Consequently

$$\begin{aligned}
\frac{f(t+\Delta t)-f(t)}{\Delta t} &= \frac{f(x+\Delta x,y+\Delta y)-f(x,y+\Delta y)+f(x,y+\Delta y)-f(x,y)}{\Delta t} \\
&= \frac{f(x+\Delta x,y+\Delta y)-f(x,y+\Delta y)}{\Delta t} \\
&\quad + \frac{f(x,y+\Delta y)-f(x,y)}{\Delta t} \\
&= \frac{f(x+\Delta x,y+\Delta y)-f(x,y+\Delta y)}{\Delta t}\frac{\Delta x}{\Delta x} \\
&\quad + \frac{f(x,y+\Delta y)-f(x,y)}{\Delta t}\frac{\Delta y}{\Delta y} \\
&= \frac{f(x+\Delta x,y+\Delta y)-f(x,y+\Delta y)}{\Delta x}\frac{\Delta x}{\Delta t} \\
&\quad + \frac{f(x,y+\Delta y)-f(x,y)}{\Delta y}\frac{\Delta y}{\Delta t} \\
&= \frac{f(x+\Delta x,y+\Delta y)-f(x,y+\Delta y)}{\Delta x}\frac{x(t+\Delta t)-x(t)}{\Delta t} \\
&\quad + \frac{f(x,y+\Delta y)-f(x,y)}{\Delta y}\frac{y(t+\Delta t)-y(t)}{\Delta t}
\end{aligned}$$

Therefore

$$\begin{aligned}
\frac{df}{dt} &= \lim_{\Delta t\to 0}\frac{f(t+\Delta t)-f(t)}{\Delta t} \\
&= \lim_{\Delta t\to 0}\frac{f(x+\Delta x,y+\Delta y)-f(x,y+\Delta y)}{\Delta x}\lim_{\Delta t\to 0}\frac{x(t+\Delta t)-x(t)}{\Delta t} \\
&\quad + \lim_{\Delta t\to 0}\frac{f(x,y+\Delta y)-f(x,y)}{\Delta y}\lim_{\Delta t\to 0}\frac{y(t+\Delta t)-y(t)}{\Delta t} \\
&= \frac{\partial f}{\partial x}\frac{dx}{dt}+\frac{\partial f}{\partial y}\frac{dy}{dt} \quad \blacksquare
\end{aligned}$$

## Implicit Differentiation using the Chain Rule.

Suppose that $f(x,y)=0$ and we want to find $dy/dx$, but we can't solve for $y$ as a function of $x$. We can solve this problem using implicit differentiation, as we did for a function of a single variable.

**Example 12.6** *Find $dy/dx$ for $x^3+2x^2y-10y^5=0$.*

*Solution.* Using implicit differentiation, we differentiate both sides of the equation with respect to $x$ and solve for $dy/dx$

$$\begin{aligned}
& x^3 + 2x^2y - 10y^5 = 0 \\
\Rightarrow\ & \frac{d}{dx}(x^3 + 2x^2y - 10y^5) = \frac{d}{dx}(0) \\
\Rightarrow\ & \frac{d}{dx}(x^3) + 2\frac{d}{dx}(x^2y) - 10\frac{d}{dx}(y^5) = 0 \\
\Rightarrow\ & 3x^2 + 2\left[x^2\frac{dy}{dx} + y\frac{d}{dx}x^2\right] - 10(5)y^4\frac{dy}{dx} = 0 \\
\Rightarrow\ & 3x^2 + 2x^2\frac{dy}{dx} + 4xy - 50y^4\frac{dy}{dx} = 0
\end{aligned}$$

Now bring all the terms that have a $dy/dx$ in them to one side of the equation:

$$2x^2\frac{dy}{dx} - 50y^4\frac{dy}{dx} = -4xy - 3x^2$$

Factor $dy/dx$ on the left:

$$\frac{dy}{dx}(2x^2 - 50y^4) = -4xy - 3x^2$$

Solving for $dy/dx$

$$\frac{dy}{dx} = \frac{-4xy - 3x^2}{2x^2 - 50y^4} \quad \blacksquare$$

Now consider the same problem of finding $dy/dx$ when $f$ is a function of two parameterized variables $f(x(t), y(t))$.

$$0 = \frac{df(x(t), y(t))}{dx} = \frac{\partial f}{\partial x}\frac{dx}{dx} + \frac{\partial f}{\partial y}\frac{dy}{dx}$$

Since $dx/dx = 1$

$$\begin{aligned}
0 &= \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx} \\
\Rightarrow \frac{\partial f}{\partial y}\frac{dy}{dx} &= -\frac{\partial f}{\partial x}
\end{aligned}$$

Hence we have the following result.

**Theorem 12.1 Implicit Differentiation** *If* $f(x, y) = 0$ *then*

$$\boxed{\frac{dy}{dx} = -\frac{\partial f/\partial x}{\partial f/\partial y}}$$

By a similar argument, we also have

**Theorem 12.2 Implicit Differentiation** *If* $f(x, y, z) = 0$ *then*

$$\boxed{\frac{\partial z}{\partial x} = -\frac{\partial f/\partial x}{\partial f/\partial z}} \quad \text{and} \quad \boxed{\frac{\partial z}{\partial y} = -\frac{\partial f/\partial y}{\partial f/\partial z}}$$

**Example 12.7** *Repeat the previous example using theorem 1 instead of implicit differentiation.*

*Solution.* We had $f(x, y) = x^3 + 2x^2y - 10y^5$ so that

$$\frac{dy}{dx} = -\frac{\partial f/\partial x}{\partial f/\partial y} = -\frac{\frac{\partial}{\partial x}\left(x^3 + 2x^2y - 10y^5\right)}{\frac{\partial}{\partial y}\left(x^3 + 2x^2y - 10y^5\right)} = -\frac{3x^2 + 4xy}{2x^2 - 50y^4} \quad \blacksquare$$

**Example 12.8** *Find $dy/dx$ if $x^2 \cos y - y^2 \sin x = 0$.*

*Solution.*Writing $f(x, y) = x^2 \cos y - y^2 \sin x$ and usingtheorem 1,

$$\frac{dy}{dx} = -\frac{\partial f/\partial x}{\partial f/\partial y} = -\frac{\frac{\partial}{\partial x}\left(x^2 \cos y - y^2 \sin x\right)}{\frac{\partial}{\partial y}\left(x^2 \cos y - y^2 \sin x\right)} = -\frac{2x \cos y - y^2 \cos x}{-x^2 \sin y - 2y \sin x} \quad \blacksquare$$

## Differentials

**Definition 12.1** *The* **differential** *of a function $f(x, y, z)$ is the quantity*

$$df = \nabla f \cdot dr$$

*where*

$$dr = (dx, dy, dz)$$

*namely,*

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz$$

The differential gives an estimate of the change in the function $f(x, y, z)$ when $\mathbf{r}$ is perturbed by a small amount from $(x, y, z)$ to $(x + dx, y + dy, z + dz)$.

**Example 12.9** *Estimate the change in $f(x, y) = x^2 + y^2$ as you move from $(1, 1)$ to $(1.01, 1.01)$ using differentials, and compare with the exact change.*

*Solution.* We have

$$\Delta r = (0.01, 0.01)$$

and

$$\nabla f = (2x, 2y)$$

so that

$$\nabla f(1, 1) = (2, 2)$$

Therefore

$$\Delta f = \nabla f \cdot \Delta r = (2, 2) \cdot (0.01, 0.01) = .04$$

The exact change is

$$f(1.01, 1.01) - f(1, 1) = 1.01^2 + 1.01^2 - 1 - 1 = 0.0402 \quad \blacksquare$$

**Example 12.10** *The effective resistance $R$ of two resistors $R_1, R_2$in parallel is given by the formula*

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2}$$

*Suppose we have two resistors*

$$R_1 = 10 \pm 1 \ ohms$$

$$R_2 = 40 \pm 2 \ ohms$$

*that are connected in parallel. Estimate the total resistance $R$ and the uncertainty in $R$.*

*Solution.* We first calculate R,

$$\frac{1}{R} = \frac{1}{10} + \frac{1}{40} = \frac{5}{40} = \frac{1}{8} \Rightarrow R = 8$$

To calculate the uncertainty we solve for $R(R_1, R_2)$ and find the differential.

$$R = \frac{R_1 R_2}{R_1 + R_2}$$

Hence

$$dR = \frac{\partial}{dR_1}\frac{R_1 R_2}{R_1 + R_2} dR_1 + \frac{\partial}{dR_2}\frac{R_1 R_2}{R_1 + R_2} dR_2$$

By the quotient rule,

$$\begin{aligned}
\frac{\partial}{dR_1}\frac{R_1 R_2}{R_1 + R_2} &= \frac{(R_1 + R_2)\frac{\partial}{\partial R_1}(R_1 R_2) - (R_1 R_2)\frac{\partial}{\partial R_1}(R_1 + R_2)}{(R_1 + R_2)^2} \\
&= \frac{(R_1 + R_2) R_2 - R_1 R_2}{(R_1 + R_2)^2} \\
&= \frac{R_2^2}{(R_1 + R_2)^2} = \frac{40^2}{(10 + 40)^2} = \left(\frac{40}{50}\right)^2 = \left(\frac{4}{5}\right)^2 = \frac{16}{25} \\
\frac{\partial}{dR_2}\frac{R_1 R_2}{R_1 + R_2} &= \frac{(R_1 + R_2)\frac{\partial}{\partial R_2}(R_1 R_2) - (R_1 R_2)\frac{\partial}{\partial R_2}(R_1 + R_2)}{(R_1 + R_2)^2} \\
&= \frac{(R_1 + R_2) R_1 - R_1 R_2}{(R_1 + R_2)^2} \\
&= \frac{R_1^2}{(R_1 + R_2)^2} = \frac{10^2}{(10 + 40)^2} = \left(\frac{10}{50}\right)^2 = \left(\frac{1}{5}\right)^2 = \frac{1}{25}
\end{aligned}$$

Thus

$$dR = \frac{\partial}{dR_1}\frac{R_1 R_2}{R_1 + R_2} dR_1 + \frac{\partial}{dR_2}\frac{R_1 R_2}{R_1 + R_2} dR_2 = \frac{16}{25}(1) + \frac{1}{25}(2) = \frac{18}{25} = 0.72$$

and we conclude that $R = 8 \pm 0.72$ ohms ■

# Lecture 13

# Tangent Planes

Since the gradient vector is perpendicular to a surface in three dimensions, we can find the tangent plane by constructing the locus of all points perpendicular to the gradient vector. In fact, we will *define* the tangent plane in terms of the gradient vector.

**Definition 13.1** *The* **tangent plane** *to a surface* $f(x, y, z) = k$ *at a point* $\mathbf{P}_0 = (x_0, y_0, z_0)$ *is the plane perpendicular to the gradient vector at* $p_0$.

To get a formula for the tangent plane, let $\mathbf{P} = (x, y, z)$ run over all points in the plane. Then any vector of the form

$$\mathbf{P} - \mathbf{P}_0 = (x - x_0, y - y_0, z - z_0) = (x - x_0)\mathbf{i} + (y - y_0)\mathbf{j} + (z - z_0)\mathbf{k}$$

Since $\nabla f(x_0, y_0, z_0)$ is normal to the tangent plane, the critical equation defining the tangent plane is

$$\boxed{(\mathbf{P} - \mathbf{P}_0) \cdot \nabla f(x_0, y_0, z_0) = 0} \tag{13.1}$$

**Example 13.1** *Find the equation of the plane tangent to the surface*

$$3x^2 + y^2 - z^2 = -20$$

*at the point* $\mathbf{P}_0 = (1, 2, 3)$.

*Solution.* First, we rewrite the equation of the surface into the form $f(x, y, z) = 0$, as

$$f(x, y, z) = 3x^2 + y^2 - z^2 + 20 = 0$$

The general form of the gradient vector at any point on the surface is then

$$\nabla f = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} + \frac{\partial f}{\partial z}\mathbf{k} = 6x\mathbf{i} + 2y\mathbf{j} - 2z\mathbf{k}$$

At the point $\mathbf{P}_0 = (1, 2, 3)$ the gradient is

$$\nabla f(\mathbf{P}_0) = 6(1)\mathbf{i} + 2(2)\mathbf{j} - 2(3)\mathbf{k} = 6\mathbf{i} + 4\mathbf{j} - 6\mathbf{k}$$

A general vector on the plane is

$$\mathbf{P}-\mathbf{P}_0=(x-1)\mathbf{i}+(y-2)\mathbf{j}+(z-3)\mathbf{k}$$

Hence

$$\begin{aligned} 0 &= (\mathbf{P}-\mathbf{P}_0)\cdot\nabla f(\mathbf{P}_0) \\ &= [(x-1)\mathbf{i}+(y-2)\mathbf{j}+(z-3)\mathbf{k}]\cdot[6\mathbf{i}+4\mathbf{j}-6\mathbf{k}] \\ &= 6(x-1)+4(y-2)-6(z-3) \\ &= 6x-6+4y-8-6z+18 \\ &= 6x+4y-6z+4 \end{aligned}$$

Solving for z, we obtain

$$z=x+(2/3)y+2/3$$

as the equation of the tangent plane. ■

We can find an explicit equation of the form $z=f(x,y)$ for the tangent plane as follows. We write the surface in the form

$$0=F(x,y,z)=f(x,y)-z.$$

It is only when the surface is in the form $F(x,y,z)=0$ that the gradient vector is normal to the surface. But then

$$\begin{aligned} \nabla F(\mathbf{P}_0) &= \mathbf{i}F_x(\mathbf{P}_0)+\mathbf{j}F_y(\mathbf{P}_0)+\mathbf{k}F_z(\mathbf{P}_0) \\ &= \mathbf{i}f_x(\mathbf{P}_0)+\mathbf{j}f_y(\mathbf{P}_0)-1 \end{aligned}$$

Therefore according to equation 13.1

$$\begin{aligned} \mathbf{P}\cdot\nabla F(\mathbf{P}_0) &= \mathbf{P_0}\cdot\nabla F(\mathbf{P}_0) \\ \mathbf{P}\cdot(\mathbf{i}f_x(\mathbf{P}_0)+\mathbf{j}f_y(\mathbf{P}_0)-1) &= \mathbf{P}_0\cdot(\mathbf{i}f_x(\mathbf{P}_0)+\mathbf{j}f_y(\mathbf{P}_0)-1) \end{aligned}$$

Writing $\mathbf{P}=x\mathbf{i}+y\mathbf{j}+z\mathbf{k}$ and $\mathbf{P}_0=x_0\mathbf{i}+y_0\mathbf{j}+z_0\mathbf{k}$,

$$\mathbf{P}\cdot(\mathbf{i}f_x(\mathbf{P}_0)+\mathbf{j}f_y(\mathbf{P}_0)-1)=\mathbf{P}_0\cdot(\mathbf{i}f_x(\mathbf{P}_0)+\mathbf{j}f_y(\mathbf{P}_0)-1)$$

$$(x\mathbf{i}+y\mathbf{j}+z\mathbf{k})\cdot(\mathbf{i}f_x(\mathbf{P}_0)+\mathbf{j}f_y(\mathbf{P}_0)-1)=(x_0\mathbf{i}+y_0\mathbf{j}+z_0\mathbf{k})\cdot(\mathbf{i}f_x(\mathbf{P}_0)+\mathbf{j}f_y(\mathbf{P}_0)-1)$$

$$xf_x(\mathbf{P}_0)+yf_y(\mathbf{P}_0)-z=x_0f_x(\mathbf{P}_0)+y_0f_y(\mathbf{P}_0)-z_0$$

Solving for $z$, we find that the **equation for the tangent plane to the surface** $z=f(x,y)$ **at the point** $\mathbf{P}_0$ is

$$\boxed{z=z_0+(x-x_0)f_x(\mathbf{P}_0)+(y-y_0)f_y(\mathbf{P}_0)} \tag{13.2}$$

**Example 13.2** *Find the equation of the tangent plane to the surface* $z=xe^{-2y}$ *at the point* $(1,0,1)$.

*Solution.* From equation 13.2, we calculate that

$$\begin{aligned}
z - z_0 &= (x - x_0) f_x(x_0, y_0, z_0) + (y - y_0) f_y(x_0, y_0, z_0) \\
z - 1 &= (x - 1)\left(e^{-2y}\big|_{(1,0,1)}\right) + y\left(-2xe^{-2y}\big|_{(1,0,1)}\right) \\
&= (x - 1)e^0 + y(-2(1)(e^0)) \\
&= x - 1 - 2y
\end{aligned}$$

Solving for $z$ gives $z = x - 2y$ as the equation of the tangent plane. ■

**Example 13.3** *Find the equation of the tangent plane to the surface*

$$z = \sqrt{x} + y^{1/3}$$

*at the point* $(1, 1, 2)$.

*Solution.* Differentiating,

$$\left.\frac{\partial f}{\partial x}\right|_{(1,1,2)} = \left.\frac{1}{2\sqrt{x}}\right|_{(1,1,2)} = \frac{1}{2}$$

$$\left.\frac{\partial f}{\partial y}\right|_{(1,1,2)} = \left.\frac{1}{3y^{2/3}}\right|_{(1,1,2)} = \frac{1}{3}$$

From equation 13.2, the tangent plane is therefore

$$\begin{aligned}
z &= z_0 + (x - x_0) f_x(x_0, y_0, z_0) + (y - y_0) f_y(x_0, y_0, z_0) \\
&= 2 + \frac{x - 1}{2} + \frac{y - 1}{3} \\
&= \left(2 - \frac{1}{2} - \frac{1}{3}\right) + \frac{x}{2} + \frac{y}{3} \\
&= \frac{7}{6} + \frac{x}{2} + \frac{y}{3}
\end{aligned}$$

Multiplying through by 7 gives $6z = 7 + 3x + 2y$. ■

**Example 13.4** *Find a point on the surface*

$$z = 2x^2 + 3y^2 \tag{13.3}$$

*where the tangent plane is parallel to the plane*

$$8x - 3y - z = 0$$

.

*Solution.* In our discussion of the equation of a plane in section 14.2 we found that when a plane is written in the form

$$ax + by + cz = d$$

then its normal vector is $\mathbf{n} = (a, b, c)$. Hence we want to find a point on the surface of $z = 2x^2 + 3y^2$ where the normal vector is parallel to

$$\mathbf{n} = (8, -3, -1)$$

Rewriting equation 13.3 in the form $F(x, y, z) = 0$,

$$F(x, y, z) = 2x^2 + 3y^2 - z = 0$$

A general form of the normal vector is given by the gradient vector

$$\nabla F(x, y, z) = (4x, 6y, -1)$$

Since the two gradient vectors are parallel

$$(4x, 6y, -1) = k(8, -3, -1)$$

for some number $k$. The solution is $k = 1$ (which we find from the $z$-component) and hence $x = 2$ and $y = -1/2$. Thus the point on the surface where the normal vector has the right direction is

$$(x, y, z) = (2, -1/2, f(2, -1/2)) = (2, -1/2, 8.75) \quad \blacksquare$$

## Multivariate Taylor Series

Recall the following result from Calculus I (see section 10.8 of the text):

**Theorem 13.1 Taylor's Theorem with Remainder** *Let $f$ be a function whose first $n+1$ derivatives exist for all $x$ in an interval $J$ containing $a$. Then for any $x \in J$,*

$$f(x) = f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2!}f''(a) + \cdots + \frac{(x-a)^n}{n!}f^{(n)}(a) + R_n(x) \quad (13.4)$$

*where the* **remainder** *is*

$$R_n(x) = \frac{(x-a)^{(n+1)}}{(n+1)!}f^{(n+1)}(c) \quad (13.5)$$

*for some (unknown) number $c \in J$.*

**Definition 13.2** *The* **Taylor Polynomial of order n at a (or about a)** *is defined as*

$$P_n(x) = f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2!}f''(a) + \cdots + \frac{(x-a)^n}{n!}f^{(n)}(a) \quad (13.6)$$

**Corollary 13.1** *$f(x) = P_n(x) + R_n(x)$ for all $x \in J$.*

**Theorem 13.2** *If $f$ is infinitely differentiable then $R_n(x) \to 0$ as $n \to \infty$, so that $f(x) = \lim_{n\to\infty} P_n(x)$ for all $x \in J$, or more explicitly,*

$$\begin{aligned} f(x) &= \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k & (13.7)\\ &= f(a) + (x-a)f'(a) + \frac{1}{2!}(x-a)^2 f''(a) + \cdots & (13.8) \end{aligned}$$

*which is called the* **Taylor Series of f about the point x=a**

**Example 13.5** *Find the Taylor Polynomial of order 3 of $f(x) = \sqrt{x+1}$ and the corresponding remainder formula about the point $a = 0$.*

*Solution.* Taking the first 4 derivatives,

$$\begin{aligned} f(x) &= (x+1)^{1/2} &\Rightarrow\quad& f(0) = 1\\ f'(x) &= \frac{1}{2}(x+1)^{-1/2} &\Rightarrow\quad& f'(0) = \frac{1}{2}\\ f''(x) &= -\frac{1}{4}(x+1)^{-3/2} &\Rightarrow\quad& f''(0) = -\frac{1}{4}\\ f'''(x) &= \frac{3}{8}(x+1)^{-5/2} &\Rightarrow\quad& f'''(0) = \frac{3}{8} \end{aligned}$$

$$f^{(4)}(x) = -\frac{15}{16}(x+1)^{-7/2} \quad \Rightarrow \quad f^{(4)}(0) = -\frac{15}{16}$$

Therefore

$$\begin{aligned}
P_3(x) &= f(0) + xf'(0) + \frac{x^2}{2}f''(0) + \frac{x^3}{3!}f^{(3)}(0) \\
&= 1 + \frac{1}{2}x + \frac{1}{2}\left(-\frac{1}{4}\right)x^2 + \frac{1}{6}\left(\frac{3}{8}\right)x^3 \\
&= 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3
\end{aligned}$$

and

$$R_3(x) = \frac{f^{(4)}(c)}{4!}x^4 = \frac{-15(c+1)^{-7/2}x^4}{384} = -\frac{15x^4}{384(c+1)^{7/2}}$$

for some $c$ between 0 and $x$. Since smaller denominators make larger numbers, the remainder is maximized for the smallest possible value of $c$, which occurs when $c = 0$. So the error at $x$ is bounded by

$$|R_3(x)| \leq \frac{15x^4}{384}$$

For example, the error at $x = 1$ is no more than 15/384. ■

Looking more closely at the Taylor series

$$f(x) = f(a) + (x-a)f'(a) + \frac{1}{2!}(x-a)^2 f''(x_0) + \cdots$$

we observe that the first two terms

$$f(x) = f(a) + (x-a)f'(a) + \cdots$$

give the equation of a line tangent to $f(x)$ at the point $(a, f(a))$. The next term gives a quadratic correction, followed by a cubic correction, and so forth, so we might write

$$\begin{aligned}
f(x \text{ near } a) &= (\text{equation of a tangent line through } a) \\
&\quad + (\text{quadratic correction at } a) \\
&\quad + (\text{cubic correction at } a) \\
&\quad + (\text{quartic correction at } a) + \cdots
\end{aligned}$$

For a function of two variables $z = f(x, y)$, the *tangent line* because a *tangent plane*; the *quadratic correction* becomes a *parabolid correction*; and so forth. The explicit result is the following.

**Theorem 13.3** *Let $f(x, y)$ be infinitely differentiable in some open set $J$ that contains the point $(a, b)$. Then the* **Taylor Series of f(x,y) about the point (a,b)** *is*

$$\begin{aligned}
f(x,y) &= f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b) \\
&\quad + \frac{1}{2}\left[f_{xx}(a,b)(x-a)^2 + 2f_{xy}(a,b)(x-a)(y-b) + f_{yy}(a,b)(y-b)^2\right] \\
&\quad + \cdots
\end{aligned}$$

*for all points $(x, y) \in J$.*

According to the text,

> "The details are best left to higher-level books." [8th ed., page 669].

Nevertheless we will at least do one example.

**Example 13.6** *Find a Taylor Approximation to* $f(x, y) = \sqrt{x + y + 1}$ *near the origin*

*Solution.* Our point $(a, b)$ is the origin, so that $(a, b) = (0, 0)$. Differentiating,

$$f_x(x, y) = f_y(x, y) = \frac{1}{2\sqrt{x + y + 1}}$$

Hence

$$f_x(a, b) = f_x(0, 0) = \frac{1}{2}, \quad f_y(a, b) = f_y(0, 0) = \frac{1}{2}$$

Similarly

$$f_{xx}(x, y) = f_{yy}(x, y) = f_{xy}(x, y) = \frac{-1}{4(x + y + 1)^{3/2}}$$

so that

$$f_{xx}(0, 0) = f_{yy}(0, 0) = f_{zz}(0, 0) = -\frac{1}{4}$$

Therefore the Taylor expansion is

$$\begin{aligned} f(x, y) &\approx f(0, 0) + [x f_x(0, 0) + y f_y(0, 0)] \\ &\quad + \frac{1}{2}\left[f_{xx}(0, 0)x^2 + 2f_{xy}(0, 0)xy + f_{yy}(0, 0)y^2\right] + \cdots \\ &= 1 + \frac{1}{2}x + \frac{1}{2}y + \frac{1}{2}\left[-\frac{1}{4}x^2 - \frac{2}{4}xy - \frac{1}{4}y^2\right] + \cdots \\ &= 1 + \frac{1}{2}(x + y) - \frac{1}{8}(x^2 + 2xy + y^2) + \cdots \\ &= 1 + \frac{1}{2}(x + y) - \frac{1}{8}(x + y)^2 + \cdots \quad \blacksquare \end{aligned}$$

# Lecture 14

# Unconstrained Optimization

Every continuous function of two variables $f(x, y)$ that is defined on a closed, bounded set attains both a minimum and a maximum. The process of finding these points is called **unconstrained optimization**. The process of finding a maximum or a minimum subject to a constraint will be discussed in the following section.

**Definition 14.1** *Let $f(x, y)$ be a function defined on some set $S$, and let $(a, b) \in S$ be a point in $S$. Then we say that*

1. *$f(a, b)$ is a* **global maximum of $f$ on $S$** *if $f(a, b) \geq f(x, y)$ for all $(x, y) \in S$.*
2. *$f(a, b)$ is a* **global minimum of $f$ on $S$** *if $f(a, b) \leq f(x, y)$ for all $(x, y) \in S$.*
3. *$f(a, b)$ is a* **global extremum** *if it is either a global maximum or a global minimum on $S$*

**Theorem 14.1** *Let $f(x, y)$ be a continuous function on some closed, bounded set $S \subset \mathbb{R}^2$. Then $f(x, y)$ has both a global maximum value and a global minimum value on $S$.*

The procedure for finding the extrema (maxima and minima) of functions of two variables is similar to the procedure for functions of a single variable:

1. Find the critical points of the function to determine candidate locations for the extrema (in the single-dimensional case, these were points where the derivative is zero or undefined, and the boundary points of the interval);
2. Examine the second derivative at the candidate points that do not lie on the border (in one-dimension, we had $f''(a) < 0$ at local maxima and $f''(a) > 0$ at local minima).
3. Compare internal extrema with the value of the function on the boundary points and at any points where the derivative or second derivative is undefined to determine absolute extrema.

**Definition 14.2** *Let $f(x, y)$ be defined on some set $J \subset \mathbb{R}^2$. Then the* **candidate extrema points** *occur at*

1. *All points on the boundary of* $J$*;*

2. *All* **stationary points***, namely, any point where* $\nabla f(x, y) = 0$*;*

3. *All* **singular points***, namely, any points where* $f(x, y)$ *or* $\nabla f(x, y)$ *are undefined.*

*The candidate extrema are sometimes called* **critical points***; some authors restrict the use of this term to describe the stationary points.*

Stationary points occur when $\nabla f(x, y) = 0$; this requires all of the derivatives to be zero:

$$\frac{\partial f}{\partial x} = 0 = \frac{\partial f}{\partial y}$$

Recall from one-variable calculus that three different things could happen at a stationary point where $f'(a) = 0$:

- $f(x)$ could have a local maximum at $x = a$, as, for example, occurs for $f(x) = -(x - a)^2$
- $f(x)$ could have a local minimum at $x = a$, as, for example, occurs for $f(x) = (x - a)^2$
- $f(x)$ could have an inflection point at $x = a$, as, for example, occurs for $f(x) = (x - a)^3$

For multivariate functions there are three types of stationary points: local maxima, local minima, and saddle points. Saddle points are the generalization of inflection points. At a saddle point, in some vertical cross-sections, the function appears to have a local minimum, while in other vertical cross-sections, the function appears to have a local maximum.

**Example 14.1** *Find the stationary points of the function* $f(x, y) = x^2 - 7xy + 12y^2 - y$

*Solution.* The function has stationary points when $f_x(x, y) = f_y(x, y) = 0$. Differentiating,

$$\begin{aligned}
\frac{\partial f}{\partial x} = 0 \quad &\Rightarrow \quad \frac{\partial}{\partial x}(x^2 - 7xy + 12y^2 - y) = 2x - 7y = 0 \\
&\Rightarrow \quad y = 2x/7 \\
\frac{\partial f}{\partial y} = 0 \quad &\Rightarrow \quad \frac{\partial}{\partial y}(x^2 - 7xy + 12y^2 - y) = -7x + 24y - 1 = 0 \\
&\Rightarrow \quad 24y = 7x + 1
\end{aligned}$$

Combining the two results,

$$\frac{48}{7}x = 7x + 1 = \frac{49x + 7}{7} \Rightarrow x = -7$$

Figure 14.1: Top: Topology of a solid around a saddle point; the canonical function is $z = y^2 - x^2$ with a saddle point at the origin, although many other functions display saddle points. Bottom: Origin of the nomenclature. The draft saddle model 1583 is sold by Midwest Leather Co. for $955 [from http://www.draftsource.com/Midwest_Leather].

Hence

$$y = 2x/7 = -2$$

So the only stationary point is at $(x, y) = (-7, -2)$. ■

**Example 14.2** *Find the stationary points of $f(x, y) = x^3 + y^3 - 6xy$*

*Solution.* Differentiating, as before,

$$0 = \frac{\partial f}{\partial x} = \frac{\partial}{\partial x}(x^3 + y^3 - 6xy) = 3x^2 - 6y \tag{14.1}$$

$$0 = \frac{\partial f}{\partial y} = \frac{\partial}{\partial y}(x^3 + y^3 - 6xy) = 3y^2 - 6x \tag{14.2}$$

From equation 14.1,

$$y = 3x^2/6 = x^2/2$$

From equation 14.2,

$$x = 3y^2/6 = y^2/2 = (x^2/2)^2/2 = x^4/8$$

$$x^4 - 8x = x(x^3 - 8) = 0$$

$$x = 0 \text{ or } x = 2$$

When $x = 0$, $y = 0$. When $x = 2$, $y = 2^2/2 = 2$. So the stationary points are $(0, 0)$ and $(2, 2)$. ■

## The Second Derivative Test: Classifying the Stationary Points

**Theorem 14.2 Second Derivative Test.** *Suppose that $f(x, y)$ has continuous partial derivatives in some neighborhood of $(a, b)$ where*

$$f_x(a, b) = f_y(a, b) = 0$$

*Define the function*

$$D(x, y) = f_{xx}f_{yy} - f_{xy}^2$$

*and the number*

$$d = D(a, b) = f_{xx}(a, b)f_{yy}(a, b) - (f_{xy}(a, b))^2$$

*Then*

1. *If $d > 0$ and $f_{xx}(a, b) < 0$, then $f(a, b)$ is a local maximum;*
2. *If $d > 0$ and $f_{xx}(a, b) > 0$, then $f(a, b)$ is a local minimum;*
3. *If $d < 0$ then $f(a, b)$ is a saddle point;*
4. *If $d = 0$ the second-derivative test is inconclusive.*

*Proof of the Second Derivative Test.* Given at the end of this section. ■

**Example 14.3** *Classify the stationary points of $f(x, y) = x^3 + y^3 - 6xy$*

*Solution.* We found in the previous example that stationary points occur at (0, 0) and (2, 2), and that

$$f_x = 3x^2 - 6y$$
$$f_y = 3y^2 - 6x$$

Hence the second derivatives are

$$f_{xx} = 6x$$
$$f_{xy} = -6$$
$$f_{yy} = 6y$$

Thus

$$D(x, y) = f_{xx}f_{yy} - f_{xy}^2 = (6x)(6y) - (-6)^2 = 36xy - 36$$

At $(a, b) = (0, 0)$,

$$d = -36 < 0$$

hence $(0, 0)$ is a saddle point.

At $(a, b) = (2, 2)$,

$$d = (36)(2)(2) - 36 > 0$$

and

$$f_{xx}2, 2 = 12 > 0$$

Hence $(2, 2)$ is a local minimum. ■

**Example 14.4** . *Find and classify the stationary points of $f(x, y) = xy$*

*Solution.* Differentiating,

$$f_x = y$$
$$f_y = x$$

The only stationary point occurs when $x = y = 0$, i.e., at the origin. Furthermore,

$$f_{xx} = f_{yy} = 0$$
$$f_{xy} = 1$$

Therefore

$$D(x, y) = f_{xx}f_{yy} - f_{xy}^2 = (0)(0) - (1)^2 = -1 < 0$$

for all $x, y$. We conclude that $(0, 0)$ is a saddle point. ■

**Example 14.5** *Find and classify all the stationary points of $f(x, y) = e^{-(x^2+y^2-4y)}$*

*Solution.* Differentiating,

$$f_x = -2xe^{-(x^2+y^2-4y)}$$
$$f_y = (-2y + 4)e^{-(x^2+y^2-4y)}$$

Setting $f_x = 0$ gives

$$-2xe^{-(x^2+y^2-4y)} = 0$$
$$\Rightarrow x = 0$$

Setting $f_y = 0$ gives

$$(-2y + 4)e^{-(x^2+y^2-4y)} = 0$$
$$\Rightarrow -2y + 4 = 0$$
$$\Rightarrow y = 2$$

The only stationary point is at (0, 2). Next, we calculate the second derivatives.

$$\begin{aligned} f_{xx} &= -2e^{-(x^2+y^2-4y)} + 4x^2e^{-(x^2+y^2-4y)} \\ &= (-2 + 4x^2)e^{-(x^2+y^2-4y)} \\ f_{xy} &= -2x(-2y + 4)e^{-(x^2+y^2-4y)} \\ f_{yy} &= -2e^{-(x^2+y^2-4y)} + (-2y + 4)^2e^{-(x^2+y^2-4y)} \\ &= (-2 + 4y^2 - 16y + 16)e^{-(x^2+y^2-4y)} \\ &= (4y^2 - 16y + 14)e^{-(x^2+y^2-4y)} \end{aligned}$$

At the critical point (0,2), we find that

$$\begin{aligned} f_{xx}(0, 2) &= (-2 + 4(0)^2)e^{-(0^2+2^2-4(2))} = -2e^4 \\ f_{xy}(0, 2) &= -2(0)(-2(2) + 4)e^{-(0^2+2^2-4(2))} = 0 \\ f_{yy}(0, 2) &= (4(2)^2 - 16(2) + 14)e^{-(0^2+2^2-4(2))} = -2e^4 \end{aligned}$$

Therefore

$$d = D(0,2) = f_{xx}f_{yy} - f_{xy}^2 = (-2e^4)(-2e^4) - (0)^2 = 4e^4 > 0$$

Since $f_{xx} < 0$ the point $(0,2)$ a local maximum. ■

**Example 14.6** *Find and classify the stationary points of*

$$f(x,y) = x^2 + 1 - 2x\cos y, -\pi < y < \pi$$

*Solution.* Proceeding as before

$$f_x = 2x - 2\cos y$$

$$f_y = 2x\sin y$$

Setting these expressions equal to zero gives

$$x = \cos y \tag{14.3}$$

$$x\sin y = 0 \Rightarrow x = 0 \text{ or } \sin y = 0 \tag{14.4}$$

Equation 14.4 can be rewritten (using the fact that $-pi < y < pi$) as

$$x = 0 \text{ or } y = 0 \tag{14.5}$$

because $\sin y = 0$ implies that $y = \pm k\pi, k = 0, 1, 2, ....$ Therefore we have two cases consider. The first case is

$$x = \cos y \text{ and } x = 0 \tag{14.6}$$

The second case is

$$x = \cos y \text{ and } y = 0 \tag{14.7}$$

From equation 14.6, if $x = 0$ then $\cos y = 0$, which means $y = \pm\pi/2$. So our first two stationary points are

$$(0, \pi/2), \quad (0, -\pi/2)$$

The second case (equation 14.7 ) has $y = 0$ and $x = \cos y = \cos 0 = 1$. So there is a third stationary point at

$$(0, 1)$$

We now classify the critical points. To do so we must first calculate the second derivatives,

$$f_{xx} = 2$$

$$f_{xy} = 2\sin y$$

$$f_{yy} = 2x\cos y$$

Hence

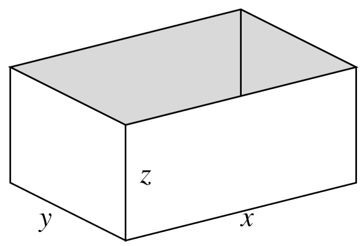
$$D(x,y) = f_{xx}f_{yy} - f_{xy}^2 = 4x\cos y - 4\sin^2 y$$

At (1,0), we have $d = 4(1)\cos(0) - 4\sin^2(0) = 4 > 0$. Since $f_{xx} = 2 > 0$ this is a local minimum.

At $(0, -\pi/2)$ we have $d = D(0, -\pi/2) = (4)(0)\cos(-\pi/2) - 4\sin^2(-\pi/2) = -4(-1)^2 = -4 < 0$ so this is a saddle point.

At $(0, \pi/2)$ $d = D(0, \pi/2) = (4)(0)\cos(\pi/2) - 4\sin^2(\pi/2) = -4(1)^2 = -4 < 0$ so this is also a saddle point. ■

Figure 14.2: Geometry for example 14.7.




**Example 14.7** *A rectangular metal tank with an open top is to hold 256 cubic feet of liquid. What are the dimensions of the tank that requires the least material to build?*

*Solution.* Assume that the material required to build the tank is proportional to the area of the cube. Let the box have dimensions $x$, $y$, and $z$ as illustrated in figure 14.2.

The area of the bottom is $xy$; the area of each of the two small sides in the figure is $yz$; and the area of each of the larger sides (in the figures) is $xz$. Therefore the total area (since there is one bottom but two of each of the additional sides) is

$$A = xy + 2xz + 2yz$$

Furthermore, since the volume is 256, and since the volume of the box is $xyz$, we have

$$256 = xyz$$

or

$$z = 256/xy$$

Combining this with our previous equation for the area,

$$\begin{aligned} A &= xy + 2xz + 2yz \\ &= xy + 2x(256/xy) + 2y(256/xy) \\ &= xy + \frac{512}{y} + \frac{512}{x} \end{aligned}$$

This is the function we need to minimize. We define

$$f(x, y) = xy + \frac{512}{y} + \frac{512}{x}$$

Taking derivatives,

$$f_x = y - 512/x^2$$

$$f_y = x - 512/y^2$$

Stationary points occur when

$$0 = y - 512/x^2 \text{ and } 0 = x - 512/y^2$$

Thus

$$y = 512/x^2 = 512/(512/y^2)^2 = y^4/512$$

$$\Rightarrow 512y - y^4 = 0$$

$$\Rightarrow y(512 - y^3) = 0$$

$$\Rightarrow y = 0 \text{ or } y = 8$$

Since we also required $x = 512/y^2$, we find that at $y = 0$, $x \to \infty$ hence there is no stationary point in this case.

At $y = 8$ we have $x = 512/8^2 = 8$. Hence the only stationary point is $(8, 8)$. We now proceed to test this stationary point. The second derivatives are

$$f_{xx} = 1024/x^3$$

$$f_{xy} = 1$$

$$f_{yy} = 1024/y^3$$

Thus

$$f_{xx}(8, 8) = 1024/(8^3) = 1024/512 = 2$$

$$f_{xy}(8, 8) = 1$$

$$f_{yy}(8, 8) = 1024/(8^3) = 1024/512 = 2$$

and

$$d = D(8, 8) = f_{xx}(8, 8)f_{yy}(8, 8) - f_{xy}^2(8, 8) = (2)(2) - (1)^2 = 3 > 0$$

Since $d > 0$ and $f_{xx}(8, 8) > 0$, we conclude that $f(x, y)$ has a local minimum at (8,8). The third dimension is

$$z = 256/xy = 256/64 = 4$$

The dimensions of the minimum area box are 8 by 8 by 4, and its total area is $A = xy + 2xz + 2yz = (8)(8) + 2(8)(4) + 2(8)(4) = 192$ ■

## Least Squares Linear Regression

Suppose we have a large set of data points in the $xy$-plane

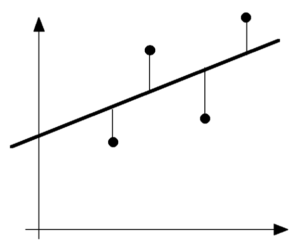
$$\{(x_i, y_i) : i = 1, 2, ..., n\}$$

and we want to find the "best fit" straight line to our data, namely, we want to find number $m$ and $b$ such that

$$y = mx + b$$

is the "best" possible line in the sense that it minimizes the total sum-squared vertical distance between the data points and the line.

Figure 14.3: The least squares procedure finds the line that minimizes the total sum of all the vertical distances as shown between the line and the data points.



The vertical distance between any point $(x_i, y_i)$ and the line, which we will denote by $d_i$, is

$$d_i = |mx_i + b - y_i|$$

Since this distance is also minimized when its square is minimized, we instead calculate

$$d_i^2 = (mx_i + b - y_i)^2$$

The total of all these square-distances (the "sum-squared-distance") is

$$f(m, b) = \sum_{i=1}^{n} d_i^2 = \sum_{i=1}^{n} (mx_i + b - y_i)^2$$

The only unknowns in this expression are the slope $m$ and $y$-intercept $b$. Thus we have written the expression as a function $f(m, b)$. Our goal is to find the values of $m$ and $b$ that correspond to the global minimum of $f(m, b)$.

Setting $\partial f/\partial b = 0$ gives

$$\begin{aligned} 0 &= \frac{\partial f}{\partial b} = \frac{\partial}{\partial b}\sum_{i=1}^{n}(mx_i + b - y_i)^2 \\ &= \sum_{i=1}^{n} 2\,(mx_i + b - y_i) \\ &= 2\sum_{i=1}^{n}(mx_i + b - y_i) \end{aligned}$$

Dividing by 2 and separating the three sums

$$\begin{aligned} 0 &= \sum_{i=1}^{n}(mx_i + b - y_i) \\ &= \sum_{i=1}^{n} mx_i + \sum_{i=1}^{n} b - \sum_{i=1}^{n} y_i \\ &= m\sum_{i=1}^{n} x_i + nb - \sum_{i=1}^{n} y_i \end{aligned}$$

Defining

$$X = \sum_{i=1}^{n} x_i \tag{14.8}$$

$$Y = \sum_{i=1}^{n} y_i \tag{14.9}$$

then we have

$$0 = mX + nb - Y \tag{14.10}$$

Next, we set $\partial f/\partial m = 0$, which gives

$$\begin{aligned} 0 &= \frac{\partial f}{\partial m} = \frac{\partial}{\partial m}\sum_{i=1}^{n}(mx_i + b - y_i)^2 \\ &= \sum_{i=1}^{n} 2x_i\,(mx_i + b - y_i) \\ &= 2\sum_{i=1}^{n} x_i\,(mx_i + b - y_i) \end{aligned}$$

Dividing by 2 and separating the three sums as before

$$\begin{aligned} 0 &= \sum_{i=1}^{n} x_i \left(mx_i + b - y_i\right) \\ &= \sum_{i=1}^{n} mx_i^2 + b\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i y_i \\ &= m\sum_{i=1}^{n} x_i^2 + bX - \sum_{i=1}^{n} x_i y_i \end{aligned}$$

where $X$ is defined in equation 14.8. Next we define,

$$A = \sum_{i=1}^{n} x_i^2 \qquad (14.11)$$

$$C = \sum_{i=1}^{n} x_i y_i \qquad (14.12)$$

so that

$$0 = mA + bX - C \qquad (14.13)$$

Equations **??** and **??** give us a a system of two linear equations in two variables $m$ and $b$. Multiplying equation 14.10 by $A$ and equation 14.13 by $X$ gives

$$0 = A\left(mX + nb - Y\right) = AXm + Anb - AY \qquad (14.14)$$

$$0 = X\left(mA + bX - C\right) = AXm + X^2 b - CX \qquad (14.15)$$

Subtracting these two equations gives

$$0 = Anb - AY - X^2 b + CX = b(An - X^2) + CX - AY$$

and therefore

$$b = \frac{AY - CX}{An - X^2} = \frac{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i y_i \sum_{i=1}^{n} x_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

If we instead multiply equation 14.10 by $X$ and equation 14.13 by $n$ we obtain

$$0 = X\left(mX + nb - Y\right) = mX^2 + nXb - YX$$

$$0 = n\left(mA + bX - C\right) = nAm + nXb - nC$$

Subtracting these two equations,

$$0 = m\left(X^2 - nA\right) - \left(YX - nC\right)$$

Solving for $m$ and substituting the definitions of $A$, $C$, $X$ and $Y$, gives

$$m = \frac{XY - nC}{X^2 - nA} = \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i - n \sum_{i=1}^{n} x_i y_i}{\left(\sum_{i=1}^{n} x_i\right)^2 - n \sum_{i=1}^{n} x_i^2}$$

Generally this algorithm needs to implemented computationally, because there are so many sums to calculate. It is also implemented on many calculators.

## Least Squares Algorithm

*To find a best-fit line to a set of n data points*

$$(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$$

*calculate,*

$$X = \sum_{i=1}^{n} x_i$$
$$Y = \sum_{i=1}^{n} y_i$$
$$A = \sum_{i=1}^{n} x_i^2$$
$$C = \sum_{i=1}^{n} x_i y_i$$

*The best fit line is*

$$y = mx + b$$

*where*

$$m = \frac{XY - nC}{X^2 - nA}$$
$$b = \frac{AY - CX}{An - X^2}$$

**Example 14.8** *Find the least squares fit to the data (3, 2), (4,3), (5,4), (6,4) and (7,5).*

*Solution.* First we calculate the numbers $X$, $Y$, $A$, and $C$,

$$\begin{aligned} X &= \sum_{i=1}^{n} x_i = 3 + 4 + 5 + 6 + 7 = 25 \\ Y &= \sum_{i=1}^{n} y_i = 2 + 3 + 4 + 4 + 5 = 18 \\ A &= \sum_{i=1}^{n} x_i^2 = 9 + 16 + 25 + 36 + 49 = 135 \\ C &= \sum_{i=1}^{n} x_i y_i = (3)(2) + (4)(3) + (5)(4) + (6)(4) + (7)(5) = 97 \end{aligned}$$

Therefore

$$m = \frac{XY - nC}{X^2 - nA} = \frac{(25)(18) - (5)(97)}{(25)^2 - 5(135)} = \frac{450 - 485}{625 - 675} = \frac{-35}{-50} = 0.7$$

and

$$b = \frac{AY - CX}{An - X^2} = \frac{(135)(18) - (97)(25)}{(135)(5) - 25^2} = \frac{2430 - 2425}{50} = \frac{5}{50} = 0.1$$

So the best fit line is $y = 0.7x + 0.1$ ■

## Derivation of the Second Derivative Test

Let $z = f(x, y)$ and suppose that $P = (a, b)$ is a critical point where all the partials are zero, namely,

$$f_x(a, b) = 0$$

$$f_y(a, b) = 0$$

An equivalent way of stating this is that

$$\nabla f(a, b) = \mathbf{i} f_x(a, b) + \mathbf{j} f_y(a, b) = 0$$

We ask the following question: what conditions are ***sufficient*** to ensure that $P$ is the location of either a maximum or a minimum?

We will study this question by expanding the function in a Taylor Series approximation about a critical point where the partials are zero.

$$\begin{aligned} f(x) &= f(P) + f_x(P)\,(x-a) + f_y(P)\,(y-b) \\ &\quad + \frac{1}{2} f_{xx}(P)\,(x-a)^2 + f_{xy}(P)\,(x-a)\,(y-b) + \frac{1}{2} f_{yy}(P)\,(y-b)^2 \end{aligned}$$

If the first partial derivatives are zero, then

$$\begin{aligned} f(x) &= f(P) + \frac{1}{2} f_{xx}(P)\,(x-a)^2 \\ &\quad + f_{xy}(P)\,(x-a)\,(y-b) + \frac{1}{2} f_{yy}(P)\,(y-b)^2 \end{aligned}$$

Next we make a change of variables

$$u = x - a$$

$$v = y - b$$

$$g(u, v) = f(x, y) - f(a, b)$$

Then

$$\frac{\partial g}{\partial u} = \frac{\partial}{\partial u}\left[f(x, y) - f(P)\right] = \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u} = \frac{\partial f}{\partial x}(1) + \frac{\partial f}{\partial y}(0) = \frac{\partial f}{\partial x}$$

and so forth gives us the relationships

$$f_{xx}(P) = g_{xx}(0) = g_{xx,0}$$

$$f_{xy}(P) = g_{xy}(0) = g_{xy,0}$$

$$f_{yy}(P) = g_{yy}(0) = g_{yy,0}$$

We have transformed the critical point of $f(x, y)$ at $(a, b)$ to a critical point of $g(u, v)$ at the origin, where

$$g(u, v) = \frac{1}{2} g_{uu,0} u^2 + g_{uv,0} uv + \frac{1}{2} g_{vv,0} v^2$$

Now we make the following substitutions,

$$A = \frac{1}{2}g_{uu,0} = \frac{1}{2}f_{xx,P}$$

$$B = g_{uv,0} = f_{xy,P}$$

$$C = \frac{1}{2}g_{vv,0} = \frac{1}{2}f_{yy,P}$$

which gives us

$$g(u,v) = Au^2 + Buv + Cv^2$$

In other words, we have approximated the original function with a quadratic at the origin (note to physics majors: this is a generalization of a property called Hooke's law to 2 dimensions).

The function

$$D(u,v) = g_{uu}g_{vv} - (g_{uv})^2$$

so that

$$d = (2A)(2C) - B^2 = 4AC - B^2$$

Completing the squares in the formula for $g$,

$$\begin{aligned}
g(u,v) &= A\left[\left(u^2 + \frac{B}{A}uv + \frac{B^2v^2}{4A^2}\right) - \frac{B^2v^2}{4A^2} + \frac{C}{A}v^2\right] \\
&= A\left[\left(u + \frac{B}{2A}v\right)^2 + \left(\frac{C}{A} - \frac{B^2}{4A^2}\right)v^2\right] \\
&= A\left[\left(u + \frac{B}{2A}v\right)^2 + \left(\frac{4AC - B^2}{4A^2}\right)v^2\right] \\
&= A\left[\left(u + \frac{B}{2A}v\right)^2 + \frac{d}{4A^2}v^2\right]
\end{aligned}$$

We have three cases to consider, depending on the value of $d$.

**Case 1: d > 0**

If $d > 0$ then everything in the brackets is positive except at the origin where it is zero. Thus the origin is a local minimum of everything inside the square brackets. The function $g$ is a paraboloid that extends upwards around the origin when $A > 0$ and downwards when $A < 0$. Thus if $A > 0$, we have a local minimum and if $A < 0$ we have a local maximum.

**Case 2: d < 0**

If $d < 0$ then write $d = -|d| = m^2 > 0$ for some number $m > 0$, so that

$$g(u, v) = A\left[\left(u + \frac{B}{2A}v\right)^2 - \left(\frac{mv}{2A}\right)^2\right]$$

If we make yet another change of variables,

$$p = u + \frac{Bv}{2A}, \ \ q = \frac{mv}{2A}$$

then

$$g(u, v) = A\left[p^2 - q^2\right]$$

This is a hyperbolic paraboloid with a saddle at $(p, q) = (0, 0)$. The original critical point is at $u = 0, v = 0$, which corresponds to $(p, q) = 0$. So this is also a saddle point in $uv$-space.

**Case 3: d=0**

Finally, if $d = 0$ then

$$g(u, v) = A\left(u + \frac{Bv}{2A}\right)^2$$

Along the line $u = -B/(2A)v$, $g(u, v)$ identically equal to zero. Thus $g(u, v)$ is a constant along this line. Otherwise, $g$ is always positive when $A > 0$ and always negative (when $A < 0$). Thus the origin is neither a maximum of a minimum. ∎

# Lecture 15

# Constrained Optimization: Lagrange Multipliers

In the previous section we learned how to find the maximum and minimum points of a function of two variables. In this section we will study how to find the maximum and minimum points on a function subject to a constraint using a technique called **Lagrange's Method** or **The Method of Langrange Multipliers**.

**Theorem 15.1 Lagrange's Method.** *An extreme value of the function $f(x, y)$ subject to the* **constraint**

$$g(x, y) = 0$$

*occurs when*

$$\nabla f(x, y) = \lambda \nabla g(x, y)$$

*for some number $\lambda$.*

We introduce the method along with the concept of a constraint via the following example.

**Example 15.1** *Find the dimensions and area of the largest rectangular area that can be enclosed with a fixed length of fence $M$.*

*Solution* Let the sides of the rectangle have length $x$ and $y$. Then the area is given by the function $f(x, y)$ where

$$f(x, y) = xy$$

. Similarly, the perimeter $M$ is

$$M = 2x + 2y$$

One way to solve this problem is to solve for one of the variables, say $y$, as a function of $x$ and $M$, substitute the result into the equation for the area, and then maximize $f$ which becomes a function of a single variable. Instead, we will use a different method: We want to maximize the function $f(x, y)$ subject to the constraint $g(x, y) = 0$ where

$$g(x, y) = M - 2x - 2y$$

According to Lagrange's method, the optimum (maximum or minimum) occurs when

$$\nabla f(x, y) = \lambda \nabla g(x, y)$$

for some number $\lambda$. Thus

$$(y, x, 0) = -\lambda\,(2, 2, 0)$$

This gives us three equations in three unknowns ($x$, $y$, and $\lambda$; the third equation is the original constraint)

$$y = -2\lambda$$

$$x = -2\lambda$$

$$M = 2x + 2y$$

We can find $\lambda$ by substituting the first two equations into the third:

$$M = 2(-2\lambda) + 2(-2\lambda) = -8\lambda \Rightarrow \lambda = -M/8$$

Hence

$$x = y = M/4$$

and the total area is

$$f(M/4, M/4) = M^2/16$$

To test whether this is a maximum or a minimum, we need to pick some other solution that satisfies the constraint, say $x = M/8$. Then $y = M/2 - x = M/2 - M/8 = 3M/8$ and the area is

$$f(M/8, 3M/8) = 3M^2/64 < 4M^2/64 = M^2/16$$

Since this area is smaller, we conclude that the point $(M/4, M/4)$ is the location of the maximum (and not the minimum) area. ■

**Example 15.2** *Verify the previous example using techniques from Calculus I.*

*Solution.* We want to maximize $f(x, y) = xy$ subject to $M = 2x + 2y$. Solving for $y$ gives

$$y = M/2 - x$$

hence

$$f(x, y) = x\left(\frac{M}{2} - x\right) = \frac{Mx}{2} - x^2$$

Setting the derivative equal to zero,

$$0 = \frac{M}{2} - 2x$$

or $x = M/4$. This is the same answer we found using Lagrange's Method. ■

Why does Lagrange's method work? The reason hinges on the fact that the gradient of a function of two variables is perpendicular to the level curves. The equation

$$\nabla f(x, y) = \lambda \nabla g(x, y)$$

says that the normal vector to a level curve of $f$ is parallel to a normal vector of the curve $g(x, y) = 0$. Equivalently, the curve $g(x, y) = 0$ is tangent to a level curve of $f(x, y)$. Why is this an extremum? Suppose that it is not an extremum. Then we can move a little to the left or the right along $g(x, y) = 0$ and we will go to a higher or lower level curve of $f(x, y)$. But this is impossible because we are tangent to a level curve, so if we move infinitesimally in either direction, we will not change the value of $f(x, y)$. Hence the value of $f(x, y)$ must be either a maximum or a minimum at the point of tangency.

**Example 15.3** . *Find the maximum and minimum value of the function*

$$f(x, y) = x + y$$

*on the ellipse*

$$3x^2 + 4y^2 = 25$$

*Solution.* We use Lagrange's method with the constraint

$$g(x, y) = 3x^2 + 4y^2 - 25 = 0$$

setting $\nabla f = \lambda \nabla g$ to give

$$(1, 1) = \lambda (6x, 8y)$$

Our system of equations is

$$x = \frac{1}{6\lambda}$$

$$y = \frac{1}{8\lambda}$$

$$3x^2 + 4y^2 = 25$$

Substituting the first two equations into the third equation gives

$$25 = \frac{3}{36\lambda^2} + \frac{4}{64\lambda^2}$$

$$25\lambda^2 = \frac{1}{12} + \frac{1}{16} = \frac{16 + 12}{(16)(12)} = \frac{28}{192} = \frac{7}{48}$$

$$\lambda^2 = \frac{7}{(48)(25)} = \frac{7}{1200}$$

$$\lambda = \pm\sqrt{\frac{7}{1200}} = \pm\frac{1}{20}\sqrt{\frac{7}{3}}$$

Thus

$$x = \frac{1}{6\lambda} = \pm\frac{20}{6}\sqrt{\frac{3}{7}} = \pm\frac{10}{3}\sqrt{\frac{3}{7}} = \pm\frac{10}{\sqrt{21}}$$

and

$$y = \frac{1}{8\lambda} = \pm\frac{20}{8}\sqrt{\frac{3}{7}} = \pm\frac{5}{2}\sqrt{\frac{3}{7}} = \pm\frac{5}{2}\sqrt{\frac{3\cdot 3}{3\cdot 7}} = \pm\frac{15}{2\sqrt{21}}$$

Hence

$$(x, y) = \pm\left(\frac{10}{\sqrt{21}}, \frac{15}{2\sqrt{21}}\right) = \pm\frac{5}{\sqrt{21}}\left(2, \frac{3}{2}\right)$$

To determine which point gives the minimum and which point gives the maximum we must evaluate $f(x, y)$ at each point.

$$f\left(\frac{10}{\sqrt{21}}, \frac{15}{2\sqrt{21}}\right) = \frac{35}{2\sqrt{21}}$$

whereas

$$f\left(-\frac{10}{\sqrt{21}}, -\frac{15}{2\sqrt{21}}\right) = -\frac{35}{2\sqrt{21}}$$

so the positive solution gives the location of the maximum and the negative solution gives the location of the minimum. ■

**Example 15.4** *Find the minimum of*

$$f(x, y, z) = x^2 + y^2 + z^2$$

*on the plane*

$$x + 3y - 2z = 12$$

*Solution.* We write the constraint as

$$g(x, y) = x + 3y - 2z - 12 = 0$$

Setting $\nabla f = \lambda\nabla g$ gives

$$(2x, 2y2z) = \lambda\,(1, 3, -2)$$

or

$$2x = \lambda$$

$$2y = 3\lambda$$

$$2z = -2\lambda$$

$$x + 3y - 2z = 12$$

Substituting each of the first three equations into the fourth,

$$\frac{\lambda}{2} + \frac{9\lambda}{2} + 2\lambda = 12 \quad\Rightarrow\quad \lambda = \frac{12}{7}$$

Hence

$$x = \frac{6}{7}, \quad , y = \frac{18}{7}, \quad z = -\frac{12}{7}$$

and

$$f\left(\frac{6}{7}, \frac{18}{7}, -\frac{12}{7}\right) = \frac{36 + 324 + 144}{49} = \frac{504}{49} \approx 10.28$$

To determine if this is a maximum or minimum we need to compare with some other point on the plane $x + 3y - 2z = 12$. Since there are two free variables and one dependent variable we can pick $x = 0$ and $y = 0$. Then $z = -6$, and

$$f(0, 0, -6) = 36 > 10.28$$

Hence the point we found is a minimum, as desired. ■

**Example 15.5** *Find the point on the plane* $2x - 3y + 5z = 19$ *that is nearest to the origin.*

*Solution.*The distance from the point $(x, y, z)$ to the origin is

$$d(x, y, z) = \sqrt{x^2 + y^2 + z^2}$$

We want to minimize this distance. Because differentiating square roots is messy, we observe that $d$ is minimized if and only if $d^2$ is also minimized. So we choose to minimize

$$f(x, y, z) = x^2 + y^2 + z^2$$

subject to the constraint

$$g(x, y, z) = 2z - 3y + 5z - 19 = 0$$

Using Lagrange's method $\nabla f = \lambda \nabla g$ so that

$$(2x, 2y, 2z) = \lambda (2, -3, 5)$$

Hence we have four equations in four unknowns,

$$x = \lambda$$

$$y = -\frac{-3\lambda}{2}$$

$$z = \frac{5\lambda}{2}$$

$$19 = 2x - 3y + 5z$$

From the fourth equation

$$19 = 2\lambda + \frac{9\lambda}{2} + \frac{25\lambda}{2} = 19\lambda$$

Using $\lambda = 1$ gives

$$x = 1, y = 3/2, z = 5/2$$

and the distance is

$$d^2 = 1 + 9/4 + 25/4 = 37/4 = 9.25$$

Picking any other point on the plane, say $x = 0, y = 0, z = 19/5$, the distance is $d = 361/5 \approx 72.2 > 9.25$. Hence the closest point to the origin is (1, -1.5, 2.5). ■

**Example 15.6** *Suppose the cost of manufacturing a particular type of box is such that the base of the box costs three times as much per square foot as the sides and top. Find the dimensions of the box that minimize the cost for a given volume.*

*Solution.* Let the bottom (and top) of the box have dimensions $a \times b$ and let the height of the box have dimension $h$. Then the area of the base $= ab$ and the area of the rest of the box is

$$\text{top} + (\text{left and right}) + (\text{front and back}) = ab + 2ah + 2bh$$

If the cost function is $c$ dollars per square foot that then total cost is

$$f(a, b, h) = 3abc + (ab + 2ah + 2bh)c = 4abc + 2ach + 2bch$$

We want to minimize this cost subject to the constraint

$$g(a, b, h) = abh - V = 0$$

The Lagrange formulation $\nabla f = \lambda \nabla g$ gives

$$\left(\frac{\partial f}{\partial a}, \frac{\partial f}{\partial b}, \frac{\partial f}{\partial h}\right) = \lambda \left(\frac{\partial g}{\partial a}, \frac{\partial g}{\partial b}, \frac{\partial g}{\partial h}\right)$$

$$(4bc + 2ch, 4ac + 2ch, 2ac + 2bc) = \lambda (bh, ah, ab)$$

We have four equations in four unknowns $a$, $b$, $h$, and $\lambda$:

$$4bc + 2ch = bh\lambda \tag{15.1}$$

$$4ac + 2ch = ah\lambda \tag{15.2}$$

$$2ac + 2bc = ab\lambda \tag{15.3}$$

$$abh = V \tag{15.4}$$

This is non-trivial to solve. We begin by solving 15.1 for $h$,

$$bh\lambda - 2ch = 4bc \Rightarrow h = \frac{4bc}{b\lambda - 2c} \tag{15.5}$$

Similarly, equation 15.2 gives

$$ah\lambda - 2ch = 4ac \Rightarrow h = \frac{4ac}{a\lambda - 2c} \tag{15.6}$$

Since both equations for $h$ must be equal,

$$\frac{4bc}{b\lambda - 2c} = \frac{4ac}{a\lambda - 2c} \tag{15.7}$$

$$\begin{aligned}
b(a\lambda - 2c) &= a(b\lambda - 2c) \\
ab\lambda - 2bc &= ab\lambda - 2ac \\
2bc &= 2ac \\
b &= a
\end{aligned} \tag{15.8}$$

Using equation 15.8 in 15.4 gives

$$h = V/a^2 \tag{15.9}$$

Using equation 15.8 in 15.3 gives

$$4ac = a^2\lambda \Rightarrow \lambda = 4c/a \tag{15.10}$$

Using equation 15.9 in 15.6 gives

$$\frac{V}{a^2} = \frac{4ac}{a\lambda - 2c} \tag{15.11}$$

Substituting equation 15.10 gives

$$\frac{V}{a^2} = \frac{4ac}{4c - 2c} = 2a \tag{15.12}$$

Hence

$$a = b = (V/2)^{1/3} \tag{15.13}$$

From equation 15.9

$$h = Va^{-2} = VV/2^{-2/3} = V^{1/3}2^{2/3} = (4V)^{1/3} \tag{15.14}$$

The solution is $a = b = (V/2)^{1/3}$ and $h = (4V)^{1/3}$. ■

**Theorem 15.2 Lagrange's Method: Two Constraints** *An extreme value of the the function $f(x, y, z)$ subject to the constraints*

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

*and*

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

*occurs when*

$$\nabla f = \lambda \nabla g + \mu \nabla h$$

*for some numbers $\lambda$ and $\mu$*

**Example 15.7** *Find the maximum and minimum of*

$$f(x, y, z) = -x + 2y + 2z$$

*on the ellipse formed by the intersection of the cylinder*

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

*with the plane*

$$y + 2z = 1$$

*Solution.* The idea is to find the extreme values of the function $f(x, y, z)$ subject to the constraints imposed by

$$g(x, y, z) = x^2 + y^2 - 2 = 0$$

and

$$h(x, y, z) = y + 2z - 1$$

Lagrange's method says that $\nabla f = \lambda \nabla g + \mu \nabla h$

$$(-1, 2, 2) = \lambda (2x, 2y, 0) + \mu (0, 1, 2)$$

Hence we have 5 equations in 5 unknowns,

$$2x\lambda = -1 \tag{15.15}$$

$$2y\lambda + \mu = 2 \tag{15.16}$$

$$2\mu = 2 \tag{15.17}$$

$$x^2 + y^2 = 2 \tag{15.18}$$

$$y + 2z = 1 \tag{15.19}$$

We can immediately elminate $\mu = 1$, to give

$$2y\lambda = 1 \tag{15.20}$$

Using equations 15.20 and 15.15 in equation 15.18 gives

$$2 = x^2 + y^2 = (1/2\lambda)^2 + (-1/2\lambda)^2 = 1/(2\lambda^2) \tag{15.21}$$

or

$$\lambda = \pm 1/2 \tag{15.22}$$

For $\lambda = 1/2$ we have $x = -1, y = 1$ and $z = (1 - y)/2 = 0$. At this point $f(-1, 1, 0) = 1 + 2(1) + 0 = 3$. For $\lambda = -1/2$ we have $x = 1, y = -1$, and $z = 1$. Here $f(1, -1, 1) = -1 - 2 + 2 = -1$. Hence the maximum of 3 occurs at (-1, 1, 0) and the minimum of -1 occurs at (1, -1, 1). ■

**Example 15.8** *The cone $z^2 = x^2 + y^2$ is cut by the plane $z = 1 + x + y$ in some curve $C$. Find the point on $C$ that is closest to the origin.*

*Solution* We need to minimize

$$f(x, y, z) = x^2 + y^2 + z^2 \tag{15.23}$$

subject to the two constraints

$$g(x, y, z) = z^2 - x^2 - y^2 = 0 \tag{15.24}$$

$$h(x, y, z) = z - 1 - x - y = 0 \tag{15.25}$$

Applying our usual technique, $\nabla f = \lambda \nabla g + \mu \nabla h$, so that

$$(2x, 2y, 2z) = \lambda(-2x, -2y, 2z) + \mu(-1, -1, 1) \tag{15.26}$$

or

$$2x = -2x\lambda - \mu \tag{15.27}$$

$$2y = -2y\lambda - \mu \tag{15.28}$$

$$2z = 2z\lambda + \mu \tag{15.29}$$

We can elminate $\mu$ by adding equation 15.29 to each of equations 15.27 and 15.28 to give

$$2x + 2z = 2\lambda(z - x) \tag{15.30}$$

$$2y + 2z = 2\lambda(z - y) \tag{15.31}$$

or

$$\lambda = \frac{z + x}{z - x} \tag{15.32}$$

$$\lambda = \frac{z + y}{z - y} \tag{15.33}$$

Equating the two expressions for $\lambda$,

$$\begin{aligned} \frac{z+x}{z-x} &= \frac{z+y}{z-y} \\ (z+x)(z-y) &= (z+y)(z-x) \\ z^2 + xz - yz - xy &= z^2 + yz - xz - xy \\ xz - yz &= yz - xz \\ yz &= xz \end{aligned}$$

Hence $y = x$ or $z = 0$.

**Case 1: z=0**

From the constraint $g$ we have $x^2 + y^2 = 0$ which means $x = 0$ and $y = 0$. The origin is the only point on the cone that satisfies $z = 0$. But this point is not on the plane $z = 1 + x + y$. So there is no solution here.

**Case 2: x=y**

The first constraint ($g$) gives $z^2 = 2x^2$. The second constraint ($h$) gives $z = 1 + 2x$, hence

$$(1 + 2x)^2 = 2x^2 \Rightarrow 1 + 4x + 4x^2 = 2x^2 \Rightarrow 2x^2 + 4x + 1 = 0$$

$$y = x = \frac{-4 \pm \sqrt{8}}{4} = -1 \pm \frac{1}{\sqrt{2}}$$

hence

$$z = 1 + 2\left(-1 \pm \frac{1}{\sqrt{2}}\right) = -1 \pm \sqrt{2}$$

The extremum occur at

$$\left(-1+\frac{1}{\sqrt{2}},-1+\frac{1}{\sqrt{2}},-1+\sqrt{2}\right)$$

and

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

The minimum is at

$$\begin{aligned}
f\left(-1+\frac{1}{\sqrt{2}},-1+\frac{1}{\sqrt{2}},-1+\sqrt{2}\right) &= 2\left(-1+\frac{1}{\sqrt{2}}\right)^2+\left(-1+\sqrt{2}\right)^2 \\
&= 2-\frac{4}{\sqrt{2}}+1+1-2\sqrt{2}+2 \\
&= 4-4\sqrt{2}
\end{aligned}$$

and the maximum occurs when

$$\begin{aligned}
f\left(-1-\frac{1}{\sqrt{2}},-1-\frac{1}{\sqrt{2}},-1-\sqrt{2}\right) &= 2\left(-1-\frac{1}{\sqrt{2}}\right)^2+\left(-1-\sqrt{2}\right)^2 \\
&= 2+\frac{4}{\sqrt{2}}+1+1+2\sqrt{2}+2 \\
&= 6+4\sqrt{2} \quad \blacksquare
\end{aligned}$$

# Lecture 16

# Double Integrals over Rectangles

Recall how we defined the Riemann integral as an area in Calculus I: to find the area under a curve from $a$ to $b$ we partitioned the interval $[a,b]$ into the set of numbers

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$$

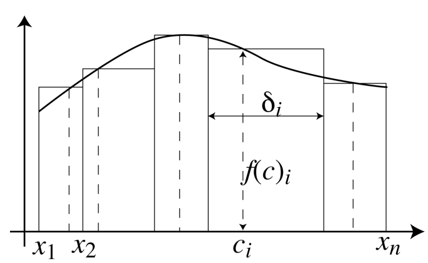
We then chose a sequence of points, one inside each interval $c_i \in (x_{i-1}, x_i)$,

$$x_{i-1} < c_i < x_i$$

and approximated the area $A_i$ under the curve from $x_{i-1}$ to $x_i$ with a rectangle with a base width of $\delta_i = x_i - x_{i-1}$ and height $f(c_i)$, so that

$$A_i = \delta_i f(c_i)$$

Figure 16.1: Calculation of the Riemann Sum to find the area under the curve from $a = x_1$ to $b = x_n$ approximates the area by a sequence of rectangles and then takes the limit as the number of rectangles becomes large.



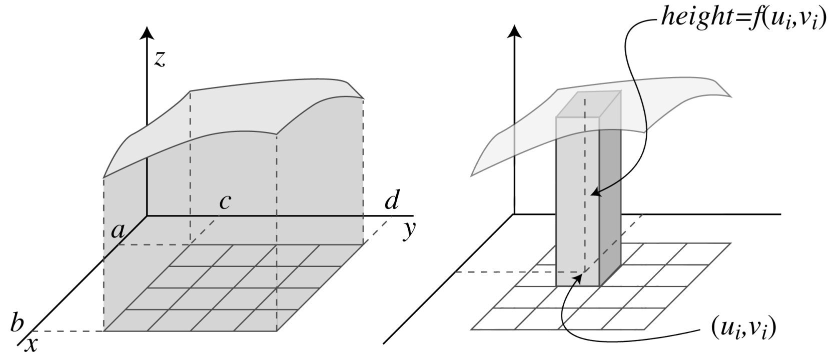

The total area under the curve from $a$ to $b$ was then approximated by a sequence of such rectangles, as the **Riemann Sum**

$$A \approx \sum_{i=1}^{n} A_i = \sum_{i=1}^{n} \delta_i f(c_i)$$

We then defined **Riemann Integral** as the limit

$$\int_a^b f(x)dx = \lim_{\delta_1,\delta_2,\dots,\delta_n \to 0, n\to\infty} \sum_{i=1}^{n} f(c_i)\delta_i$$

Figure 16.2: Left: A Riemann Sum in 3D can also be used to estimate the *volume* between a surface described some some function $f(x, y)$ and a rectangle beneath it in the $xy$-plane. Right: construction of the boxes of height $f(u_i, v_i)$ with base in rectangle $i$.




We now want to generalize this procedure to find the volume of the solid between the surface $z = f(x, y)$ and the $x$-$y$ plane over some rectangle $R$ as illustrated in figure 16.2. The surface $z = f(x, y)$ forms the top of the volume, and rectangle

$$R = [a, b] \times [c, d]$$

in the $x$-$y$ plane forms the bottom of the volume. We can approximate this volume by filling it up with rectangular boxes, as illustrated on the right-hand sketch in figure 16.2.

1. Divide the $[a, b] \times [c, d]$ rectangle in the $x$-$y$ plane into little rectangles. Although they are illustrated as squares in the figure, they do not have to be squares.

2. Number the little rectangles in the $xy$ plane from $i = 1$ to $i = n$.

3. Let the area of rectangle $i$ be $\Delta A_i$

4. Pick one point in each rectangle and label it $(u_i, v_i)$. The $i^{th}$ point does not have to be in the center of mini-rectangle $i$, just somewhere within the rectangle. The points in rectangle $i$ and rectangle $j$ can be in different locations within the rectangle.

5. The distance of the shaded surface about the point $u_i, v_i$ is $f(u_i, v_i)$. Draw a box of height $f(u_i, v_i)$ whose base is given by rectangle $i$.

6. The volume of box $i$ is $V_i = f(u_i, v_i)\Delta A_i$.

7. The total volume under the surface is

$$\sum_{i=1}^{n} V_i = \sum_{i=1}^{n} f(u_i, v_i)\Delta A_i$$

We then obtain the double integral by taking the limit as the number of squares becomes large and their individual sizes approach zero.

**Definition 16.1** *Let $f(x, y) : R \subset \mathbb{R}^2 \mapsto \mathbb{R}$ be a function of two variables defined on a rectangle $R$. Then the* **Double Integral of $f$ over $R$** *is defined as*

$$\iint_R f(x, y)dA = \lim_{n\to\infty, \Delta A_i \to 0} \sum_{i=1}^{n} f(u_i, v_i)\Delta A_i$$

*if this limit exists. The summation on the right is called a* **Riemann Sum**. *If the integral exits, the function $f(x, y)$ is said to be* **integrable over $R$**.

**Theorem 16.1** *If $f(x, y)$ is bounded on a closed rectangle $R$, and is continuous everywhere except for possibly only finitely many smooth curves in $R$, then $f(x, y)$ is integrable on $R$. Furthermore, if $f(x, y)$ is continuous everywhere on $R$ then $f(x, y)$ is integrable on $R$.*

**Example 16.1** *Approximate*

$$\iint_R f(x, y)dA$$

*using a Riemann Sum for*

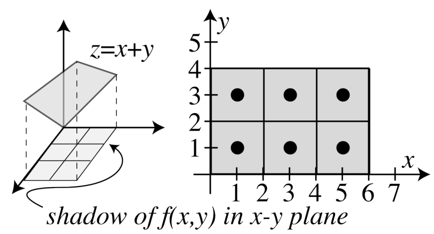
$$f(x, y) = x + y$$

*and $R$ is the rectangle $[0, 6] \times [0, 4]$, with a partition that breaks $R$ into squares by the lines $x = 2$, $x = 4$, and $y = 2$, and choosing points at the center of each square to define the Riemann sum.*

*Solution.* See figure 16.3. Choose points $(u_i, v_i)$ in the center of each square, at $(1, 1), (3, 1), (5, 1), (1, 3), (3, 3), (5, 3)$, as shown. Since each $\Delta A_i = 4$, the Riemann Sum is then

$$\begin{aligned}
\iint_R f(x, y)dA &\approx \textstyle\sum_{i=1}^{n} f(u_i, v_i)\Delta A_i \\
&= 4\,(f(1, 1) + f(3, 1) + f(5, 1) + f(1, 3) + f(3, 3) + f(5, 3)) \\
&= 4\,(2 + 4 + 6 + 4 + 6 + 8) = 120 \quad \blacksquare
\end{aligned}$$

Figure 16.3: A partition of the rectangle $R = [0,6] \times [0,4]$ into size equally-sized squares can be used to calculate $\int_R f(x,y)dA$ as described in example 16.1.




**Theorem 16.2** *Suppose that $R = U + V$ where $U$, $V$, and $R$ are rectangles such that $U$ and $V$ do not overlap (except on their boundary) then*

$$\iint\limits_R f(x,y)dA = \iint\limits_U f(x,y)dA + \iint\limits_V f(x,y)dA$$

*in other words, the integral over the union of non-overlapping rectangles is the sum of the integrals.*

**Theorem 16.3 Iterated Integrals.** *If $R = [a,b] \times [c,d]$ is a rectangle in the xy plane and $f(x,y)$ is integrable on $R$, then*

$$\iint\limits_R f(x,y,)dA = \int_a^b \left( \int_c^d f(x,y)dy \right) dx = \int_c^d \left( \int_a^b f(x,y)dx \right) dy$$

*The parenthesis are ordinarily omitted in the double integral, hence the order of the differentials dx and dy is crucial,*

$$\iint\limits_R f(x,y,)dA = \int_a^b \int_c^d f(x,y)dydx = \int_c^d \int_a^b f(x,y)dxdy$$

**Note:** If $R$ is not a rectangle, then you can not reverse the order of integration as we did in theorem 16.3.

**Example 16.2** *Repeat example 1 as an iterated integral.*

*Solution.* The corresponding iterated integral is

$$\begin{aligned} \iint_R f(x,y)dA &= \int_0^6 \left( \int_0^4 (x+y)dy \right) dx \\ &= \int_0^6 \left( \int_0^4 xdy + \int_0^4 ydy \right) dx \end{aligned}$$

In the first integral inside the parenthesis, $x$ is a constant because the integration is over $y$, so we can bring that constant out of the integral:

$$\iint_R f(x,y)dA = \int_0^6 \left( x \int_0^4 dy + \int_0^4 ydy \right) dx$$

Next, we evaluated each of the two inner integrals:

$$\begin{aligned}
\iint_R f(x,y)dA &= \int_0^6 \left( x\left( y|_0^4 \right) + \left( \frac{1}{2}y^2 \Big|_0^4 \right) \right) dx \\
&= \int_0^6 \left( x(4-0) + \left( \frac{1}{2}(16) - \frac{1}{2}(0) \right) \right) dx \\
&= \int_0^6 (4x+8)dx
\end{aligned}$$

After the inside integrals have been completed what remains only depends on the other variable, which is x. There is no more y in the equation:

$$\begin{aligned}
\iint_R f(x,y)dA &= \int_0^6 (4x+8)dx \\
&= (2x^2+8x)\big|_0^6 \\
&= 2(36)+8(6)-2(0)-8(0) = 72+48 = 120
\end{aligned}$$

Usually the Riemann sum will only yield a good approximation, but not the exact answer, with the accuracy of the approximation improving as the squares get smaller. It is only a coincidence that in this case we obtained the exact, correct answer. ■

**Example 16.3** *Evaluate*

$$\int_4^7 \left( \int_1^5 (3x+12y)dx \right) dy$$

*Solution.* We evaluate the inner integral first. Since the inner integral is an integral over $x$, within the inner integral, we can treat $y$ as a constant:

$$\begin{aligned}
\int_4^7 \left( \int_1^5 (3x+12y)dx \right) dy &= \int_4^7 \left( 3\int_1^5 xdx + 12y\int_1^5 dx \right) dy \\
&= \int_4^7 \left( 3\left( \frac{1}{2}x^2 \Big|_1^5 \right) + 12y\left( x|_1^5 \right) \right) dy \\
&= \int_4^7 \left( 3\left( \frac{25}{2} - \frac{1}{2} \right) + 12y(5-1) \right) dy \\
&= \int_4^7 (36+48y)\,dy
\end{aligned}$$

After the integration over $x$ is completed, there should not be any $x$'s left in the equation – they will all have been "integrated out" and what remains will only depend on $y$.

$$\begin{aligned}
\int_4^7 \left( \int_1^5 (3x+12y)dx \right) dy &= \int_4^7 (36+48y)\,dy \\
&= (36y+24y^2)\big|_4^7 \\
&= [(36)(7)+24(49)-(36)(4)-24(16)] \\
&= 252+1176-144-384 = 900
\end{aligned}$$

Whenever we calculate a definite integral, the final answer will always be a number and will not have any variables left in it. ■

**Example 16.4** *Find* $\int_0^2 \int_0^2 (x^2 + y^2) dx dy$

*Solution.*

$$\begin{aligned}
\int_0^2 \int_0^2 (x^2 + y^2) dx dy &= \int_0^2 \left( \int_0^2 x^2 dx + \int_0^2 y^2 dx \right) dy \\
&= \int_0^2 \left[ \frac{1}{3} x^3 \Big|_0^2 + y^2 \int_0^2 dx \right] dy \\
&= \int_0^2 \left[ \frac{8}{3} + 2y^2 \right] dy \\
&= \left( \frac{8}{3} y + \frac{2}{3} y^3 \right) \Big|_0^2 \\
&= \frac{8}{3}(2) + \frac{2}{3}(8) = \frac{32}{3} \quad \blacksquare
\end{aligned}$$

## Volumes and Areas

The **Area** $A$ of a rectangle $R$ in the $xy$ plane is

$$A = \iint_R dx dy$$

The **Volume** $V$ between the surface $z = f(x, y)$ and the rectangle $R$ in the $xy$ plane is

$$V = \iint_R f(x, y) dx dy$$

These formulas for area and volume still hold even if $R$ is not a rectangle (even though we have not yet defined the concept of an integral over a non-rectanglular domain).

**Example 16.5** *Find the area of the rectangle* $[0, l] \times [0, w]$.

*Solution.* Using double integrals,

$$A = \iint_R dx dy = \int_0^w \int_0^l dx dy = lw \quad \blacksquare$$

**Example 16.6** *Find the volume of the solid under the plane* $z = 2x + 4y$ *and over the rectangle* $[3, 12] \times [2, 4]$.

*Solution.* The volume is given by the iterated integral

$$\begin{aligned}
\int_3^{12}\int_2^4 (2x+4y)dydx &= \int_3^{12}\int_2^4 2xdydx + \int_3^{12}\int_2^4 4ydydx \\
&= \int_3^{12} 2x\int_2^4 dydx + \int_3^{12} 4\int_2^4 ydydx \\
&= \int_3^{12} 2x(4-2)dx + \int_3^{12} 4[(y^2/2)\big|_2^4\,dx \\
&= \int_3^{12} 4xdx + \int_3^{12} 4(8-2)dx \\
&= 4\int_3^{12} xdx + 24\int_3^{12} dx \\
&= 4\ (x^2/2)\big|_3^{12} + 24(12-3) \\
&= 4(72-9/2) + 216 = 288 - 18 + 216 = 486 \ \blacksquare
\end{aligned}$$

**Example 16.7** *Find* $\displaystyle\int_0^1\int_0^2 \frac{y}{1+x^2}dydx$

*Solution.* Since the inner integral is over $y$, the denominator, which depends only on $x$, can be brought out of the inner integral:

$$\begin{aligned}
\int_0^1\int_0^2 \frac{y}{1+x^2}dydx &= \int_0^1 \frac{1}{1+x^2}\int_0^2 ydydx \\
&= \int_0^1 \frac{1}{1+x^2}\ (y^2/2)\big|_0^2\,dx \\
&= 2\int_0^1 \frac{1}{1+x^2}dx \\
&= 2\ \tan^{-1} x\big|_0^1 \\
&= 2(\tan^{-1} 1 - \tan^{-1} 0) \\
&= 2(\pi/4 - 0) = \pi/2 \ \blacksquare
\end{aligned}$$

**Example 16.8** *Find the iterated integral* $\displaystyle\int_2^5\int_1^3 \frac{x}{(x^2+y)^2}dxdy$

*Solution.* There is nothing that can be factored out of the inner integral. We will first consider just the inside integral,

$$\int_1^3 \frac{x}{(x^2+y)^2}dx$$

where $y$ is a constant. If we make the substitution

$$u = x^2 + y$$

then $du = 2dx$, because $y$ is a constant. When $x$=1, $u$=1+$y$; When $x$=3, $u$=9+$y$. Therefore,

$$\begin{aligned}\int_1^3 \frac{x}{(x^2+y)^2}dx &= \frac{1}{2}\int_{1+y}^{9+y}\frac{du}{u^2}\\ &= \frac{1}{2}\int_{1+y}^{9+y} u^{-2}du\\ &= \frac{1}{2}\left(-u^{-1}\right)\Big|_{1+y}^{9+y}\\ &= \frac{1}{2}\left[\frac{-1}{9+y}+\frac{1}{1+y}\right]\end{aligned}$$

Substituting this back into the original iterated integral, we find that

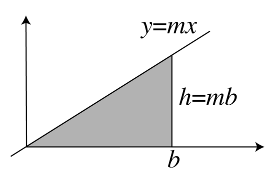
$$\begin{aligned}\int_2^5\int_1^3 \frac{x}{(x^2+y)^2}dxdy &= \int_2^5 \frac{1}{2}\left[\frac{-1}{9+y}+\frac{1}{1+y}\right]dy\\ &= \frac{1}{2}\left[-\ln(9+y)+\ln(1+y)\right]\Big|_2^5\\ &= \frac{1}{2}[(-\ln 14+\ln 6)-(-\ln 11+\ln 3)]\\ &= \frac{1}{2}\ln\frac{(6)(11)}{(14)(3)} = \frac{1}{2}\ln\frac{11}{7} = \ln\sqrt{11/7} \approx 0.226 \blacksquare\end{aligned}$$

When we integrate over non-rectangular regions, the limits on the inner integral will be functions that depend on the variable in the outer integral.

**Example 16.9** *Find the area of the triangle*

$$R = \{(x,y) : 0 \le x \le b, 0 \le y \le mx\}$$

*using double integrals and show that it gives the usual formula y = (base)(height)/2*




*Solution.*

$$A = \iint_R dxdy = \int_0^b\int_0^{mx} dydx = \int_0^b mxdx = mb^2/2$$

We can find the height of the triangle by oberserving that at $x = b$, $y = mb$, therefore the height is $h = mb$, and hence

$$A = mb^2/2 = (mb)(b)/2 = (height)(base)/2 \blacksquare$$

We will consider integrals over non-rectangular regions in greater detail in the following section.

# Lecture 17

# Double Integrals over General Regions

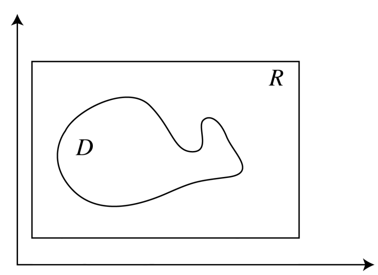
**Definition 17.1** *Let $f(x, y) : D \subset \mathbb{R}^2 \mapsto \mathbb{R}$. Then we say that that $f$ is integrable on $D$ if for some rectangle $R$, which is oriented parallel to the $xy$ axes, the function*

$$F(x, y) = \begin{cases} f(x, y) & \textit{if } (x, y) \in D \\ 0 & \textit{otherwise} \end{cases}$$

*is* **integrable***, and we definitionine the* **double integral over the set *D*** *by*

$$\iint_D f(x, y) dA = \iint_R F(x, y) dA$$

Figure 17.1: The integral over a non-rectangular region is definitionined by extending the domain to an enclosing rectangle.




In other words, we extend the domain of $f(x, y)$ from $D$ to some rectangle including $D$, by definitionining a new function $F(x, y)$ that is zero everywhere outside $D$, and equal to $f$ on $D$, and then use our previous definitioninition of the integral over a rectangle. The double integral has the following properties.

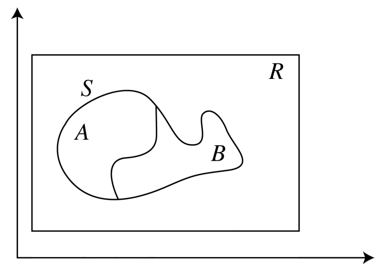
**Theorem 17.1 Linearity**. *Suppose that $f(x, y)$ and $g(x, y)$ are integrable over $D$. The for any constants $\alpha, \beta \in \mathbb{R}$*

$$\iint_D (\alpha f(x, y) + \beta g(x, y))\, dA = \alpha \iint_D f(x, y) dA + \beta \iint_D g(x, y) dA$$

**Theorem 17.2 Additivity**. *Suppose that $f(x,y)$ is integrable on $S$, that $S = A \cup B$, and that any overlap between $A$ and $B$ occurs only on a smooth curve. Then*

$$\iint_S f(x,y)dA = \iint_A f(x,y)dA + \iint_B f(x,y)dA$$

Figure 17.2: The integral is additive over regions that share (at most) a common smooth curve. Here $\iint_S f = \iint_A f + \iint_B f$.




**Theorem 17.3 Comparison Property**. *Suppose that $f(x,y)$ and $g(x,y)$ are integrable functions on a set $S$ such that*

$$f(x,y) \le g(x,y)$$

*for all $(x,y) \in S$. Then*

$$\iint_S f(x,y)dA \le \iint_S g(x,y)dA$$

The simplest non-rectangular regions to integrate over are called ***x*-simple** and ***y*-simple** regions, as illustrated in the following example. A region is called **x-simple** if it can be expressed as a union of line segments parallel to the $x$-axis; it is called ***y*-simple** if it can be expressed as a union of line segments parallel to the $y$-axis. The following example considers one $y$-simple set.

**Example 17.1** *Set up the integral*

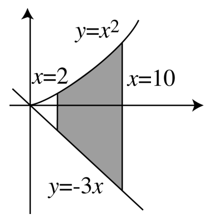
$$\iint_s f(x,y)dA$$

*where $S$ is the set bounded above by the curve $y = x^2$, below by the line $y = -3x$, on the left by the line $x = 2$, and on the right by the line x=10.*

*Solution.* As $x$ increases from $x$=2 to $x$=10, we can trace out the figure with a vertical rectangle that goes from the lower boundary $y = -3x$ to the upper boundary $y = x^2$. In other words: For all $x$ between 2 and 10 (this is the outer integral), include all $y$ between $y = -3x$ and $y = x^2$. Therefore

$$\iint_s f(x,y)dA = \int_2^{10} \int_{-3x}^{x^2} f(x,y)dx \ \blacksquare$$

Figure 17.3: The integral over this y-simple region is examined in the example.




**Definition 17.2** *A region $S$ is called* **x-simple** *if there are numbers $c, d \in \mathbb{R}$ and functions $g(y), h(y) : [c, d] \mapsto \mathbb{R}$ such that*

$$S = \{(x, y) : g(y) \le y \le h(y) \;\; and \;\; y \in [c, d]\}$$

**Definition 17.3** *A region $S$ is called* **y-simple** *if there are numbers $a, b \in \mathbb{R}$ and functions $f(x), g(x) : [a, b] \mapsto \mathbb{R}$ such that*

$$S = \{(x, y) : f(x) \le x \le g(x) \;\; and \;\; x \in [a, b]\}$$

Figure 17.4: Examples of x-simple and y-simple regions.

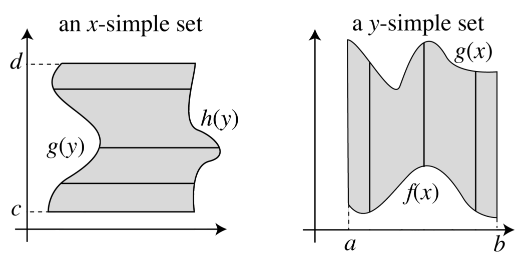
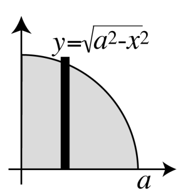



Figure 17.5: The quarter of the circle of radius $a$ centered at the origin that is in the first quadrant.




**Example 17.2** *Set up the integral $\iint_S f(x, y)dA$ over the portion of the interior of a circle of radius a centered at the origin that lies in the first quadrant.*

*Solution.* The equation of a circle is

$$x^2 + y^2 = a^2$$

Solving for $y$,

$$y = \pm\sqrt{a^2 - x^2}$$

The set $S$ can be described as follows: As $x$ increases from $x = 0$ to $x = a$, $y$ at $x$ increases from $y = 0$ to $y = \sqrt{a^2 - x^2}$. Therefore

$$\iint_s f(x,y)dA = \int_0^a \int_0^{\sqrt{a^2-x^2}} f(x,y)dydx \quad \blacksquare$$

**Example 17.3** *Find a formula for the area of a circle of radius a using double integrals.*

*Solution.* We can place the center of the circle at the origin, so that its equation is given by $x^2 + y^2 = a^2$, as in the previous example. The circle can be thought of as the set

$$\{(x,y) : -a \le x \le a, -\sqrt{a^2 - x^2} \le y \le \sqrt{a^2 - x^2}\}$$

so that

$$\begin{aligned}
A &= \iint_R dxdy = \int_{-a}^{a} \int_{-\sqrt{a^2-x^2}}^{\sqrt{a^2-x^2}} dydx \\
&= \int_{-a}^{a} \left( y|_{-\sqrt{a^2-x^2}}^{\sqrt{a^2-x^2}} \right) dx \\
&= \int_{-a}^{a} \left( \sqrt{a^2 - x^2} - -\sqrt{a^2 - x^2} \right) dx \\
&= 2\int_{-a}^{a} \sqrt{a^2 - x^2} dx
\end{aligned}$$

By symmetry, since the integrand is even, we also have

$$A = 4\int_0^a \sqrt{a^2 - x^2} dx$$

According to integral formula (54) on the inside book jacket

$$\int \sqrt{a^2 - x^2} du = \frac{x}{2}\sqrt{a^2 - x^2} + \frac{x^2}{2}\sin^{-1}\frac{x}{a}$$

Therefore

$$\begin{aligned}
A &= 4\int_0^a \sqrt{a^2 - x^2} dx = \left( \frac{x}{2}\sqrt{a^2 - x^2} + \frac{x^2}{2}\sin^{-1}\frac{x}{a} \right)\Bigg|_0^a \\
&= 4\left[ \left( \frac{a}{2}\sqrt{a^2 - a^2} + \frac{a^2}{2}\sin^{-1}\frac{a}{a} \right) - \left( \frac{0}{2}\sqrt{a^2 - 0^2} + \frac{0^2}{2}\sin^{-1}\frac{0}{a} \right) \right] \\
&= 4[0 + \frac{a^2}{2}\sin^{-1}(1) - 0 - 0] = 4(a^2/2)(\pi/2) \\
&= \pi a^2 \quad \blacksquare
\end{aligned}$$

**Example 17.4** *Find the iterated integral*

$$\int_0^1 \int_0^{y^2} 2ye^x dxdy$$

*and sketch the set S.*

*Solution* We write the integral as

$$\begin{aligned}\int_0^1 \int_0^{y^2} 2ye^x dxdy &= \int_0^1 2y \int_0^{y^2} e^x dxdy = \int_0^1 2y \left\{ e^x|_0^{y^2} \right\} dy \\ &= \int_0^1 2y[e^{y^2} - e^0]dy = \int_0^1 2ye^{y^2} dy - \int_0^1 2ydy\end{aligned}$$

The first integral on the right can be solve with the substitution

$$u = y^2, du = 2dy$$

when $y = 0$, $u = 0$, and when $y = 1$, $u = 1$. Thus $y$=0, $u$=0; when $y$=1, $u$=1. Thus

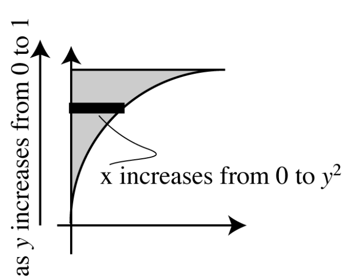
$$\int_0^1 2ye^{y^2} dy = \int_0^1 e^u du = e^u|_0^1 = e^1 - e^0 = e - 1$$

The second integral is straightforward:

$$\int_0^1 2ydy = y^2\Big|_0^1 = 1 - 0 = 1$$

Therefore

Figure 17.6: The region in example 4.




$$\begin{aligned}\int_0^1 \int_0^{y^2} 2ye^x dxdy &= \int_0^1 2ye^{y^2} dy - \int_0^1 2ydy \\ &= (e-1) - 1 = e - 2\end{aligned}$$

We can draw the domain by observing that the outer integral – the one over y – increase from zero to 1. For any fixed $y$ within this region, $x$ increases for 0 to $y^2$. As we move up along the $y$ axis (the outer integral) we need to increase from $x = 0$ to $x = y^2$, i.e., the curve of $y = \sqrt{x}$. ■

**Example 17.5** *Find the iterated integral*

$$\int_{-2}^{12}\int_{-x^2}^{x+3}(x+3y)dxdy$$

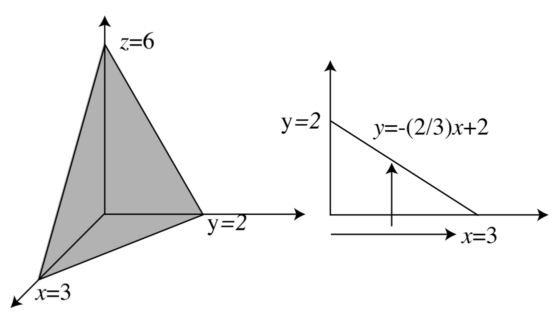
*Solution.* Separating the integral into two parts and factoring what we can gives

$$\begin{aligned}
\int_{-2}^{12}\int_{-x^2}^{x+3}(x+3y)dydx &= \int_{-2}^{12}\int_{-x^2}^{x+3}xdydx+\int_{-2}^{12}\int_{-x^2}^{x+3}3ydydx\\
&= \int_{-2}^{12}x\int_{-x^2}^{x+3}dydx+3\int_{-2}^{12}\int_{-x^2}^{x+3}ydydx\\
&= \int_{-2}^{12}x(y|_{-x^2}^{x+3})dx+3\int_{-2}^{12}(\frac{1}{2}y^2|_{-x^2}^{x+3})dx\\
&= \int_{-2}^{12}x(x+3+x^2)dx+\frac{3}{2}\int_{-2}^{12}[(x+3)^2-(-x^2)^2]dx\\
&= \int_{-2}^{12}(x^2+3x+x^3)dx+\frac{3}{2}\int_{-2}^{12}(x+3)^2dx-\frac{3}{2}\int_{-2}^{12}x^4dx
\end{aligned}$$

Integrating and plugging in numbers, we find that

$$\begin{aligned}
\int_{-2}^{12}\int_{-x^2}^{x+3}(x+3y)dydx &= (\frac{1}{3}x^3+\frac{3}{2}x^2+\frac{1}{4}x^4)|_{-2}^{12}+\frac{3}{2}\frac{1}{3}(x+3)^3|_{-2}^{12}-\frac{3}{2}\frac{1}{5}x^5|_{-2}^{12}\\
&= \frac{1}{3}(12)^3+\frac{3}{2}(12)^2+\frac{1}{4}(12)^4-\frac{1}{3}(-8)-\frac{3}{2}(4)-\frac{1}{2}(16)\\
&\quad +\frac{1}{2}(15^3-(-8))-\frac{3}{10}(12^5-(-32))\\
&= 576+216+5184+\frac{8}{3}-6-8+1691.5-124432\\
&\approx -116775.83 \ \blacksquare
\end{aligned}$$

Figure 17.7: The plane $z = 6 - 2x - 3y$ and its cross-section on the $xy$ plane.




**Example 17.6** *Find the volume of the tetrahedron bounded by the coordinate planes and the plane $z = 6 - 2x - 3y$.*

*Solution* Dividing the equation of the plane by 6 and rearranging gives

$$\frac{x}{3}+\frac{y}{2}+\frac{z}{6}=1$$

This tells us that the plane intersects the coordinate axes at $x$= 3, $y$=2, and $z$=6 as illustrated in figure 17.7. We can find the intersection of the given plane with the $xy$-plane by setting $z = 0$. The cross-section is illustrated on the right-had side of figure 17.7, and the intersection is the line

$$2x+3y=6 \Rightarrow \frac{x}{3}+\frac{y}{2}=1$$

Solving for $y$,

$$y=-\frac{2}{3}x+2$$

If we integrate first over $y$ (in the inner integral), then going from the bottom to the top of the triangle formed by the $x$ and $y$ axes and the line $y = (2/3)x + 2$, our limits are $y = 0$ (on the bottom) to $y = -(2/3)x + 2$ (on the top).. Then we integrate over $x$ (in the outer integral) going from left to right, with limits $x = 0$ to $x = 3$,

$$\begin{aligned}
V &= \int_0^3 \int_0^{-(2/3)x+2} (6-2x-3y)dx \\
&= \int_0^3 \left(6y-2xy-\frac{3}{2}y^2\right)\Big|_0^{-(2/3)x+2} dx \\
&= \int_0^3 \left[6\left(-\frac{2}{3}x+2\right)-2x\left(-\frac{2}{3}x+2\right)-\frac{3}{2}\left(-\frac{2}{3}x+2\right)^2\right] dx
\end{aligned}$$

Expanding the factors in the integrand gives

$$\begin{aligned}
V &= \int_0^3 \left[-4x+12+\frac{4}{3}x^2-4x-\frac{3}{2}\left(\frac{4}{9}x^2-\frac{8}{3}x+4\right)\right] dx \\
&= \int_0^3 \left[-4x+12+\frac{4}{3}x^2-4x-\frac{2}{3}x^2+4x-6\right] dx \\
&= \int_0^3 \left[-4x+6+\frac{2}{3}x^2\right] dx \\
&= \left(-2x^2+6x+\frac{2}{9}x^3\right)\Big|_0^3 \\
&= -2(9)+6(3)+\frac{2}{9}(27) = -18+18+6=6 \ \blacksquare
\end{aligned}$$

**Example 17.7** *Find the volume of the tetrahedron bounded by the coordinate planes and the plane* $3x + 4y + z - 12 = 0$.

*Solution.* This example is exactly like the previous example. Dividing the equation through by 12 we determine that the given plane intersects the coordinate axes at $x = 4$,$y = 3$, and $z = 12$. The tetrahedron drops a shadow in the $xy$ plane onto a triangle formed by the origin and the points (0,3) and (4,0). The equation of the line between these last two points is

$$y = -(3/4)x + 3$$

. Hence the volume of the tetrahedron is given by the integral

$$V = \int_0^4 \int_0^{9-(3/4)x+3} (12 - 3x - 4y) dydx$$

Alternatively, we could integrate first over $x$ and then over $y$ as

$$V = \int_0^3 \int_0^{-(4/3)y+4} (12 - 3x - 4y) dxdy$$

Either integral will give the correct solution. Looking at the first integral,

$$\begin{aligned}
V &= \int_0^4 \int_0^{-(3/4)x+3} (12 - 3x - 4y) dydx \\
&= \int_0^4 \left(12y - 3xy - 2y^2\right)\Big|_0^{-(3/4)x+2} dx \\
&= \int_0^4 \left[12\left(-\frac{3}{4}x + 3\right) - 3x\left(-\frac{3}{4}x + 3\right) - 2\left(-\frac{3}{4}x + 3\right)^2\right] dx \\
&= \int_0^4 \left[-9x + 36 + \frac{9}{4}x^2 - 9x - 2\left(\frac{9}{16}x^2 - \frac{9}{2}x + 9\right)\right] dx \\
&= \int_0^4 \left[36 - 18x + \frac{9}{4}x^2 - \frac{9}{8}x^2 + 9x - 18\right] dx \\
&= \int_0^4 \left[18 - 9x + \frac{9}{8}x^2 +\right] dx \\
&= \left[18x - \frac{9}{2}x^2 + \frac{9}{(8)(3)}x^3\right]\Big|_0^4 \\
&= 18(4) - \frac{9}{2}(16) + \frac{3(64)}{8} \\
&= 72 - 72 + 24 = 24 \ \blacksquare
\end{aligned}$$

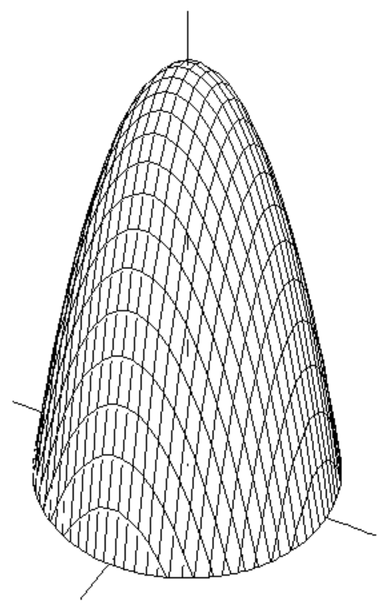
**Example 17.8** *Find the volume of the solid in the first octant bounded by the paraboloid*

$$z = 9 - x^2 - y^2$$

*and the xy-plane.*

*Solution* The domain of the integral is the intersection of the paraboloid with the

Figure 17.8: The solid formed by the paraboloid $z = 9 - x^2 - y^2$ and the $xy$ plane.



$xy$ plane. This occurs when $z = 0$. So that means the domain is

$$0 = 9 - x^2 - y^2 \Rightarrow y^2 + x^2 = 3^2$$

which is a circle of radius 3 centered at the origin. Since we only want the part of the circle that is in the first quadrant, we need $x$ and $y$ both positive. To cover the upper right hand corner of a circle of radius 3 centered at the origin, we can let $x$ increase from 0 to 3 and $y$ increase from 0 to $y = \sqrt{9 - x^2}$ Therefore the volume is

$$\begin{aligned} V &= \int_0^3 \int_0^{\sqrt{9-x^2}} (9 - x^2 - y^2) dy dx \\ &= \int_0^3 (9 - x^2) \int_0^{\sqrt{9-x^2}} dy dx - \int_0^3 \int_0^{\sqrt{9-x^2}} y^2 dy dx \\ &= \int_0^3 (9 - x^2) \left( y|_0^{\sqrt{9-x^2}} \right) dx - \frac{1}{3} \int_0^3 y^3|_0^{\sqrt{9-x^2}} dx \\ &= \int_0^3 (9 - x^2)^{3/2} dx - \frac{1}{3} \int_0^3 (9 - x^2)^{3/2} dx \\ &= \frac{2}{3} \int_0^3 (9 - x^2)^{3/2} dx \end{aligned}$$

Making the substitution $x = 3\cos u$ in $\int(9-x^2)^{3/2}dx$ we have $dx = -3\sin u du$ hence

$$\begin{aligned}\int(9-x^2)^{3/2}dx &= \int(9-9\cos^2 u)^{3/2}(-3\sin u)du\\ &= \int 9^{3/2}(1-\cos^2 u)^{3/2}(-3)\sin u du\\ &= -3(3^2)^{3/2}\int(\sin^2 u)^{3/2}\sin u du\\ &= -3(3^3)\int\sin^4 u du\\ &= -81\int(\sin^2 u)(\sin^2 u)du\end{aligned}$$

Using the trigonometric relationship

$$\sin^2 u = \frac{1}{2}(1-\cos 2u)$$

gives

$$\begin{aligned}\int(9-x^2)^{3/2}dx &= -81\int\frac{1}{2}(1-\cos 2u)\frac{1}{2}(1-\cos 2u)du\\ &= -\frac{81}{4}\int(1-2\cos 2u+\cos^2 2u)du\end{aligned}$$

Use the substitution

$$\cos^2 u = \frac{1}{2}(1+\cos 2u)$$

to get

$$\begin{aligned}\int(9-x^2)^{3/2}dx &= -\frac{81}{4}\left[u-\sin 2u+\frac{1}{2}\int(1+\cos 4u)du\right]\\ &= -\frac{81}{4}\left[\frac{3}{2}u-\sin 2u-\frac{1}{8}\sin 4u\right]\end{aligned}$$

The limits on the definite integral over $x$ were $[0,3]$. With the substitution $x = 3\cos u$, we have $u = \pi/2$ when $x = 0$ and $u = 0$ when $x = 3$. Hence

$$\begin{aligned}\int_0^3(9-x^2)^{3/2}dx &= -\frac{81}{4}\left[\frac{3}{2}u-\sin 2u-\frac{1}{8}\sin 4u\right]\Bigg|_{\pi/2}^{0}\\ &= -\frac{81}{4}\left[\frac{3}{2}(0)-\sin(0)-\frac{1}{8}\sin 0-\frac{3}{2}\frac{\pi}{2}-\sin\pi-\frac{1}{8}\sin 2\pi\right]\\ &= \frac{81}{4}\frac{3}{2}\frac{\pi}{2}=\frac{243\pi}{8}\end{aligned}$$

and consequently

$$V = \frac{2}{3}\int_0^3(9-x^2)^{3/2}dx = \frac{2}{3}\frac{243\pi}{8} = \frac{81\pi}{4} \quad \blacksquare$$

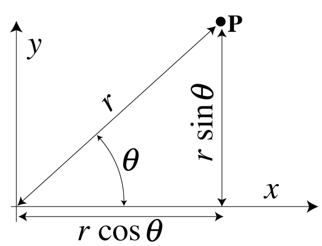
# Lecture 18

# Double Integrals in Polar Coordinates

Recall that in polar coordinates the point $\mathbf{P} = (x, y)$ is represented by the numbers:

$r$, the distance of $\mathbf{P}$ from the origin; and
$\theta$, the angle, measured counter-clockwise from the axis
between the $x$-axis and the line segment from the origin to $\mathbf{P}$.

Figure 18.1: The point $\mathbf{P}$ is represented by the cartesian coordinates $(x, y)$ and the polar coordinates $(r, \theta)$.




To convert from Cartesian coordinates to polar coordinates,

$$x = r\cos\theta$$

$$y = r\sin\theta$$

To convert from polar coordinates to Cartesian coordinates,

$$r^2 = x^2 + y^2$$

$$\tan\theta = y/x$$

**Example 18.1** *Find the polar coordinates of the point (3, 7).*

*Solution.*

$$r^2 = 3^2 + 7^2 = 58 \Rightarrow r = \sqrt{58} \approx 7.62$$

$$\tan\theta = 7/3 \Rightarrow \theta \approx 1.165 \text{ radians} \approx 66.8 \deg \quad \blacksquare$$

**Example 18.2** *Find the Cartesian coordinates corresponding to the point* $\left(\sqrt{3}, -\frac{7\pi}{6}\right)$.

*Solution*

$$x = r\cos\theta = \sqrt{3}\cos\left(-7\pi/6\right) = -\frac{3}{2}$$

$$y = r\sin\theta = \sqrt{3}\sin\left(-7\pi/6\right) = \frac{\sqrt{3}}{2} \quad \blacksquare$$

Our goal here is to define a double integral over a region $R$ that is given in polar rather than rectangular coordinates. We will write this integral as

$$\iint_R f(r,\theta)dA$$

Here $dA$ is the area element as calculated in polar coordinates. In Cartesian coordinates $dA$ represented the area of an infinitesimal rectangle of width $dx$ and height $dy$, hence we were able to write

$$dA = dxdy$$

However, in polar coordinates we can not merely multiple coordinates, because that would give us units of distance $\times$ radians rather than (distance)$^2$ as we require for area. Suppose that our region $R$ is defined as the set

$$R = \{(r,\theta) : a \le r \le b, \alpha \le \theta \le \beta\}$$

We can calculate a formula for the area element $dA$ by breaking the set up into small bits by curves of constant $r$, namely concentric circles about the origin, and curves of constant $\theta$, namely, rays emanating from the origin. Consider one such "bit" extending a length $\Delta r$ radially and $\Delta\theta$ angularly. The length of an arc of a circle of radius $r$ is $r\Delta\theta$, hence the area of the "bit" is approximately

$$\Delta A \approx \Delta\theta\Delta r$$

To see why this is so, observe that since the area of a circle of radius $r$ is $\pi r^2$, then the area of a circular ring from $r$ to $r + \Delta r$ is $\pi r^2 - \pi(r + \delta r)^2$. The fraction of this ring in a wedge of angle $\Delta\theta$ is $\Delta\theta/(2\pi)$, hence

$$\begin{aligned}
\Delta A &= \frac{\Delta\theta}{2\pi}\left[\pi(r+\Delta r)^2 - \pi r^2\right] \\
&= \frac{\Delta\theta}{2}\left[2r\Delta r + (\Delta r)^2\right] \\
&= r\Delta\theta\Delta r\left[1 + \frac{\Delta r}{2r}\right]
\end{aligned}$$

In the limit as $\Delta r$ and $\Delta\theta \to 0$, the term $\Delta r/r \to 0$ much faster than the other terms and hence

$$dA = \lim_{\Delta\theta,\Delta r \to 0} r\Delta\theta\Delta r \left[1 + \frac{\Delta r}{2r}\right] = r dr d\theta$$

Hence we write

$$\iint_R f(x,y)dxdy = \iint_R f(rcos\theta, rsin\theta) r dr d\theta$$

As with cartesian coordinates, we can distinguish between $\theta$-simple and $r$-simple domains in the $xy$-plane.

**Definition 18.1** *A region $R$ in the $xy$ plane is called* **$\theta$-simple** *if it can be expressed in the form*

$$R = \{(r,\theta) : a \le r \le b, g(r) \le \theta \le h(r)\}$$

Double integrals over $\theta$-simple regions have the simple form

$$\iint_R f(x,y)dA = \int_a^b \int_{g(r)}^{h(r)} f(r\cos\theta, r\sin\theta) d\theta \; r \; dr$$

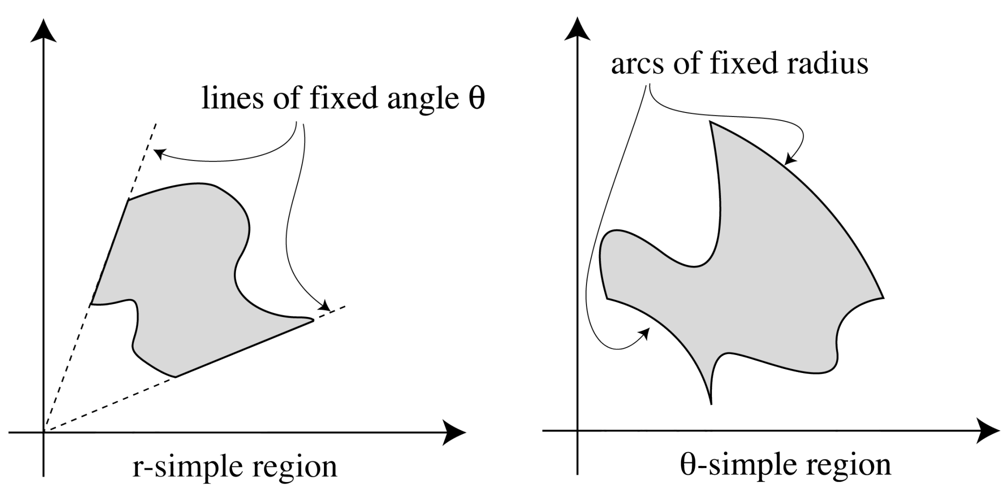
**Definition 18.2** *A region $R$ in the $xy$ plane is called* **r-simple** *if it can be expressed in the form*

$$R = \{(r,\theta) : \alpha \le \theta \le \beta, g(\theta) \le r \le h(\theta)\}$$

Double integrals over $\theta$-simple regions have the simple form

$$\iint_R f(x,y)dA = \int_\alpha^\beta \int_{g(\theta)}^{h(\theta)} f(r\cos\theta, r\sin\theta) r \; dr d\theta$$

Figure 18.2: Examples of $r$-simple (left) and $\theta$-simple (right) regions.



**Example 18.3** *Find the integral*

$$\int_0^{\sqrt{2}} \int_y^{\sqrt{4-y^2}} xdxdy$$

*using polar coordinates.*

*Solution.* The integrand is $f(x, y) = x$, hence

$$f(r\cos\theta, r\sin\theta) = r\cos\theta$$

The area element is

$$dA = rdrd\theta$$

Hence (except for the limits, which we still have to determine), the integral is

$$\int_?^? \int_?^? r^2 \cos\theta drd\theta$$

To determine the limits of integration, we need to figure out what the domain is. This is usually facilitated by sketching the domain. The outer integral has limits

$$0 \leq y \leq \sqrt{2}$$

while the inner integral has limits

$$y \leq x \leq \sqrt{4-y^2}$$

Since the limits on $x$ are functions of $y$ the region is $x$ simple. The equation $x = \sqrt{4-y^2}$ lies on the same curve $x^2 + y^2 = 4$, i.e., the upper limit on the $x$-integral is part of the circle of radius 2 centered at the origin. Since $x$ is less than $\sqrt{4-y^2}$ then the integral is over part of the inside of this circle, and lies to the left of the arc of the circle (smaller values of $x$ lie to the left of larger values of $x$), as illustrated in figure 18.3a.

The region is also bounded on the left by the line $x = y$, since we have $x > y$; equivalently, $y < x$ hence the region is beneath as well to the right of the line, as illustrated in figure 18.3b. Next, we can fill in the top and bottom of the region as being between the lines $y = 0$ and and $\sqrt{2}$ (figure 18.3c). We see that the domain of integration is a circular wedge, of the circle of radius 2, centered in the origin, that lies between the $x$-axis and the line $y = x$. In polar coordinates we end up with

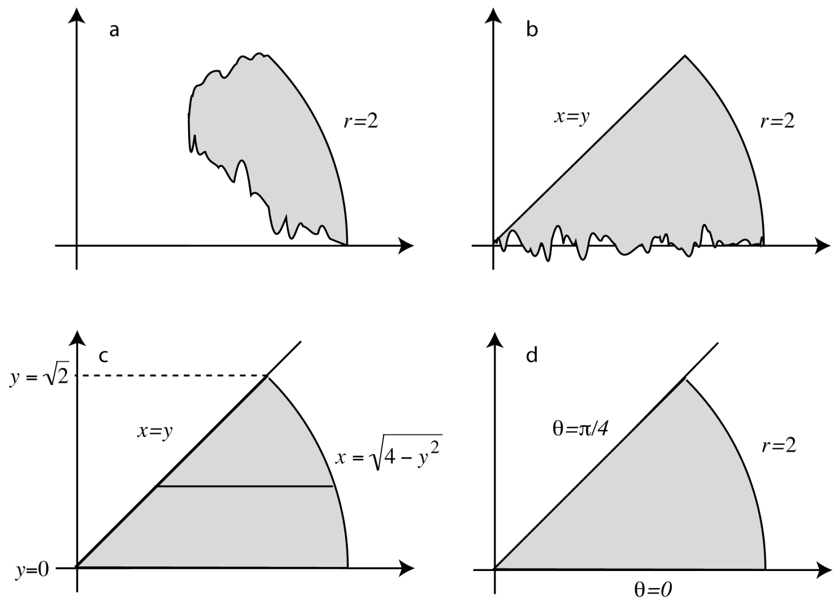
$$0 \leq r \leq \sqrt{2}$$

$$0 \leq \theta \leq \pi/4$$

The region is both $\theta$-simple and $r$-simple, and hence the order of integration does not matter, and we have

$$\begin{aligned}\int_0^{\sqrt{2}} \int_y^{\sqrt{4-y^2}} xdxdy &= \int_0^{\sqrt{2}} \int_0^{\pi/4} r^2 \cos\theta d\theta dr \\ &= \int_0^{\pi/4} \int_0^{\sqrt{2}} r^2 \cos\theta drd\theta\end{aligned}$$

Figure 18.3: The domain of integration of the integral in example 18.3.




We emphasize the fact that when the order of integration is reversed, the order of the differentials $d\theta dr$ reverses to $drd\theta$ **and** the order two integrals is reversed.

Finally, to solve the integral we arbitrarily choose one of these forms, say the second one, to give

$$\begin{aligned}
\int_0^{\sqrt{2}} \int_y^{\sqrt{4-y^2}} xdxdy &= \int_0^{\pi/4} \cos\theta \int_0^2 r^2 drd\theta \\
&= \int_0^{\pi/4} \cos\theta \left( \frac{1}{3} r^3 \Big|_0^2 \right) d\theta \\
&= (8/3) \int_0^{\pi/4} \cos\theta d\theta \\
&= (8/3) \left( \sin\theta|_0^{\pi/4} \right) = (8/3)(\sqrt{2}/2) = 4\sqrt{2}/3 \approx 1.88562 \ \blacksquare
\end{aligned}$$

**Example 18.4** *Find the area of a circle of radius a using double integrals in polar coordinates.*

*Solution.* Consider a circle of radius $a$ whose center is at the origin. Then the interior of the circle is the set

$$C = \{(r,\theta) : 0 \le r \le a, 0 \le \theta \le 2\pi\}$$

Since the area of any region $C$ is $\iint_C f dA$, the area of the circle is

$$\begin{aligned}
\iint_C r dr d\theta &= \int_0^{2\pi}\int_0^a r dr d\theta \\
&= \int_0^{2\pi} \frac{1}{2}r^2\bigg|_0^a d\theta \\
&= \frac{a^2}{2}\int_0^{2\pi} d\theta \\
&= \frac{a^2}{2}(2\pi) \\
&= \pi a^2 \ \blacksquare
\end{aligned}$$

**Example 18.5** *Find the volume $V$ of the solid under the function*

$$f(x, y) = e^{-x^2-y^2}$$

*and over the region*

$$S = \{(r, \theta) : 5 \le r \le 7, 0 \le \theta \le \pi/2\}$$

*Solution.* This region is both r-simple and theta-simple, so it is not difficult to "set-up" the integral:

$$V = \iint\limits_R f(r, \theta) r dr d\theta = \int_0^{\pi/2}\int_5^7 f(r, \theta) r dr d\theta$$

To convert the function $f(x, y)$ to a function in polar coordinates we observe that since

$$x^2 + y^2 = r^2$$

then

$$f(x, y) = e^{-(x^2+y^2)} = e^{-r^2}$$

Thus the integral becomes

$$V = \int_0^{\pi/2}\int_5^7 e^{-r^2} r dr d\theta$$

In the inner integral, make the substitution

$$u = -r^2$$

so that

$$du = -2dr$$

With this change of variables, when $r = 5$, $u = -25$; when $r = 7$, $u = -49$. Then the inner integral is then

$$\begin{aligned}\int_5^7 e^{-r^2} r dr &= -\frac{1}{2}\int_{-25}^{-49} e^u du \\ &= -\frac{1}{2}\, e^u|_{-25}^{-49} \\ &= -\frac{1}{2}\left(e^{-49} - e^{-25}\right) \\ &= \frac{1}{2}\left(e^{-25} - e^{-49}\right) \\ &\approx 6.94 \times 10^{-12}\end{aligned}$$

Finally, we can calculate the double integral,

$$\begin{aligned} V &= \int_0^{\pi/2}\int_5^7 e^{-r^2} r dr d\theta \\ &\approx 6.94 \times 10^{-12}\int_0^{\pi/2} d\theta \\ &\approx (6.94 \times 10^{-12})\left(\frac{\pi}{2}\right) \\ &\approx 1.1 \times 10^{-11} \quad \blacksquare\end{aligned}$$

**Example 18.6** *Find the integral*

$$\iint_s \sqrt{4 - x^2 - y^2} dA$$

*where $S$ is the first-quadrant sector of the circle of radius 2 centered at the orgin between $y = 0$ and $y = x$.*

*Solution.* The domain is the same 45 degree sector of a circle as in example 18.3. Since

$$\begin{aligned} f(r\cos\theta, r\sin\theta) &= \sqrt{4 - (r\cos\theta)^2 - (r\sin\theta)^2} \\ &= \sqrt{4 - r^2}\end{aligned}$$

the integral is

$$\begin{aligned}\iint_S \sqrt{4 - x^2 - y^2}\, dA &= \int_0^2\int_0^{\pi/2} \sqrt{4 - r^2}\, r d\theta dr \\ &= \int_0^2 r\sqrt{4 - r^2}\left(\int_0^{\pi/2} d\theta\right) dr \\ &= \frac{\pi}{2}\int_0^2 r\sqrt{4 - r^2}\, dr\end{aligned}$$

We can solve this final integral by making the substitution $u = 4 - r^2$, hence $du = -2rdr$. When $r = 0$, $u = 4$ and when $r = 2$, $u = 0$. Therefore

$$\begin{aligned}
\iint_S \sqrt{4 - x^2 - y^2}\, dA &= -\frac{\pi}{2} \int_4^0 u^{1/2}(du/2) \\
&= -\frac{\pi}{4} \int_4^0 u^{1/2} du \\
&= -\frac{\pi}{4}\frac{2}{3} u^{3/2} \Big|_4^0 \\
&= \frac{\pi}{6} \left(4^{3/2}\right) \\
&= \frac{4\pi}{3} \ \blacksquare
\end{aligned}$$

**Example 18.7** *Find the value of*

$$\int_{-\infty}^{\infty} e^{-x^2} dx$$

*Solution.* There is no possible substitution that will allow us to find a closed form solution to the indefinite integral

$$\int e^{-x^2} dx$$

However, by using polar coordinates we can find the definite integral $\int_{-\infty}^{\infty} e^{-x^2} dx$, as follows. First, we define the quantity

$$I = \int_{-\infty}^{\infty} e^{-x^2} dx$$

which is the definite integral we are trying to find. Then

$$\begin{aligned}
I^2 &= \left(\int_{-\infty}^{\infty} e^{-x^2} dx\right)\left(\int_{-\infty}^{\infty} e^{-x^2} dx\right) \\
&= \int_{-\infty}^{\infty} e^{-x^2} dx \int_{-\infty}^{\infty} e^{-y^2} dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-x^2} e^{-y^2} dxdy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)} dxdy
\end{aligned}$$

This is an integral over the entire $xy$ plane; in polar coordinates this is the set

$$S = \{(r, \theta) : 0 \le r \le \infty, 0 \le \theta \le 2\pi\}$$

Transforming the integral to polar coordinates and using the fact that $r^2 = x^2 + y^2$ gives

$$\begin{aligned} I^2 &= \int_0^\infty \int_0^{2\pi} e^{-r^2} r d\theta dr \\ &= \int_0^\infty e^{-r^2} r \int_0^{2\pi} d\theta dr \\ &= \int_0^\infty e^{-r^2} r(2\pi) dr \\ &= 2\pi \int_0^\infty e^{-r^2} r dr \end{aligned}$$

At this point we can make the substitution $u = -r^2$. Then $du = -2rdr$, when $r = 0$, then $u = 0$ and when $r = \infty$, $u = -\infty$

$$\begin{aligned} I^2 &= 2\pi \int_0^\infty e^{-r^2} r dr \\ &= -2\pi \int_0^{-\infty} e^u du \\ &= -2\pi e^u \big|_0^{-\infty} \\ &= -2\pi(e^{-\infty} - e^0) \\ &= -2\pi(0 - 1) = 2\pi \end{aligned}$$

Therefore

$$\int_{-\infty}^\infty e^{-x^2} dx = I = \sqrt{I^2} = \sqrt{2\pi} \ \blacksquare$$

# Lecture 19

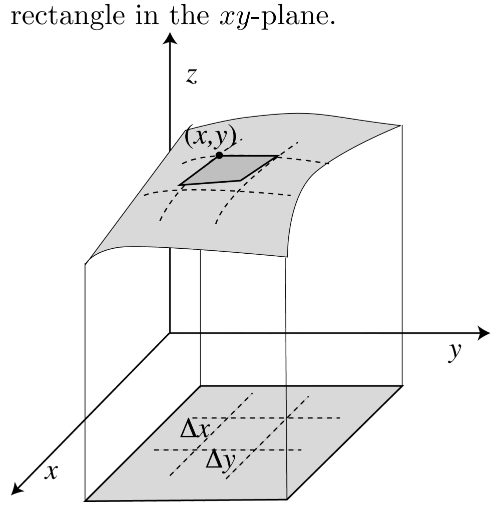
# Surface Area with Double Integrals

To find the area of a surface described by a function $f(x, y) : \mathbb{R}^2 \mapsto \mathbb{R}$ using double integrals, we first break the domain (in the $xy$-plane beneath the surface) into small rectangles of size

$$\Delta x \times \Delta y$$

Above each rectangle in the $xy$-plane there will be some infinitesmal "bit" of the surface that is approximately (but not quite) flat; we can approximate this infinitesimal "bit" of the surface by an infinitesimal tangent plane. Pick an arbitrary point within each surface "bit," say the corner, and label this point $(x, y)$. Then there is a "bit" of the tangent plane that lies over the $\Delta x \times \Delta y$ rectangle in the $xy$-plane at the point $(x, y)$. This "bit" of the tangent plane is a parallelogram. The back edge of the parallelogram is formed by the intersection of a plane parallel to

Figure 19.1: A surface can be broken into infinitesimal "bits," each of which lies above an infinitesimal rectangle in the $xy$-plane.

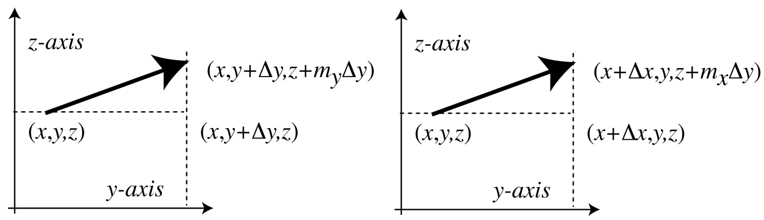


the $yz$ plane with the tangent plane; the slope of this line is therefore

$$m = f_y(x, y)$$

The vector $\mathbf{v}$ from $(x, y, z)$ to $(x, y + \Delta y, z + m\Delta y)$ is

Figure 19.2: Calculation of vectors that edge the infinitesimal tangent plane.




Plane through (x,y,z) that is parallel to y-z plane

Plane through (x,y,z) that is parallel to x-z plane

$$\mathbf{v} = (0, \Delta y, m\Delta y) = (0, \Delta y, f_y(x, y)\Delta y)$$

Similarly, the left edge of the parallelogram is formed by the intersection of a plane parallel to the $xz$-plane with the tangent plane; the slope of this line is therefore

$$m_x = f_x(x, y)$$

By a similar construction, the vector $\mathbf{u}$ from $(x, y, z)$ to $(x + \Delta x, y, z + m_x\Delta x)$ is

$$\mathbf{u} = (\Delta x, 0, m_x\Delta x) = (\Delta x, 0, f_x(x, y)\Delta x)$$

The area of the parallelogram formed by $\mathbf{u}$ and $\mathbf{v}$ is given by their cross product:

$$\begin{aligned}
\Delta A &= \|\mathbf{u} \times \mathbf{v}\| \\
&= \left\| \begin{pmatrix} 0 & -f_x(x,y)\Delta x & 0 \\ f_x(x,y) & 0 & -\Delta x \\ 0 & \Delta x & 0 \end{pmatrix} \begin{pmatrix} 0 \\ \Delta y \\ f_y(x,y)\Delta y \end{pmatrix} \right\| \\
&= \left\| \begin{pmatrix} -f_x(x,y)\Delta x\Delta y \\ -f_y(x,y)\Delta x\Delta y \\ \Delta x\Delta y \end{pmatrix} \right\| \\
&= \Delta x\Delta y\sqrt{1 + f_x^2(x,y) + f_y^2(x,y)}
\end{aligned}$$

Taking the limit as the size of the infinitesmal parallelograms go to zero, we have

$$dA = dA = \sqrt{1 + f_x^2(x,y) + f_y^2(x,y)}\, dxdy$$

**Theorem 19.1** . *Suppose that $z = f(x, y) : D \cup \mathbb{R}^2 \mapsto \mathbb{R}$ Then the surface area of $f$ over $D$ is*

$$A = \iint_D \sqrt{1 + f_x^2(x,y) + f_y^2(x,y)}\, dxdy$$

**Example 19.1** *Find the surface area of the region of the plane* $3x - 2y + 6z = 12$ *that is bounded by the planes* $x = 0$*,* $y = 0$ *and* $3x + 2y = 12$*.*

*Solution* The domain is the region in the $xy$ plane bounded by the $x$-axis, the $y$-axis, and the line

$$3x + 2y = 12$$

Solving the line for $y$,

$$y = 6 - 1.5x$$

This line intersects the $x$-axis at $x = 12/3 = 4$, and the $y$-axis at $y = 12/2 = 6$. Therefore we can write the domain of the integral as

$$\{(x, y) : 0 \le x \le 4, 0 \le y \le 6 - 1.5x\}$$

And the surface area as

$$A = \int_0^4 \int_0^{6-1.5x} \sqrt{1 + f_x^2(x, y) + f_y^2(x, y)}dydx$$

where $f(x, y)$ is the solution of $3x - 2y + 6z = 12$ for $z = f(x, y)$, namely

$$f(x, y) = 2 - \frac{x}{2} + \frac{y}{3}$$

Differentiating, $f_x = -1/2$ and $f_y = 1/3$, thus

$$\begin{aligned}
A &= \int_0^4 \int_0^{6-1.5x} \sqrt{1 + f_x^2(x, y) + f_y^2(x, y)}dydx \\
&= \int_0^4 \int_0^{6-1.5x} \sqrt{1 + (-1/2)^2 + (1/3)^2}dydx \\
&= \sqrt{49/36} \int_0^4 \int_0^{6-1.5x} dydx \\
&= \frac{7}{6} \int_0^4 \int_0^{6-1.5x} dydx \\
&= \frac{7}{6} \int_0^4 (6 - 1.5x)dx \\
&= \frac{7}{6}(6x - 0.75x^2)\Big|_0^4 \\
&= \frac{7}{6}(24 - 12) = 14 \;\blacksquare
\end{aligned}$$

**Example 19.2** *Find the surface area of the part of the surface*

$$z = \sqrt{9 - y^2}$$

*that lies above the quarter of the circle*

$$x^2 + y^2 = 9$$

*in the first quadrant.*

*Solution.* The domain is the set

$$\{(x, y) : 0 \leq y \leq 3, 0 \leq x \leq \sqrt{9 - y^2}\}$$

Therefore the area is

$$A = \int_0^3 \int_0^{\sqrt{9-y^2}} \sqrt{1 + f_x^2 + f_y^2} dxdy$$

Differentiating $f(x, y) = (9 - y^2)^{1/2}$ gives $f_x = 0$ and

$$f_y = (1/2)(9 - y^2)^{-1/2}(-2y) = -y(9 - y^2)^{-1/2}$$

and therefore

$$\begin{aligned} A &= \int_0^3 \int_0^{\sqrt{9-y^2}} \sqrt{1 + f_x^2 + f_y^2} dxdy \\ &= \int_0^3 \int_0^{\sqrt{9-y^2}} \sqrt{1 + [-y(9 - y^2)^{-1/2}]^2} dxdy \\ &= \int_0^3 \int_0^{\sqrt{9-y^2}} \sqrt{1 + y^2(9 - y^2)^{-1}} dxdy \\ &= \int_0^3 \int_0^{\sqrt{9-y^2}} \sqrt{1 + \frac{y^2}{9 - y^2}} dxdy \\ &= \int_0^3 \int_0^{\sqrt{9-y^2}} \sqrt{\frac{9}{9 - y^2}} dxdy \\ &= 3\int_0^3 \int_0^{\sqrt{9-y^2}} \frac{1}{\sqrt{9 - y^2}} dxdy \end{aligned}$$

Since the integrand is only a function of $y$, we can move it from the inner integral to the outer integral:

$$\begin{aligned} A &= 3\int_0^3 \frac{1}{\sqrt{9 - y^2}} \int_0^{\sqrt{9-y^2}} dxdy \\ &= 3\int_0^3 \frac{1}{\sqrt{9 - y^2}} \sqrt{9 - y^2} dy \\ &= 3\int_0^3 dy = 9 \quad \blacksquare \end{aligned}$$

**Example 19.3** *Find a formula for the surface area of a sphere of radius a.*

*Solution.* The equation of a sphere of radius a centered at the origin is

$$x^2 + y^2 + z^2 = a^2$$

Solving for $z$

$$z = \pm\sqrt{a^2 - x^2 - y^2}$$

The total area is the sum of the area of the top half and the area of the bottom half of the plane; the equation of the sphere in the top half of the plane is

$$f(x, y) = \sqrt{a^2 - x^2 - y^2}$$

Differentiating

$$f_x = \frac{-x}{\sqrt{a^2 - x^2 - y^2}} \quad \text{and} \quad f_y = \frac{-y}{\sqrt{a^2 - x^2 - y^2}}$$

The area of the top half of the sphere is then

$$\begin{aligned}
A &= \iint_C \sqrt{1 + f_x^2 + f_y^2}\,dxdy \\
&= \iint_c \sqrt{1 + \frac{x^2}{a^2 - x^2 - y^2} + \frac{y^2}{a^2 - x^2 - y^2}}\,dxdy \\
&= \iint_c \sqrt{\frac{a^2 - x^2 - y^2}{a^2 - x^2 - y^2} + \frac{x^2}{a^2 - x^2 - y^2} + \frac{y^2}{a^2 - x^2 - y^2}}\,dxdy \\
&= \iint_c \sqrt{\frac{a^2}{a^2 - x^2 - y^2}}\,dxdy \\
&= a \iint_c \frac{1}{\sqrt{a^2 - x^2 - y^2}}\,dxdy
\end{aligned}$$

where $C$ is the circle of radius $a$ in the $xy$-plane. In polar coordinates, then

$$A = a \int_0^{2\pi} \int_0^a \frac{1}{\sqrt{a^2 - r^2}}\,rdrd\theta$$

Letting $u = a^2 - r^2$ we have $du = -2rdr$; when $r = 0$, $u = a^2$ and when $r = a$, $u = 0$.Thus the area of the top half of the sphere is

$$\begin{aligned}
A &= a \int_0^{2\pi} \int_{a^2}^0 \frac{1}{\sqrt{u}}(-du/2)d\theta \\
&= -\frac{a}{2} \int_0^{2\pi} \int_{a^2}^0 u^{-1/2}dud\theta \\
&= -\frac{a}{2} \int_0^{2\pi} \left.\frac{u^{1/2}}{1/2}\right|_{a^2}^0 d\theta \\
&= -a \int_0^{2\pi} [0 - (a^2)^{1/2}]d\theta \\
&= a^2 \int_0^{2\pi} d\theta = 2\pi a^2
\end{aligned}$$

Therefore the area of the sphere is $2A = 4\pi a^2$ ■

**Example 19.4** *Find the area of the surface $z = x^2/4 + 4$ that is cut off by the planes $x = 0$, $x = 1$, $y = 0$, $y = 2$.*

*Solution.* Differentiating $f_x = x/2, f_y = 0$; hence the area is

$$\begin{aligned}
A &= \iint_R \sqrt{1 + f_x^2 + f_y^2}\, dA \\
&= \int_0^1 \int_0^2 \sqrt{1 + \frac{x^2}{4}}\, dydx \\
&= \int_0^2 \sqrt{\frac{4 + x^2}{4}} \int_0^2 \, dydx \\
&= \frac{1}{2}(2) \int_0^2 \sqrt{4 + x^2}\, dx \\
&= \int_0^2 \sqrt{4 + x^2}\, dx
\end{aligned}$$

From integral 44 in the back of the book,

$$\int \sqrt{x^2 + a^2}\, dx = \frac{x}{2}\sqrt{x^2 + a^2} + \frac{a^2}{2} \ln \left| x + \sqrt{x^2 + a^2} \right|$$

Setting $a = 2$,

$$\int \sqrt{x^2 + 4}\, dx = \frac{x}{2}\sqrt{x^2 + 4} + 2 \ln |x + \sqrt{x^2 + 4}|$$

Therefore

$$\begin{aligned}
A &= \int_0^2 \sqrt{4 + x^2}\, dx \\
&= \left[ \frac{x}{2}\sqrt{x^2 + 4} + 2 \ln |x + \sqrt{x^2 + 4}| \right] \Big|_0^2 \\
&= \left[ \frac{2}{2}\sqrt{4 + 4} + 2 \ln |2 + \sqrt{4 + 4}| \right] - \left[ \frac{0}{2}\sqrt{0 + 4} + 2 \ln |0 + \sqrt{0^2 + 4}| \right] \\
&= \sqrt{8} + 2 \ln |2 + \sqrt{8}| - 2 \ln 2 \\
&= 2\sqrt{2} + 2 \ln(1 + \sqrt{2}) \quad \blacksquare
\end{aligned}$$

# Lecture 20

# Triple Integrals

## Triple Integrals in Cartesian Coordinates

We can easily extend our definition of a double integral to a triple integral. Suppose that $f(x, y, z) : V \cup \mathbb{R}^3 \mapsto \mathbb{R}$. Then we extend our Riemann sum so that it covers the volume $V$ with small boxes $V_j$ of dimension

$$\Delta x_j \times \Delta y_j \times \Delta z_j$$

The volume $\Delta V_j$ of the $i$th box is then

$$\Delta V_j = \Delta x_j \Delta y_j \Delta z_j$$

Then the Riemann Sum representing $\iiint_V f(x, y, z)dV$ is

$$V \approx \sum_{j=1}^{n} f(x_j, y_j, z_j)\Delta V_i$$

and the **triple integral of $f$ over V** is

$$\iiint_V f(x, y, z)dV = \lim_{n\to\infty, \Delta V_j \to 0} \sum_{j=1}^{n} f(x_j, y_j, z_j)\Delta V_i$$

We can also make general definitions of $x$-simple and $y$-simple sets. These definitions are exactly the same as they were previously but with an added dimension in the domain.

**Example 20.1** *Find*

$$\int_0^2 \int_{-2}^3 \int_0^{2x-y} dzdydx$$

*Solution* We integrate first over $z$, then over $y$, then over $x$, because that is the order of the "d"s inside the integral.

$$\begin{aligned}
\int_0^2 \int_{-2}^3 \int_0^{2x-y} dz dy dx &= \int_0^2 \int_{-2}^3 \left( \int_0^{2x-y} dz \right) dy dx \\
&= \int_0^2 \int_{-2}^3 \left( z|_0^{2x-y} \right) dy dx \\
&= \int_0^2 \int_{-2}^3 (2x-y) dy dx \\
&= \int_0^2 \left( 2xy - \frac{1}{2}y^2 \right)\Big|_{-2}^3 dx \\
&= \int_0^2 [6x - 9/2 + 4x + 2] dx \\
&= \int_0^2 (10x - 5/2) dx \\
&= \left[ 5x^2 - (5/2)x \right]_0^2 \\
&= 5(4) - (5/2)(2) = 15 \quad \blacksquare
\end{aligned}$$

**Example 20.2** *Find*

$$\int_0^{\pi/2} \int_0^z \int_0^y \sin(x+y+z) dx dy dz$$

*Solution* We integrate first over $x$, then over $y$, then over $z$. Since $\int \sin u du = -\cos u$,

$$\begin{aligned}
\int_0^{\pi/2} \int_0^z \int_0^y \sin(x+y+z) dx dy dz &= \int_0^{\pi/2} \left[ \int_0^z \left( \int_0^y \sin(x+y+z) dx \right) dy \right] dz \\
&= \int_0^{\pi/2} \left[ \int_0^z \left( -\cos(x+y+z)|_0^y \right) dy \right] dz \\
&= \int_0^{\pi/2} \left[ \int_0^z \left( -\cos(2y+z) + \cos(y+z) \right) dy \right] dz
\end{aligned}$$

Next we integrate over $y$, using the fact that $\int \cos u du = \sin u$

$$\begin{aligned}
\int_0^{\pi/2} \int_0^z \int_0^y \sin(x+y+z) dx dy dz &= \int_0^{\pi/2} \left[ \int_0^z \left( -\cos(2y+z) + \cos(y+z) \right) dy \right] dz \\
&= \int_0^{\pi/2} \left[ -\frac{1}{2} \sin(2y+z) + \sin(y+z) \right]\Big|_0^z dz \\
&= \int_0^{\pi/2} \left[ \left( -\frac{1}{2} \sin 3z + \sin 2z \right) - \left( -\frac{1}{2} \sin z + \sin z \right) \right] dz \\
&= \int_0^{\pi/2} \left[ -\frac{1}{2} \sin 3z + \sin 2z - \frac{1}{2} \sin z \right] dz
\end{aligned}$$

Finally, we an integrate over $z$.

$$\int_0^{\pi/2}\int_0^z\int_0^y \sin(x+y+z)dxdydz = \int_0^{\pi/2}\left[-\frac{1}{2}\sin 3z+\sin 2z-\frac{1}{2}\sin z\right]dz$$
$$=\left[\frac{1}{6}\cos 3z-\frac{1}{2}\cos 2z+\frac{1}{2}\cos z\right]_0^{\pi/2}$$
$$=\left[\frac{1}{6}\cos\frac{3\pi}{2}-\frac{1}{2}\cos\pi+\frac{1}{2}\cos\frac{\pi}{2}\right]$$
$$-\left[\frac{1}{6}\cos 0-\frac{1}{2}\cos 0+\frac{1}{2}\cos 0\right]$$
$$=\frac{1}{6}(0)-\frac{1}{2}(-1)+\frac{1}{2}(0)-\frac{1}{6}=\frac{1}{2}\ \blacksquare$$

**Definition 20.1** *If $S\cup\mathbb{R}^3$ is any solid object and $f(x,y,z): S\mapsto\mathbb{R}$ then the* **volume integral of $f$ over $S$**

$$I=\iiint_s f(x,y,z)dxdydz$$

**Theorem 20.1** *The volume of $S$ is the volume integral of $f(x,y,z)=1$ over $S$*

$$V=\iiint_s dxdydz$$

The volume of S is the volume integral of the function $f(x,y,z)=1$.

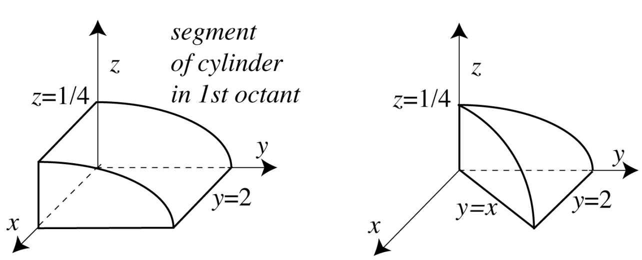
**Example 20.3** . *Find the volume of the solid in the first octant bounded by the surfaces*

$$y^2+64z^2=4$$

*and*

$$y=x$$

Figure 20.1: Left: the portion of the elliptic cylinder $y^2+64z^2=4$ in the first octant . Right: the portion bounded by $y=x$ in the first octant.




*Solution.* We can solve the equation of the surface for $z$,

$$z=\pm\frac{1}{8}\sqrt{4-y^2}$$

Since we are interested in the first octant, we chose the positive solution.

$$z = \pm\frac{1}{8}\sqrt{4-y^2}$$

The ellipse where the cylinder crosses the $yz$-plane intersects the $y$-axis at $y = 2$ and the $z$-axis at $z = 1/4$. The base of the object in the $xy$-plane is a triangle formed by the lines $y = x$, $y = 2$, and the $y$-axis.Therefore the domain of integration is the set

$$V = \left\{(x, y, z) : 0 \le z \le \frac{\sqrt{4-y^2}}{8}, x \le y \le 2, 0 \le x \le 2\right\}$$

The volume of this set is $\iiint_V (1) \times dV$, namely

$$\begin{aligned} V &= \int_0^2 \int_x^2 \int_0^{(4-y^2)^{1/2}/8} dz dy dx \\ &= \frac{1}{8}\int_0^2 \int_x^2 \sqrt{4-y^2} dy dx \end{aligned}$$

From integral formula (54) in the inside back cover

$$\int \sqrt{a^2-y^2}\, dy = \frac{y}{2}\sqrt{a^2-y^2} + \frac{a^2}{2}\sin^{-1}\frac{y}{a}$$

Setting $a = 2$,

$$\int \sqrt{4-y^2}\, dy = \frac{y}{2}\sqrt{4-y^2} + 2\sin^{-1}\frac{y}{2}$$

Therefore

$$\begin{aligned} \int_x^2 \sqrt{4-y^2} dy &= \left[\frac{y}{2}\sqrt{4-y^2} + 2\sin^{-1}\frac{y}{2}\right]_x^2 \\ &= \left(\frac{2}{2}\sqrt{4-2^2} + 2\sin^{-1}\frac{2}{2}\right) - \left(\frac{x}{2}\sqrt{4-x^2} + 2\sin^{-1}\frac{x}{2}\right) \\ &= \pi - \frac{x}{2}\sqrt{4-x^2} - 2\sin^{-1}\frac{x}{2} \end{aligned}$$

and hence the volume is

$$\begin{aligned} V &= \frac{1}{8}\int_0^2 \int_x^2 \sqrt{4-y^2}\, dy dx = \frac{1}{8}\int_0^2 \left(\pi - \frac{x}{2}\sqrt{4-x^2} - 2\sin^{-1}\frac{x}{2}\right) dx \\ &= \frac{1}{8}\left[\pi \int_0^2 dx - \frac{1}{2}\int_0^2 x\sqrt{4-x^2} dx - 2\int_0^2 \sin^{-1}\frac{x}{2} dx\right] \\ &= \frac{\pi}{4} - \frac{1}{16}\int_0^2 x\sqrt{4-x^2} dx - \frac{1}{4}\int_0^2 \sin^{-1}\frac{x}{2} dx \end{aligned}$$

The first integral we can solve with the substitution $u = 4 - x^2$, which implies that $du = -2dx$. Furthermore, when $x = 0$, $u = 4$, and when $x = 2$, $u = 0$. Thus

$$\int_0^2 x\sqrt{4-x^2} dx = -\frac{1}{2}\int_4^0 u^{1/2} du = \frac{1}{2}\int_0^4 u^{1/2} du = \left.\frac{u^{3/2}}{3}\right|_0^4 = \frac{(4)^{3/2}}{3} = \frac{8}{3}$$

so that

$$V = \frac{\pi}{4} - \frac{1}{16}\frac{8}{3} - \frac{1}{4}\int_0^2 \sin^{-1}\frac{x}{2}dx = \frac{\pi}{4} - \frac{1}{6} - \frac{1}{4}\int_0^2 \sin^{-1}\frac{x}{2}dx$$

Using formula (69) in the textbook's integral table,

$$\int \sin^{-1} u du = u \sin^{-1} u + \sqrt{1-u^2}$$

so that

$$\begin{aligned} \int_0^2 \sin^{-1}(x/2)dx &= 2\left[(x/2)\sin^{-1}(x/2) + \sqrt{1-x^2/4}\right]_0^2 \\ &= 2\left(\sin^{-1} 1 + \sqrt{1-1} - 0\sin^{-1} 0 - \sqrt{1-0}\right) \\ &= \pi - 2 \end{aligned}$$

so that the desired volume is

$$V = \frac{\pi}{4} - \frac{1}{6} - \frac{1}{4}(\pi - 2) = \frac{\pi}{4} - \frac{1}{6} - \frac{\pi}{4} + \frac{1}{2} = \frac{1}{3} \quad \blacksquare$$

## Triple Integrals in Cylindrical Coordinates

Recall the conversion between Cartesian and cylindrical coordinates: to find $(x, y, z)$ given $(r, \theta, z)$,

$$x = r\cos\theta, \quad y = r\sin\theta$$

with $z$ unchanged; to find $(r, \theta, z)$ given $(x, y, z)$

$$r^2 = x^2 + y^2, \quad , \tan\theta = y/x$$

again with $z$ unchanged. We observe that cylindrical coordinates are identical to polar coordinates in the $xy$-plane, and identical to Cartesian coordinates in the $z$-direction. We can fill up volume $V$ with micro-volumes that have bases given by the volume element in polar coordinates

$$dA = rdrd\theta$$

and height $dz$, hence

$$dV = dA \times dz = rdrd\theta dz$$

and

$$\iiint\limits_V f(x, y, z)dV = \iiint\limits_V f(r\cos\theta, r\sin\theta, z)\ rdr\ d\theta\ dz$$

**Example 20.4** *Use cylindrical coordinates to find the volume of solid bounded by the paraboloid*

$$z = 9 - x^2 - y^2$$

*and the xy-plane.*

*Solution.* The paraboloid intersects the $xy$-plane in a circle of radius 3 centered about the origin.In cylindrical coordinates,

$$z = 9 - x^2 - y^2 = 9 - r^2$$

and therefore

$$\begin{aligned} V &= \int_0^{2\pi}\int_0^3\int_0^{9-r^2} dzrdrd\theta \\ &= \int_0^{2\pi}\int_0^3 (9 - r^2)rdrd\theta \\ &= \int_0^{2\pi}\int_0^3 (9r - r^3)drd\theta \\ &= \int_0^{2\pi}\left(\frac{9}{2}r^2 - \frac{1}{4}r^4\right)_0^3 d\theta \\ &= \left(\frac{81}{2} - \frac{81}{4}\right)(2\pi) \\ &= \frac{81\pi}{4} \end{aligned}$$

## Triple Integrals in Spherical Coordinates

Recall that each point in space is represented by a triple $(\rho, \theta, \phi)$ where

$$x = \rho \sin\phi \cos\theta$$

$$y = \rho \sin\phi \sin\theta$$

$$z = \rho \cos\phi$$

and

$$\rho^2 = x^2 + y^2 + z^2$$

$$\tan\theta = y/z$$

$$\cos\phi = z/\sqrt{x^2 + y^2 + z^2}$$

The volume element in spherical coordinates is

$$dV = \rho^2 \sin\phi d\rho d\theta d\phi$$

and hence the triple integral is

$$\iiint_V f(x, y, z) dV = \iiint_s f(\rho \sin\phi cos\theta, \rho \sin\phi \sin\theta, \rho \cos\phi)\rho^2 \sin\phi \; d\rho \; d\theta \; d\phi$$

The total volume of V is thus

$$V = \iiint_s \rho^2 \sin\phi d\rho d\theta d\phi$$

**Example 20.5** *Find the volume of ball of radius a centered at the origin.*

*Solution.*

$$\begin{aligned} V &= \iiint_s \rho^2 \sin\phi d\rho d\theta d\phi \\ &= \int_0^\pi \int_0^{2\pi} \int_0^a \rho^2 \sin\phi d\rho d\theta d\phi \\ &= \frac{1}{3} a^3 \int_0^\pi \int_0^{2\pi} d\theta \sin\phi d\phi \\ &= \frac{1}{3} a^3 (2\pi) \int_0^\pi \sin\phi d\phi \\ &= \frac{2\pi}{3} a^3 (-\cos\phi)|_0^\pi \\ &= \frac{2\pi}{3} a^3 (-(-1) - -(1)) = \frac{4\pi}{3} a^3 \quad \blacksquare \end{aligned}$$

## Triple Integrals in Any Coordinate Frame

Suppose we have a coordinate frame $(u, v, w)$ where

$$x = X(u, v, w)$$

$$y = Y(u, v, w)$$

$$z = Z(u, v, w)$$

Then the volume element is

$$dV = J(u, v, w)dudvdw$$

where

$$J(u, v, w) = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} & \dfrac{\partial x}{\partial w} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} & \dfrac{\partial y}{\partial w} \\ \dfrac{\partial z}{\partial u} & \dfrac{\partial a}{\partial v} & \dfrac{\partial a}{\partial w} \end{vmatrix}$$

is called the **Jacobian of the transformation**.

# Lecture 21

# Vector Fields

**Definition 21.1** *A* **vector field on** $\mathbb{R}^3$ *is a function* $\boldsymbol{F}(x, y, z) : D \subset \mathbb{R}^3 \mapsto \mathbb{R}^3$ *that assigns a vector* $\mathbf{v} = (v_x, v_y, v_z)$ *to every point* $(x, y, z)$ *in* $D$*.*

Of course, it is also possible to define a vector field on $\mathbb{R}^2$, as a function that assigns a vector $\mathbf{v} = (v_x, v_y)$ to each point $(x, y)$ in some domain $D \subset \mathbb{R}^2$. We will limit ourselves to a discussion 2D fields at first because they are easier to visualize on paper. One way to visualize a vector field is the following:

1. Pick an arbitrary set of points in the domain at which you want to know the vector field.

2. For each point $(x, y)$ in your set, calculate the vector $\mathbf{v} = \mathbf{F}(x, y)$ at $(x, y)$.

3. For each vector you have calculated, draw an arrow starting at $(x, y)$ in the direction of $\mathbf{v}$ and whose length is proportional to $\|\mathbf{v}\|$.

**Example 21.1** *Plot the vector field*

$$\mathbf{F}(x, y) = \left(-\frac{y}{2}, \frac{x}{4}\right)$$

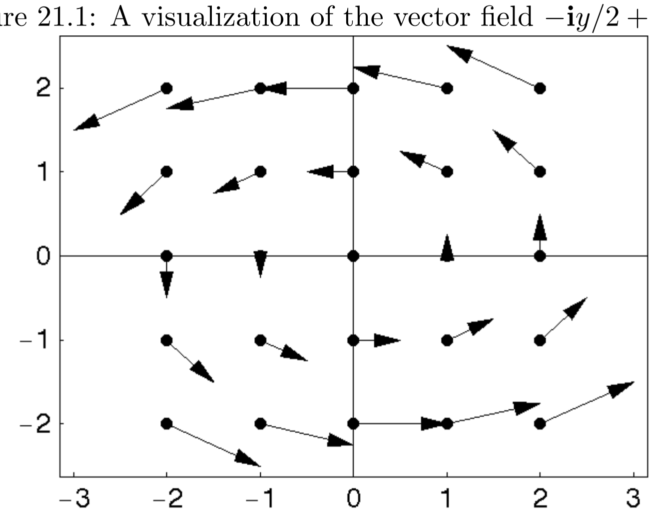
*on the domain* $[-2, 2] \times [-2, 2]$*.*

*Solution.* First, we construct a table of values for the function. We will pick the points at integer coordinates, e.g., $(-2, -2), (-2, -1), ..., (2, 2)$.

| $x$ | $y$ | $v_x$ | $v_y$ |
|---|---|---|---|
| -2 | -2 | 1 | -1/2 |
| -2 | -1 | 1/2 | -1/2 |
| -2 | 0 | 0 | -1/2 |
| -2 | 1 | -1/2 | -1/2 |
| -2 | 2 | -1 | -1/2 |
| -1 | -2 | 1 | -1/4 |
| -1 | -1 | 1/2 | -1/4 |
| -1 | 0 | 0 | -1/4 |
| -1 | 1 | -1/2 | -1/4 |
| -1 | 2 | -1 | -1/4 |
| 0 | -2 | 1 | 0 |
| 0 | -1 | 1/2 | 0 |
| 0 | 0 | 0 | 0 |
| 0 | 1 | -1/2 | 0 |
| 0 | 2 | -1 | 0 |
| 1 | -2 | 1 | 1/4 |
| 1 | -1 | 1/2 | 1/4 |
| 1 | 0 | 0 | 1/4 |
| 1 | 1 | -1/2 | 1/4 |
| 1 | 2 | -1 | 1/4 |
| 2 | -2 | 1 | 1/2 |
| 2 | -1 | 1/2 | 1/2 |
| 2 | 0 | 0 | 1/2 |
| 2 | 1 | -1/2 | 1/2 |
| 2 | 2 | -1 | 1/2 |

We can then sketch the plot using the data in the table. ■

Figure 21.1: A visualization of the vector field $-\mathbf{i}y/2 + \mathbf{j}x/4$.



**Definition 21.2** *A* **scalar field** *is a function* $f(x,y,z) : D \subset \mathbb{R}^3 \mapsto \mathbb{R}$ *that associates a scalar (number) with each point in a subset* $D$ *of* $\mathbb{R}^3$.

**Definition 21.3** *The* **gradient** *of a scalar field* $f(x,y,z)$ *is the vector field*

$$\mathbf{grad} f(x,y,z) = \nabla f(x,y,z) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right) \tag{21.1}$$

**Definition 21.4** *The* **gradient operator** *is the vector operator*

$$\mathbf{grad} = \nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \tag{21.2}$$

The gradient operator is the 3-dimensional analogue of the differential operators such as

$$\frac{\partial}{\partial y} \quad \text{or} \quad \frac{d}{dx}$$

Just as with these scalar operators, the gradient operator works by operating on anything written to the right of it. The operations are distributed through to the different components of the vector operator. Thus, for example,

$$\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right)$$

and

$$\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) uv = \left(\frac{\partial (uv)}{\partial x}, \frac{\partial (uv)}{\partial y}, \frac{\partial (uv)}{\partial z}\right) \tag{21.3}$$

Equation 21.3 can be used to derive the product rule for gradients:

$$\nabla(uv) = u\nabla v + v\nabla u \tag{21.4}$$

**Definition 21.5** *A* **gradient field** $\mathbf{F}(x,y,z)$ *is a vector field that is the gradient of some scalar field* $f(x,y,z)$*. If such a function* $f$ *exists, it is called the* **potential function of the gradient field F** *and* **F** *is said to be a* **conservative vector field**.

**Example 21.2** *Find the gradient field* $\mathbf{F}(x,y,z)$ *corresponding to the function* $f(x,y,z) = x^2z + 3xy$.

*Solution.* The gradient field is

$$\begin{aligned}
\mathbb{F}(x,y,z) &= \nabla f(x,y,z) \\
&= \nabla\left(x^2z + 3xy\right) \\
&= \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)\left(x^2z + 3xy\right) \\
&= \left(\frac{\partial}{\partial x}\left(x^2z + 3xy\right), \frac{\partial}{\partial y}\left(x^2z + 3xy\right), \frac{\partial}{\partial z}\left(x^2z + 3xy\right)\right) \\
&= \left(2xz + 3y, 3x, x^2\right) \quad \blacksquare
\end{aligned}$$

Not all vector fields are gradient field. In the next lecture we will derive a test for determining if a vector field is a gradient field, and a method for determining the potential function that gives rise to the gradient field.

**Example 21.3** *Find the gradient field for the scalar function (given in polar coordinates)*

$$f(r, \theta) = a/r \tag{21.5}$$

*where a is a constant.*

*Solution.* We begin by converting to rectangular coordinates, using the substitution $r = \sqrt{x^2 + y^2}$, to give

$$f(x, y) = \frac{a}{\sqrt{x^2 + y^2}}$$

Differentiating,

$$\begin{aligned}
\mathbf{F}(x, y) &= \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right) \frac{a}{\sqrt{x^2 + y^2}} \\
&= \left(\frac{\partial}{\partial x} \frac{a}{\sqrt{x^2 + y^2}}, \frac{\partial}{\partial y} \frac{a}{\sqrt{x^2 + y^2}}\right) \\
&= \left(-\frac{ax}{(x^2 + y^2)^{3/2}}, -\frac{ay}{(x^2 + y^2)^{3/2}}\right) \\
&= -\frac{a}{(x^2 + y^2)^{3/2}} (x, y)
\end{aligned}$$

Returning to polar coordinates, we observe that since $\mathbf{r} = (x, y)$ and $r^2 = x^2 + y^2$, we have

$$\mathbf{F}(r, \theta) = -\frac{a\mathbf{r}}{r^3} \tag{21.6}$$

Equations 21.5 and 21.6 are the formulas for the Newtonian Gravitational potential and force, respectively, between two bodies of mass $m_1$ and $m_2$ if we make the substitution $a = Gm_1m_2$, where $G \approx 6.6742 \times 10^{-11}$ meters$^3$ / (second$^2$ kilogram) is the universal constant of gravity. ■

We can treat the gradient operator

$$\mathbf{grad} = \nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$$

in many ways just like a vector. We can use it, for example, in a dot product or cross product to define new vector operators.

**Definition 21.6** *The* **divergence of a vector field F** $= (F_1, F_2, F_3)$ *is given by the dot product*

$$\mathbf{div\ F} = \nabla \cdot \mathbf{F} = \left(\frac{\partial F_1}{\partial x}, \frac{\partial F_2}{\partial y}, \frac{\partial F_3}{\partial z}\right) \tag{21.7}$$

*The divergence of a vector field is a scalar field.*

**Example 21.4** *Find the divergence of the vector field*

$$\mathbf{F}(x, y, z) = x\mathbf{i} + \left(y^2x + z\right)\mathbf{j} + e^x\mathbf{k}$$

*Solution.*

$$\begin{aligned}
\nabla \cdot \mathbf{F} &= \left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}\right) \cdot \left(x\mathbf{i} + \left(y^2x + z\right)\mathbf{j} + e^x\mathbf{k}\right) \\
&= \frac{\partial x}{\partial x} + \frac{\partial}{\partial y}\left(y^2x + z\right) + \frac{\partial}{\partial z}e^x \\
&= 1 + 2xy \quad \blacksquare
\end{aligned}$$

**Example 21.5** *Show that* $\nabla \cdot (f\mathbf{a}) = \mathbf{a} \cdot \nabla f$

*Solution.* Suppose that $\mathbf{a} = (a_1, a_2, a_3)$. Then since $a_1$, $a_2$, and $a_3$ are all constants,

$$\begin{aligned}
\nabla \cdot (f\mathbf{a}) &= \nabla \cdot (a_1 f, a_2 f, a_3 f) \\
&= \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \cdot \\
&= \frac{\partial(a_1 f)}{\partial x} + \frac{\partial(a_2 f)}{\partial y} + \frac{\partial(a_3 f)}{\partial z} \\
&= a_1\frac{\partial f}{\partial x} + a_2\frac{\partial f}{\partial y} + a_3\frac{\partial}{\partial z} \\
&= (a_1, a_2, a_3) \cdot \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right) \\
&= \mathbf{a} \cdot \nabla f \quad \blacksquare
\end{aligned}$$

**Definition 21.7** *The* **curl of a vector field** $\mathbf{F} = (F_1, F_2, F_3)$ *is given by the cross product*

$$\textbf{curl } \mathbf{F} = \nabla \times \mathbf{F} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \times (F_1, F_2, F_3) \tag{21.8}$$

*The curl of a vector field is another vector field.*

We can derive a formula for the curl as follows:

$$\begin{aligned}
\nabla \times \mathbf{F} &= \begin{pmatrix} 0 & -\frac{\partial}{\partial z} & \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} & 0 & -\frac{\partial}{\partial x} \\ -\frac{\partial}{\partial y} & \frac{\partial}{\partial x} & 0 \end{pmatrix} \begin{pmatrix} F_1 \\ F_2 \\ F_3 \end{pmatrix} \\
&= \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}, \frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}, \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)
\end{aligned}$$

**Example 21.6** *Find* $\nabla \times F$ *for*

$$\mathbf{F} = \left(e^{y^2},\ 2xye^{y^2},\ 1\right)$$

*Solution.* The **curl** is

$$\begin{aligned}
\nabla \times \mathbf{F} &= \begin{pmatrix} 0 & -\frac{\partial}{\partial z} & \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} & 0 & -\frac{\partial}{\partial x} \\ -\frac{\partial}{\partial y} & \frac{\partial}{\partial x} & 0 \end{pmatrix} \begin{pmatrix} e^{y^2} \\ 2xye^{y^2} \\ 1 \end{pmatrix} \\
&= \left( \frac{\partial}{\partial y}(1) - \frac{\partial}{\partial z}\left(2xye^{y^2}\right), \frac{\partial}{\partial z}\left(e^{y^2}\right) - \frac{\partial}{\partial x}(1), \frac{\partial}{\partial x}\left(2xye^{y^2}\right) - \frac{\partial}{\partial y}\left(e^{y^2}\right) \right) \\
&= \left(0, 0, 2ye^{y^2} - 2ye^{y^2}\right) \\
&= \mathbf{0} \quad \blacksquare
\end{aligned}$$

**Example 21.7** *Find* $\nabla \times \mathbf{F}$ *for*

$$\mathbf{F} = \left(x^2 - y^2, 2xy, 0\right)$$

*Solution.* The **curl** is

$$\begin{aligned}
\nabla \times \mathbf{F} &= \begin{pmatrix} 0 & -\frac{\partial}{\partial z} & \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} & 0 & -\frac{\partial}{\partial x} \\ -\frac{\partial}{\partial y} & \frac{\partial}{\partial x} & 0 \end{pmatrix} \begin{pmatrix} x^2 - y^2 \\ 2xy \\ 0 \end{pmatrix} \\
&= \left( \frac{\partial}{\partial y}(0) - \frac{\partial}{\partial z}(2xy), \frac{\partial}{\partial z}\left(x^2 - y^2\right) - \frac{\partial}{\partial x}(0), \frac{\partial}{\partial x}(2xy) - \frac{\partial}{\partial y}\left(x^2 - y^2\right) \right) \\
&= (0, 0, 4y) \quad \blacksquare
\end{aligned}$$

**Definition 21.8** *The* **Laplacian of a scalar field** $f(x, y, z)$ *is given by the product*

$$\nabla^2 f = \nabla \cdot \nabla f = \mathbf{div\ grad} f = \frac{\partial^2 f}{\partial x} + \frac{\partial^2 f}{\partial y} + \frac{\partial^2 f}{\partial z} \tag{21.9}$$

*The Laplacian of a scalar field is another scalar field.*

**Theorem 21.1 Properties of Vector Operators** *Let* $f : \mathbb{R}^3 \mapsto \mathbb{R}$ *be a scalar field and* $\mathbf{F} : \mathbb{R}^3 \mapsto \mathbb{R}^3$ *be a vector field. Then all of the following properties hold:*

(a) $\nabla \cdot \nabla \times \mathbf{F} = 0$

(b) $\nabla \times \nabla f = \mathbf{0}$

(c) $\nabla \cdot (f\mathbf{F}) = f\nabla \cdot \mathbf{F} + \mathbf{F} \cdot \nabla f$

(d) $\nabla \times (f\mathbf{F}) = f\nabla \times \mathbf{F} - \mathbf{F} \times \nabla f$

(e) $\nabla \cdot (\mathbf{F} \times \mathbf{G}) = \mathbf{G} \cdot \nabla \times \mathbf{F} - \mathbf{F} \cdot \nabla \times \mathbf{G}$

(f) $\nabla \cdot (\nabla f \times \nabla g) = 0$

# Lecture 22

# Line Integrals

Suppose that we are traveling along a path from a point **P** towards a second point **Q**, and that are position $\mathbf{r}(t)$ on this trajectory is described parametrically in terms of a parameter $t$. We can associate two "directions" of travel along this path: one from **P** to **Q**, and the second back in the other direction from **Q** to **P**. We can visualize these directions by drawing arrows along the path.

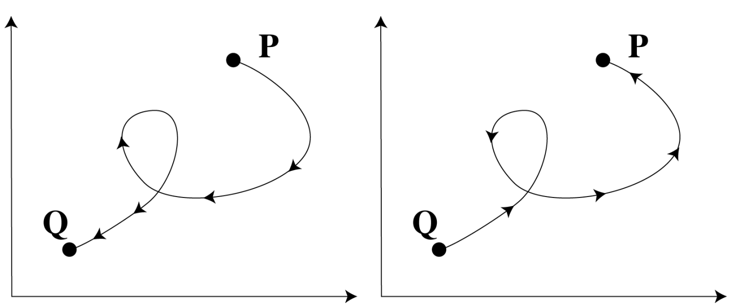
**Definition 22.1** *Let $C$ be a curve connecting two points* **P** *and* **Q** *that is parameterized as*

$$\mathbf{r}(t), \quad a \leq t \leq b$$

*where $\mathbf{r}(a) = \mathbf{P}$ and $\mathbf{r}(b) = \mathbf{Q}$. Then the curve is said to* **oriented** *if we associate a direction of travel with it, and the path is said to be* **directed**.

*Furthermore, the path is said to be* **positively oriented** *if the direction of motion corresponds to increasing $t$, and* **negatively oriented** *otherwise.*

Figure 22.1: Examples of oriented curves, where $a \leq t \leq b$, $\mathbf{P} = \mathbf{r}(a)$, and $\mathbf{Q} = \mathbf{r}(b)$. The curve on the left is positively oriented, because the motion follows the direction of increasing $t$. The curve on the right is oriented but it is not positively oriented.




The arrows on the oriented curve really represent the direction of the tangent vectors; we can approximate them by dividing the parameterization up into finite intervals

$$a = t_0 < t_1 < t_2 < \cdots < t_n = b \tag{22.1}$$

and breaking the path itself up into points

$$\mathbf{r}(t_1), \mathbf{r}(t_2), ..., \mathbf{r}(t_n) \tag{22.2}$$

Let us define the differential vector

$$d\mathbf{r}_i = \mathbf{r}(t_{i+1}) - \mathbf{t_i} \tag{22.3}$$

and the differential time element

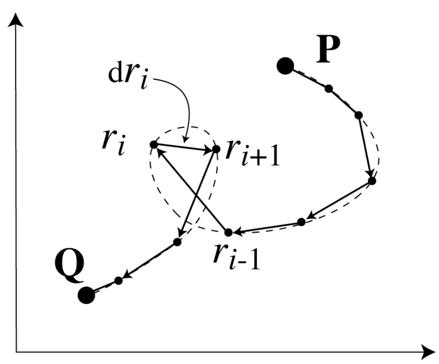
$$dt_i = t_{i+1} - t_i \tag{22.4}$$

Then in the limit as $dt_i \to 0$, we have

$$\lim_{dt_i \to 0} d\mathbf{r}_i = \mathbf{r}'(t_i)dt = \mathbf{v}(t_i)dt \tag{22.5}$$

Now suppose that the curve is embedded in some vector field $\mathbf{F}(\mathbf{r})$. Then at any

Figure 22.2: Partitioning of an oriented curve into small segments that approximate the tangent vectors.




point $\mathbf{r}$ along the curve we can calculate the dot product

$$\mathbf{r}(t) \cdot \mathbf{F}(\mathbf{r}(t))$$

This product is a measure of how much alignment there is between the motion along the curve and the direction of the vector field. The dot product is maximized when the angle between the two vectors is zero, and we are moving "with the flow" of the field. If we are moving completely in the opposite direction, the dot product is negative, and if we move in a direction perpendicularly to the vector field than the dot product is zero, as there is no motion with with or against the flow. We define the **line integral of $f$ along $C$** as the Riemann Sum

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \lim_{dr_i \to 0} \sum_{i=1}^{n} d\mathbf{r}_i \cdot \mathbf{F}(\mathbf{r}_i) \tag{22.6}$$

or equivalently

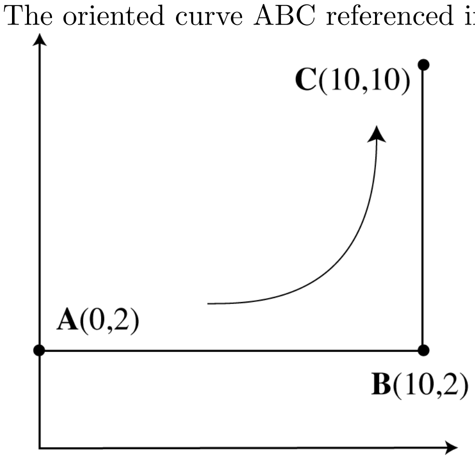
$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_a^b \left[ \frac{d\mathbf{r}}{dt} \cdot \mathbf{F}(\mathbf{r}(t)) \right] dt \tag{22.7}$$

**Example 22.1** *Find the line integral of the vector field*

$$\mathbf{F}(\mathbf{r}(t)) = (2x, 3y, 0)$$

*over the path ABC as illustrated in figure 22.3*

Figure 22.3: The oriented curve ABC referenced in example 22.1




*Solution.* We can parameterize the curve as

$$\mathbf{r}(t) = \begin{cases} (t, 2, 0) & 0 \leq t \leq 10 \\ (10, t-8, 0) & 10 \leq t \leq 18 \end{cases} \tag{22.8}$$

Based on this parameterization, we can write equations for the individual coordinates $x(t)$ and $y(t)$

$$x(t) = \begin{cases} t & 0 \leq t \leq 10 \\ 10 & 10 \leq t \leq 18 \end{cases} \tag{22.9}$$

$$y(t) = \begin{cases} 2 & 0 \leq t \leq 10 \\ t-8 & 10 \leq t \leq 18 \end{cases} \tag{22.10}$$

and for the velocity

$$\mathbf{r}'(t) = \begin{cases} (1, 0, 0) & 0 \leq t \leq 10 \\ (0, 1, 0) & 10 \leq t \leq 18 \end{cases} \tag{22.11}$$

Therefore

$$\mathbf{F} \cdot \mathbf{r}' = \begin{cases} (2x, 3y, 0) \cdot (1, 0, 0) & 0 \leq t \leq 10 \\ (2x, 3y, 0) \cdot (0, 1, 0) & 10 \leq t \leq 18 \end{cases} \tag{22.12}$$

$$= \begin{cases} 2x & 0 \leq t \leq 10 \\ 3y & 10 \leq t \leq 18 \end{cases} \tag{22.13}$$

$$= \begin{cases} 2t & 0 \leq t \leq 10 \\ 3(t-8) & 10 \leq t \leq 18 \end{cases} \tag{22.14}$$

Therefore the line integral is

$$\int_{ABC} \mathbf{F}\cdot d\mathbf{r} = \int_0^{10} 2tdt + \int_{10}^{18} (3t-24)dt \tag{22.15}$$

$$= t^2\Big|_0^{10} + \left(\frac{3}{2}t^2 - 24t\right)\Big|_{10}^{18} \tag{22.16}$$

$$= 10 - 0 + \frac{3\times 18^2}{2} - (24)(18) - \frac{3\times 10^2}{2} + 240 \tag{22.17}$$

$$= 10 + 486 - 432 + 240 = 304 \ \blacksquare \tag{22.18}$$

**Definition 22.2** *The* **work done on an object** *when it is moved along a path $C$ and subjected to a force* $\mathbf{F}$ *is*

$$W = \int_C \mathbf{F}\cdot\mathbf{r} \tag{22.19}$$

**Example 22.2** *Find the work required to lift a satellite of mass 1000 kg from the earth's surface (*$r = 6300$ km*) to a geostationary orbit (*$r = 42,000$ km*) under the influence of the Earth's gravity*

$$\mathbf{F} = -\frac{\mu m}{r^2}\mathbf{k}$$

*where $r$ is the distance of the object from the center of the Earth; $m$ is its mass; and* $\mu = GM = 3.986\times 10^{14}$ meter$^3$/second$^2$ *is the product of the universal constant of gravity and the mass of the Earth. (Work is measured in units of* joules*, where one* joule *equals one* kilogram meter$^2$/second$^2$.*)*

*Solution.*The work is the line integral

$$W = \int_{6.3\times 10^6}^{42\times 10^6} \mathbf{F}\cdot d\mathbf{r}$$

where we have converted the distances from meters to kilometers. Substituting the equation for $\mathbf{F}$ and assuming the trajectory moves completely in the $z$ directions,

$$\begin{aligned} W &= -\mu m \int_{6.3\times 10^6}^{42\times 10^6} \frac{1}{r^2}dz \\ &= \frac{\mu m}{r}\Big|_{6.3\times 10^6}^{42\times 10^6} \\ &= (3.986\times 10^{14})\times(1000)\left(\frac{1}{42\times 10^6} - \frac{1}{6.3\times 10^6}\right) \text{ Joules} \\ &= -5.4\times 10^{10} \text{ Joules} \end{aligned}$$

The work is negative because the line integral measured the work done by the force field on the satellite, i.e., work must be done by the satellite against the field to get it into orbit. Converting to day-to-day units, since one watt is the same as one Joule/second, one kilowatt-hour is the energy used by expending 1000 Joules per second (one kilowatt) for one hour (3600 seconds), namely $3.6\times 10^6$ Joules.Thus the energy required to lift 1000 kg is $(5.4\times 10^{10})/(3.6\times 10^6) = 15,000$ kW-hours. That's equivalent to leaving a 60-Watt light on for 28 years; running a typical 1000-Watt hair dryer continuously for about 20 months; or running a typical home dryer (2800 Watts) for about 7 months. ■

**Definition 22.3** *A* **closed curve** *is a curve that follows a path from a point back to itself, such as a circle.*

**Definition 22.4** *Let $C$ be a closed curve, and $\mathbf{F}$ be a vector field. Then the* **circulation of the vector field is**

$$\oint_C \mathbf{F}\cdot d\mathbf{r}$$

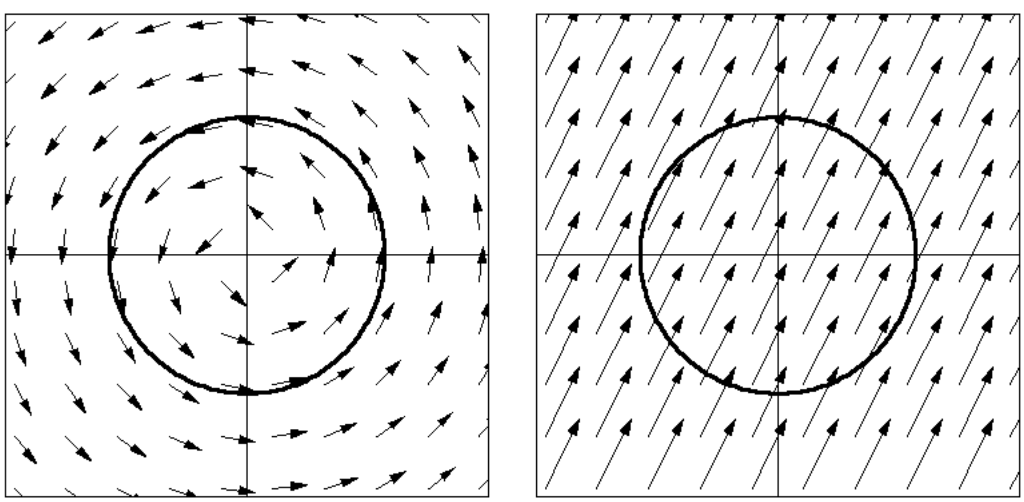
*where the special integral symbol $\oint$ indicates that the path of integration is closed.*

The circulation of a vector field is a measure of the "circularity" of its flow. For example consider the two vector fields in figure 22.4. Suppose we follow the illustrated circular paths in a counter-clockwise fashion in both cases, such as

$$\mathbf{r}(t) = (\cos t, \sin t), \quad 0 \le t \le 2\pi$$

In the vector field on the left, the magnitude of the field is the same everywhere,

Figure 22.4: Illustration of the same path through two different vector fields.



but is direction is circular, giving a flow around the origin.

$$\mathbf{F}(x, y) = \frac{1}{\sqrt{x^2+y^2}}(-y, x)$$

the direction of the field is nearly parallel to the direction of motion, during the entire path. Along this path, $x = \cos t$ and $y = \sin t$ and $\mathbf{r}' = (-\sin t, \cos t)$, so

$$\begin{aligned}\oint_c \mathbf{F} \cdot d\mathbf{r} &= \int_0^{2\pi} \frac{1}{\sqrt{\sin^2 t + \cos^2}} (-\sin t, \cos t) \cdot (-\sin t, \cos t) dt \\ &= \int_0^{2\pi} (\sin^2 t + \cos^2)^{1/2} dt \\ &= \int_0^{2\pi} dt \\ &= 2\pi\end{aligned}$$

Now consider the vector field on the right hand side of figure 22.4. In this field the vector field is a fixed constant everywhere,

$$\mathbf{G} = (1, 2)$$

The path is the same, which is a circle. Corresponding to any point on the path, the tangent vector takes on some value $\mathbf{T}$, there is a point directly opposite it on the path where the tangent vector is $-\mathbf{T}$, and hence the dot products $\mathbf{T} \cdot \mathbf{G}$ and $-\mathbf{T} \cdot \mathbf{G}$ cancel out on these two points of the path.

$$\begin{aligned}\oint_c \mathbf{G} \cdot d\mathbf{r} &= \int_0^{2\pi} (1, 2) \cdot (-\sin t, \cos t) dt \\ &= \int_0^{2\pi} (-\sin t + 2\cos t) dt \\ &= 0\end{aligned}$$

because the integral of the sine and cosine functions over a complete period is always zero.

The question naturally arises, does the circulation decline towards this limit if we were to move our path away from the origin in the first case? As we move further and further from the origin, we would observe that the field becomes more and more nearly constant, and we would expect this occur. We can check this by moving the path away from the origin, to a new center $(x_0, y_0)$, so that

$$x = x_0 + \cos t, y = y_0 + \sin t$$

The integral then becomes

$$\begin{aligned}
\oint_c \mathbf{F}\cdot d\mathbf{r} &= \int_0^{2\pi} \frac{1}{\sqrt{(y_0+\sin t)^2+(x_0+\cos t)^2}}(-y_0-\sin t, x_0+\cos t)\cdot(-\sin t,\cos t)dt \\
&= \int_0^{2\pi} \frac{y_0\sin t + x_0\cos t + \sin^2 t + \cos^2}{\sqrt{y_0^2+2y_0\sin t+\sin^2 t + x_0^2 + 2x_0\cos t + \cos^2 t}}dt \\
&= \int_0^{2\pi} \frac{1+y_0\sin t + x_0\cos t}{\sqrt{1+x_0^2+y_0^2+2x_0\cos t + 2y_0\sin t}}dt \\
&= \frac{1}{\sqrt{x_0^2+y_0^2}}\int_0^{2\pi} \frac{1+y_0\sin t+x_0\cos t}{\sqrt{1+(1+2x_0\cos t+2y_0\sin t)/(x_0^2+y_0^2)}}dt
\end{aligned}$$

If $x_0 \gg 1$ and $y0 \gg 1$, then the quantity

$$\frac{1+2x_0\cos t+2y_0\sin t}{x_0^2+y_0^2} \ll 1$$

and we can use the approximation

$$(1+u)^{-1/2} \approx 1 - \frac{u}{2} + \cdots$$

to give

$$\frac{1}{\sqrt{1+(1+2x_0\cos t+2y_0\sin t)/(x_0^2+y_0^2)}} \approx 1 + \frac{1+x_0\cos t+y_0\sin t}{x_0^2+y_0^2} + \cdots$$

Hence the circulation becomes

$$\begin{aligned}
\oint_c \mathbf{F}\cdot d\mathbf{r} = \frac{1}{r_0}\int_0^{2\pi} &(1+y_0\sin t+x_0\cos t)\times \\
&\left(1+\frac{1+x_0\cos t+y_0\sin t}{r_0^2}+\cdots\right)dt
\end{aligned}$$

where $r_0^2 = x_0^2+y_0^2$. Each successive term in the expansion falls off by a factor of $1/r_0$. Using the facts that $\int_0^{2\pi}\sin t dt = 0$ and $\int_0^{2\pi}\cos t dt = 0$,

$$\begin{aligned}
\oint_c \mathbf{F}\cdot d\mathbf{r} &= \frac{1}{r_0}\int_0^{2\pi}\left[(1+y_0\sin t+x_0\cos t) + \frac{(1+y_0\sin t+x_0\cos t)^2}{r_0^2}+\cdots\right]dt \\
&= \frac{2\pi}{r_0} + \frac{1}{r_0^3}\int_0^{2\pi}\left[1+2(y_0\sin t+x_0\cos t)+(y_0\sin t+x_0\cos t)^2+\cdots\right]dt \\
&= \frac{2\pi}{r_0}+\frac{2\pi}{r_0^3}+\frac{1}{r_0^3}\int_0^{2\pi}(y_0\sin t+x_0\cos t)^2 dt+\cdots
\end{aligned}$$

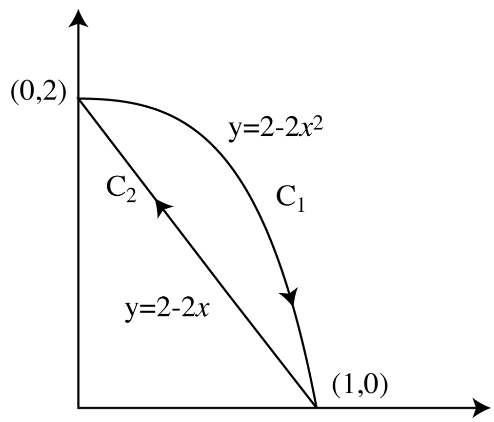
Now use the facts that $\int_0^{2\pi}\cos^2 t dt = \int_0^{2\pi}\sin^2 t dt = \pi$ and $\int_0^{2\pi}\sin t\cos t dt = 0$,

$$\begin{aligned}
\oint_c \mathbf{F}\cdot d\mathbf{r} &= \frac{2\pi}{r_0}+\frac{2\pi}{r_0^3}+\frac{1}{r_0^3}\int_0^{2\pi}(y_0^2\sin^2 t+2x_0y_0\sin t\cos t+x_0^2\cos^2 t)dt+\cdots \\
&= \frac{2\pi}{r_0}+\frac{2\pi}{r_0^3}+\frac{\pi r_0^2}{r_0^3}+\cdots = \frac{3\pi}{r_0}+\cdots
\end{aligned}$$

Thus the circulation falls off inversely as we move away from the origin, approaching zero as the arrows become more and more parallel, as our intuitive notion suggested.

**Example 22.3** *Compute the line integral of the function* $\mathbf{F} = (x, 2y)$ *over the curve* $C$ *shown in figure 22.5.*

Figure 22.5: Integration path for example 22.3.




*Solution.* To complete this integral, we divide the curve into two parts. Let us call the parabola $y = 2-2x^2$ from $(0, 2)$ to $(1, 0)$ by the name $C_1$, and let us call the line back along $y = 2 - 2x$ from $(1, 0)$ to $(0, 2)$ by the name $C_2$, as indicated in figure 22.5.

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \int_{C_1} \mathbf{F} \cdot d\mathbf{r} + \int_{C_2} \mathbf{F} \cdot d\mathbf{r}$$

We can parameterize the path $C_1$ by

$$x = t, \quad y = 2 - 2t^2, \quad 0 \le t \le 1$$

Thus on $C_1$, $\mathbf{r} = (t, 2 - 2t^2)$, $\mathbf{r}' = (1, -4t)$, and $\mathbf{F} = (x, 2y) = (t, 4 - 4t^2)$ and

$$\begin{aligned}
\int_{C_1} \mathbf{F} \cdot d\mathbf{r} &= \int_0^1 (t, 4 - 4t^2) \cdot (1, -4t) dt \\
&= \int_0^1 (t - 16t + 16t^3) dt \\
&= \left( -\frac{15t^2}{2} + \frac{16t^4}{4} \right) \Bigg|_0^1 \\
&= -\frac{15}{2} + 4 = -\frac{7}{2}
\end{aligned}$$

On the line $y = 2 - 2x$, we start at the point (1,0) and move toward the point (0, 2). We can parameterize this as

$$x = 1 - t, \quad y = 2 - 2(1 - t) = 2t, \quad 0 \le t \le 1$$

so that the second integral becomes

$$\begin{aligned}\int_{C_2} \mathbf{F}\cdot d\mathbf{r} &= \int_0^1 (x,2y)\cdot \frac{d}{dt}(1-t,2t)dt \\ &= \int_0^1 (1-t,4t)\cdot(-1,2)dt \\ &= \int_0^1 (-1+9t)dt \\ &= \left(-t+\frac{9t^2}{2}\right)\bigg|_0^1 = \frac{7}{2}\end{aligned}$$

Therefore the integral over C is

$$\oint_C \mathbf{F}\cdot d\mathbf{r} = \int_{C_1} \mathbf{F}\cdot d\mathbf{r} + \int_{C_2} \mathbf{F}\cdot d\mathbf{r} = -\frac{7}{2}+\frac{7}{2} = 0 \quad \blacksquare$$

**Example 22.4** *A particle travels along the helix*

$$\mathbf{r} = (\cos t, \sin t, 2t)$$

*in the vector field*

$$\mathbf{F} = (x, z, -xy)$$

*Find the total work done over the time period* $0 \le t \le 3\pi$.

*Solution.* The work is given by

$$\begin{aligned}W &= \int_0^{3\pi} \mathbf{F}\cdot\mathbf{r}' dt \\ &= \int_0^{3\pi} (x,z,-xy)\cdot\frac{d}{dt}(\cos t,\sin t,2t)dt \\ &= \int_0^{3\pi} (\cos t, 2t, -\cos t\sin t)\cdot(-\sin t,\cos t,2)dt \\ &= \int_0^{3\pi} (-\cos t\sin t + 2t\cos t - 2\cos t\sin t)\, dt \\ &= -3\int_0^{3\pi} \cos t\sin t dt + 2\int_0^{3\pi} t\cos t dt\end{aligned}$$

Using the integral formulas $\int \sin t\cos t dt = \frac{1}{2}\sin^2 t$ and $\int t\cos t = \cos t + t\sin t$ gives

$$\begin{aligned}W &= -3\left(\frac{1}{2}\sin^2 t\right)\bigg|_0^{3\pi} + 2\,(\cos t + t\sin t)|_0^{3\pi} \\ &= 2(\cos 3\pi - \cos 0) = -4 \quad \blacksquare\end{aligned}$$

Recall from integral calculus the following statement of the fundamental theorem of calculus: If $f$ is differentiable and $f'(t)$ is integrable on some interval $(a, b)$ then

$$\int_a^b f'(t)dt = f(b) - f(a) \tag{22.20}$$

In the generalization to line integrals, the derivative becomes a directional derivative, but otherwise the statement of the theorem remains almost unchanged.

**Theorem 22.1 Fundamental Theorem of Calculus for Line Integrals** *Suppose that $C$ is a piecewise smooth curve that can be parameterized as*

$$C = \{\mathbf{r}(t), a \leq t \leq b\} \subset S \tag{22.21}$$

*for some open set $S \subset \mathbb{R}^3$, and let*

$$\mathbf{A} = \mathbf{r}(a) \tag{22.22}$$

$$\mathbf{B} = \mathbf{r}(b) \tag{22.23}$$

*Then if $f : S \mapsto \mathbf{R}^3$ is continuously differentiable on $S$,*

$$\boxed{\int_C \nabla f(\mathbf{r}) \cdot d\mathbf{r} = f(\mathbf{B}) - f(\mathbf{A})} \tag{22.24}$$

Since the directional derivative in the direction of $\mathbf{v}(t)$ is

$$\begin{aligned} D_{\mathbf{v}} f(\mathbf{r}) &= \nabla f(\mathbf{r}(\mathbf{t})) \cdot \mathbf{v}(t) \\ &= \nabla f(\mathbf{r}(\mathbf{t})) \cdot \frac{d\mathbf{r}(t)}{dt} \end{aligned}$$

then the left-hand side of equation 22.24 can be rewritten to give

$$\int_C \nabla f(\mathbf{r}) \cdot \mathbf{r} = \int_a^b D_{\mathbf{v}} f(\mathbf{r}(t))dt \tag{22.25}$$

and the fundamental theorem of calculus for line integrals becomes

$$\boxed{\int_a^b D_{\mathbf{v}} f(\mathbf{r}(t))dt = f(\mathbf{B}) - f(\mathbf{A})} \tag{22.26}$$

*Proof.* Expanding the directional derivative,

$$\begin{aligned} D_v f(\mathbf{r}(t)) &= \nabla f(\mathbf{r}(t)) \cdot \mathbf{v} \\ &= \left( \frac{\partial f(\mathbf{r}(t))}{\partial x}, \frac{\partial f(\mathbf{r}(t))}{\partial y}, \frac{\partial f(\mathbf{r}(t))}{\partial z} \right) \cdot \left( \frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt} \right) \\ &= \frac{\partial f(\mathbf{r}(t))}{\partial x}\frac{dx}{dt} + \frac{\partial f(\mathbf{r}(t))}{\partial y}\frac{dy}{dt} + \frac{\partial f(\mathbf{r}(t))}{\partial z}\frac{dz}{dt} \\ &= \frac{df(\mathbf{r}(t))}{dt} \end{aligned}$$

where the last step follows from the chain rule. By the fundamental theorem of calculus,

$$\int_C \nabla f(\mathbf{r}) \cdot \mathbf{r} = \int_a^b D_{\mathbf{v}} f(\mathbf{r}(t)) dt \quad (22.27)$$

$$= \int_a^b \frac{df(\mathbf{r}(t))}{dt} dt \quad (22.28)$$

$$= f(\mathbf{r}(b)) - f(\mathbf{r}(a)) \quad (22.29)$$

$$= f(\mathbf{B}) - f(\mathbf{A}) \ \blacksquare \quad (22.30)$$

**Theorem 22.2 Path Independence Theorem**. *Let $D$ be an open connected set, and suppose that $F : D \mapsto \mathbb{R}^2 3$ is continous on $D$. Then $\mathbf{F}$ is a gradient field, i.e, there exists some scalar function $f$ such that $\mathbf{F} = \nabla f$, if and only if $\int_C \mathbf{F}(\mathbf{r}) \cdot d\mathbf{r}$ is independent of path, i.e.,*

$$\int_C \mathbf{F}(\mathbf{r}) \cdot d\mathbf{r} = f(\mathbf{B}) - f(\mathbf{A})$$

*for all smooth paths $C$ in $D$,*

$$C = \{\mathbf{r}(t), a \le t \le b\} \subset S$$

*with $\mathbf{A} = \mathbf{r}(a)$ and $\mathbf{B} = \mathbf{r}(b)$.*

*Proof.* Suppse that $\mathbf{F} = \nabla f$, i.e., that $\mathbf{F}$ is a gradient field. Then

$$\int_C \mathbf{F}(\mathbf{r}) \cdot d\mathbf{r} = \int_C \nabla f(\mathbf{r}) \cdot d\mathbf{r}$$

$$= f(\mathbf{B}) - f(\mathbf{A})$$

by the fundamental theorem of calculus for line integrals. Hence the line integral depends only its path.

Now suppose that the line integral depends only on its path. To complete the proof we need to show that there exists some function $f$ such that $\mathbf{F} = \nabla f$. Let $A = (x_a, y_a, z_a)$ and let $B = (x, y, z)$ denote the end points of $C$. Then

$$f(x,y,z) - f(\mathbf{A}) = f(\mathbf{B}) - f(\mathbf{A}) = \int_C \mathbf{F} \cdot d\mathbf{r} = \int_{(x_a,y_a,z_a)}^{(x,y,z)} \mathbf{F} \cdot d\mathbf{r}$$

Define the components of $\mathbf{F}$ as

$$\mathbf{F}(x,y,z) = (M(x,y,z),\ \ N(x,y,z),\ \ P(x,y,z)) \quad (22.31)$$

Since the domain is connected we can pick a point $(x_1, y, z)$ (close to $(x, y, z)$) so that

$$f(x,y,z) - f(\mathbf{A}) = \int_{(x_a,y_a,z_a)}^{(x_1,y,z)} \mathbf{F} \cdot d\mathbf{r} + \int_{(x_1,y,z)}^{(x,y,z)} \mathbf{F} \cdot d\mathbf{r}$$

The first integral does not depend on $x$; in fact the only place that $x$ appears on the right hand side of the equation is in the upper limit of the second integral. But in the second integral the values of $y$ and $z$ are constant, hence $\mathbf{F} \cdot d\mathbf{r} = M(x, y, z)dx$.

$$\frac{\partial f(x, y, z)}{\partial x} = \frac{\partial}{\partial x} \int_{x_1}^{x} M(u, y, z)du = M(x, y, z) \tag{22.32}$$

where the last equality follows from the fundamental theorem of calculus. Similarly we can pick a point $(x, y_1, z)$ such that

$$f(x, y, z) - f(\mathbf{A}) = \int_{(x_a, y_a, z_a)}^{(x, y_1, z)} \mathbf{F} \cdot d\mathbf{r} + \int_{(x, y_1, z)}^{(x, y, z)} \mathbf{F} \cdot d\mathbf{r}$$

Now the first integral does not depend on $y$, and the second integral has $x$ and $z$ fixed, so that $\mathbf{F} \cdot d\mathbf{r} = N(x, y, z)dy$ in the second integrand and

$$\frac{\partial f(x, y, z)}{\partial y} = \frac{\partial}{\partial y} \int_{y_1}^{y} N(x, u, z)du = N(x, y, z) \tag{22.33}$$

Finally, we can pick a point $(x, y, z_1)$ such that

$$f(x, y, z) - f(\mathbf{A}) = \int_{(x_a, y_a, z_a)}^{(x, y, z_1)} \mathbf{F} \cdot d\mathbf{r} + \int_{(x, y, z_1)}^{(x, y, z)} \mathbf{F} \cdot d\mathbf{r}$$

Now the first integral does not depend on $z$, and the second integral has $x$ and $y$ fixed, so that $\mathbf{F} \cdot d\mathbf{r} = P(x, y, z)zy$ in the second integrand and

$$\frac{\partial f(x, y, z)}{\partial z} = \frac{\partial}{\partial z} \int_{z_1}^{z} P(x, y, u)du = P(x, y, z) \tag{22.34}$$

Combining equations 22.31, 22.32, 22.33, and 22.34, we have

$$\mathbf{F} = (M, N, P) = \left( \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right) = \nabla f \tag{22.35}$$

Thus $\mathbf{F} = \nabla f$, i.e, $\mathbf{F}$ is a gradient field. ■

**Theorem 22.3** *A vector field* $\mathbf{F}$ *is a gradient field if and only if*

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = 0$$

*for every closed curved* $C$.

*Proof.* $\mathbf{F}$ is a gradient field if and only if it is path independent. Pick any two points on the closed curve $C$, and call them $A$ and $B$. Let $C_1$ be one path from $A$ to $B$ along $C$, and let $C_2$ be the path from $A$ to $B$ along the other half of $C$. Then

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

The second integral is negative because the $C_2$ is oriented from $A$ to $B$ and not from $B$ to $A$ as it would have to be to complete the closed curves.

But the integrals over $C_1$ and $C_2$ are two paths between the same two points; hence by the path independence theorem,

$$\int_{C_1} \mathbf{F} \cdot d\mathbf{r} = \int_{C_2} \mathbf{F} \cdot d\mathbf{r}$$

and conseequently

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = 0. \quad \blacksquare$$

The following theorem gives an easy test to see if any given field is a gradient field.

**Theorem 22.4** $\mathbf{F}$ *is a gradient field if and only if* $\nabla \times \mathbf{F} = 0$.

*Proof.* First, suppose $\mathbf{F}$ is a gradient field. Then there exists some scalar function $f$ such that $\mathbf{F} = \nabla f$. But

$$\nabla \times \mathbf{F} = \nabla \times \nabla f = 0$$

because **curl grad**$f = 0$ for any function $f$.

The proof in the other direction, namely, that if $\nabla \times \mathbf{F} = 0$ then $\mathbf{F}$ is a gradient function, requires the use of Stoke's theorem, which we shall prove in the next section. Stoke's theorem gives us the formula

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \iint_A (\nabla \times \mathbf{F}) \cdot d\mathbf{r}$$

where $A$ is the area encloses by a closed path $C$. But if the **curl F** is zero, then

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = 0$$

Since the integral over a closed path is zero, the integral must be path-independent, and hence the integrand must be a gradient field. ■

**Theorem 22.5** *The following statements are equivalent:*

1. $\mathbf{F}$ *is a gradient field.*
2. $\nabla \times \mathbf{F} = 0$
3. *There exists some scalar function* $f$ *such that* $\mathbf{F} = \nabla f$.
4. *The line integral is path independent: for some scalar function* $f$,

$$\int_{\mathbf{A}}^{\mathbf{B}} \mathbf{F} \cdot d\mathbf{r} = f(\mathbf{B}) - f(\mathbf{A})$$

5. *The line integral around a closed curve is zero, i.e.,*

$$\oint_C \mathbf{F}\cdot d\mathbf{r} = 0$$

*for all closed curves $C$.*

**Example 22.5** *Determine if*

$$\mathbf{F} = \left(x^2 - y^2, -2xy, z\right)$$

*is a gradient field.*

*Solution.* We calculate the **curl F**,

$$\begin{aligned}\nabla\mathbf{F} &= \begin{pmatrix} 0 & -D_z & D_y \\ D_z & 0 & -D_x \\ -D_y & D_x & 0 \end{pmatrix}\begin{pmatrix} x^2-y^2 \\ -2xy \\ z\end{pmatrix} = \begin{pmatrix} -D_z(-2xy) + D_y(z) \\ D_z(x^2-y^2) - D_x(z) \\ -D_y(x^2-y^2) + D_x(-2xy)\end{pmatrix} \\ &= (0, 0, 2y - 2y) = \mathbf{0}\end{aligned}$$

Hence the field is a gradient field. ■

**Example 22.6** *Determine if* $\mathbf{F} = (-y, z, x)$ *is a gradient field.*

*Solution.*

$$\begin{aligned}\nabla\mathbf{F} &= \begin{pmatrix} 0 & -D_z & D_y \\ D_z & 0 & -D_x \\ -D_y & D_x & 0 \end{pmatrix}\begin{pmatrix} -y \\ z \\ x\end{pmatrix} = \begin{pmatrix} -D_z(z) + D_y(x) \\ D_z(-y) - D_x(x) \\ -D_y(-y) + D_x(z)\end{pmatrix} \\ &= (-1, -1, 1) \neq \mathbf{0}\end{aligned}$$

Hence the field is not a gradient field. ■

**Example 22.7** *Find the potential function $f(x, y, z)$ for the gradient field in example 22.5.*

*Solution.* The vector field is

$$\mathbf{F} = (x^2 - y^2, -2xy, z) = (M, N, P)$$

where $M = x^2 - y^2$, $N = -2xy$, and $P = z$. Since $\mathbf{F}$ is a gradient field, then for some scalar function $f$,

$$f_x = M = x^2 - y^2 \tag{22.36}$$

$$f_y = N = -2xy \tag{22.37}$$

$$f_z = P = z \tag{22.38}$$

Integrating the first equation with respect to $x$,

$$f(x, y, z) = \int f_x dx = \int (x^2 - y^2)dx = \frac{1}{3}x^3 - y^2x + h(y, z) \tag{22.39}$$

The function $h(y, z)$ is the constant of integration: we integrated over $x$, so the constant may still depend on $y$ or $z$. Differentiate with respect to $y$:

$$f_y = -2xy + h_y(y, z) \tag{22.40}$$

By equation 22.37

$$\begin{aligned} -2xy + h_y(y, z) &= -2xy \\ h_y(y, z) &= 0 \end{aligned}$$

Therefore $h$ does not depend on $y$, only on $z$, and we can write

$$f(x, y, z) = \frac{1}{3}x^3 - y^2x + h(z) \tag{22.41}$$

Differentiating with respect to $z$,

$$f_z = \frac{dh}{dz} \tag{22.42}$$

By equation 22.38

$$\frac{dh}{dz} = f_z = z \tag{22.43}$$

Hence

$$h(z) = \int_z h'(z)dz = \int_z zdz = \frac{1}{2}z^2 + C \tag{22.44}$$

From equation 22.41

$$f(x, y, z) = \frac{1}{3}x^3 - y^2x + \frac{1}{2}z^2 + C \tag{22.45}$$

where $C$ is any arbitrary constant. ■

# Lecture 23

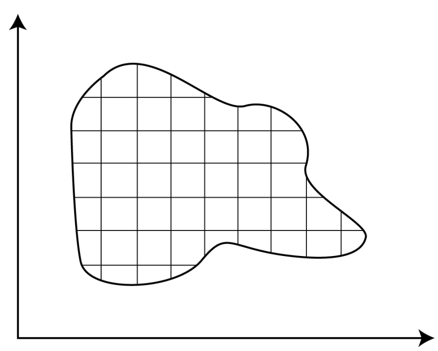
# Green's Theorem

**Theorem 23.1 Green's Theorem**. *Let $C$ be a closed, oriented curve enclosing a region $R \subset \mathbb{R}^2$ with no holes, and let $F = (M(x, y), N(x, y)$ be a vector field in $\mathbb{R}^2$. The*

$$\boxed{\oint_C \mathbf{F} \cdot d\mathbf{r} = \oint_C (M dx + N dy) = \iint_R \left( \frac{\partial N}{\partial x} - \frac{\partial M}{\partial y} \right) dA} \tag{23.1}$$

*Proof.* One way to evaluate the line integral about a closed path is to divide up the enclosed area into small rectangles. Focusing on a pair of small rectangles,

Figure 23.1: A closed curve may be filled with an large number of tiny rectangles.

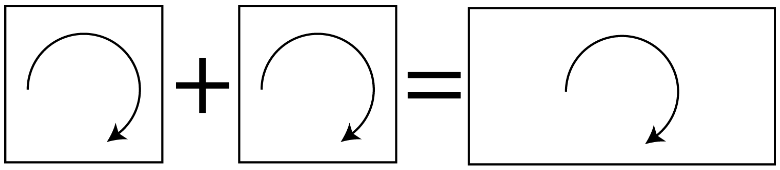


say $R_1$ and $R_2$, the integral over the enclosing rectangle $R_3$ is equal to the sum of the line integrals over the individual rectangles because the common side cancels out, assuming the orientation of each individual small square is the same as the orientation of the enclosing curve.

$$\int_{R_3} \mathbf{F} \cdot d\mathbf{r} = \int_{R_1} \mathbf{F} \cdot d\mathbf{r} + \int_{R_2} \mathbf{F} \cdot d\mathbf{r}$$

The integral over the shared path cancels out because the two times the path is traversed it is traversed in different directions - hence the two integrals differ by a minus sign factor. This cancellation repeats itself as we add additional squares,

Figure 23.2: The integral over two adjacent rectangles is the sum over the individual rectangles because the integrals over the common boundary cancels out.



eventually filling out the entire original closed curve. The only edges that do not cancel are the ones belonging to the original curve.

$$\oint_C \mathbf{F}\cdot d\mathbf{r} = \sum_{i=1}^{n} \oint_{R_i} \mathbf{F}\cdot d\mathbf{r} \tag{23.2}$$

Let us consider the path integral over a single rectangle as illustrated in figure 23.3

Figure 23.3: The path around a single tiny rectangle.

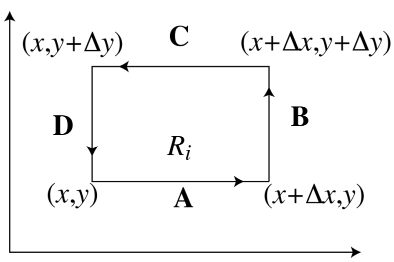



Figure 23.4: 17-4,5,6,7Fig6.pict about here.

$$\oint_{R_i} \mathbf{F}\cdot d\mathbf{r} = \int_A \mathbf{F}\cdot d\mathbf{r} + \int_B \mathbf{F}\cdot d\mathbf{r} + \int_C \mathbf{F}\cdot d\mathbf{r} + \int_D \mathbf{F}\cdot d\mathbf{r} \tag{23.3}$$

Writing the components of $\mathbf{F}$ as

$$\mathbf{F} = (M(x,y), N(x,y)) \tag{23.4}$$

Then

$$\int_A \mathbf{F}\cdot d\mathbf{r} = \int_x^{x+\Delta x} M(u,y)du \tag{23.5}$$

$$\int_B \mathbf{F}\cdot d\mathbf{r} = \int_y^{y+\Delta y} N(x+\Delta x, v)dv \tag{23.6}$$

$$\int_C \mathbf{F}\cdot d\mathbf{r} = \int_{x+\Delta x}^{x} M(u, y+\Delta y)du \tag{23.7}$$

$$\int_D \mathbf{F}\cdot d\mathbf{r} = \int_{y+\Delta y}^{y} N(x,v)dv \tag{23.8}$$

Therefore

$$\begin{aligned}
\oint_{R_i} \mathbf{F}\cdot d\mathbf{r} &= \int_x^{x+\Delta x} M(u,y)du + \int_y^{y+\Delta y} N(x+\Delta x, v)dv \\
&\quad + \int_{x+\Delta x}^{x} M(u, y+\Delta y)du + \int_{y+\Delta y}^{y} N(x,v)dv \\
&= \int_x^{x+\Delta x} \left[M(u,y) - M(u,y+\Delta y)\right] du + \\
&\quad \int_y^{y+\Delta y} \left[N(x+\Delta x, v) - N(x,v)\right] dv \\
&= \int_x^{x+\Delta x} \frac{M(u,y)-M(u,y+\Delta y)}{\Delta y}\Delta y\, du + \\
&\quad \int_y^{y+\Delta y} \frac{N(x+\Delta x,v)-N(x,v)}{\Delta x}\Delta x\, dv
\end{aligned}$$

Hence

$$\begin{aligned}
\lim_{\Delta x, \Delta y \to 0} \oint_{R_i} \mathbf{F}\cdot d\mathbf{r} &= -\int_x^{x+dx} \frac{\partial M}{\partial y} dydu + \int_y^{y+dy} \frac{\partial N}{\partial x} dxdv \\
&= -\frac{\partial M}{\partial y} dy \int_x^{x+dx} du + \frac{\partial N}{\partial x} dx \int_y^{y+dy} dv \\
&= -\frac{\partial M}{\partial y} dydx + \frac{\partial N}{\partial x} dxdy \\
&= \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) dxdy
\end{aligned}$$

Hence from equation 23.2

$$\begin{aligned}
\oint_C \mathbf{F}\cdot d\mathbf{r} &= \lim_{n\to\infty} \sum_{i=1}^{n} \oint_{R_i} \mathbf{F}\cdot d\mathbf{r} \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) dxdy \\
&= \iint_R \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) dA \quad \blacksquare
\end{aligned}$$

**Example 23.1** *Let $C$ be a circle of radius $a$ centered at the origin. Find $\oint_C \mathbf{F}\cdot d\mathbf{r}$ for $F = (-y, x)$ using Green's theorem.*

*Solution.* We have $M = -y, N = x$. Hence

$$M_y = -1, \quad N_x = -1$$

From Green's theorem

$$\oint_C \mathbf{F}\cdot d\mathbf{r} = \iint_R (N_x - M_y)dA = 2\iint_R dA = 2\pi a^2 \quad \blacksquare$$

**Example 23.2** *Let $C$ be a circle of radius $a$ centered at the origin. Find $\oint_C \mathbf{F}\cdot d\mathbf{r}$ for $F = (-x^2y, xy^2)$ using Green's theorem.*

*Solution.* Since $M = -x^2y$ and $N = xy^2$, we have

$$M_y = -x^2, \;\; N_x = y^2$$

From Green's theorem

$$\oint_C \mathbf{F}\cdot d\mathbf{r} = \iint\limits_R (N_x - M_y)dA = \iint\limits_R (x^2+y^2)dA$$

Converting to polar coordinates

$$\oint_C \mathbf{F}\cdot d\mathbf{r} = \int_0^{2\pi}\int_0^a r^3 dr d\theta = \int_0^{2\pi} \frac{a^4}{4}d\theta = \frac{a^4\pi}{2} \;\; \blacksquare$$

**Theorem 23.2 Gauss' Theorem on a Plane** *Let $C$ be a simple closed curve enclosing a region $S \subset \mathbb{R}^2$, and let*

$$\mathbf{F} = (M(x,y), N(x,y))$$

*be a smoothly differentiable vector field on $S$. Then*

$$\boxed{\oint_C \mathbf{F}\cdot\mathbf{n}ds = \iint\limits_S \nabla\cdot\mathbf{F}dA} \tag{23.9}$$

*where $\mathbf{n}(x,y)$ is a nunit tangent vector at $(x,y)$.*

*Proof.* Suppose that $C$ is described by a parameter $t$ as Recall that we derived earlier that if $s$ is the arclength,

$$\mathbf{T} = \frac{d\mathbf{r}}{ds} = \left(\frac{dx}{ds}, \frac{dy}{ds}\right) \tag{23.10}$$

is a unit tangent vector, and an outward pointing unit normal vector is

$$\mathbf{n} = \left(\frac{dy}{ds}, -\frac{dx}{ds}\right) \tag{23.11}$$

Hence

$$\begin{aligned}\oint_C \mathbf{F}\cdot\mathbf{n}ds &= \oint_C (M(x,y), N(x,y))\cdot\left(\frac{dy}{ds}, -\frac{dx}{ds}\right)ds \\ &= \oint_C (M(x,y)dy - N(x,y)dx) \\ &= \iint\limits_R \left(\frac{\partial M}{\partial x} + \frac{\partial N}{\partial y}\right)dA \\ &= \int_R \nabla\cdot\mathbf{F}dA \;\; \blacksquare\end{aligned}$$

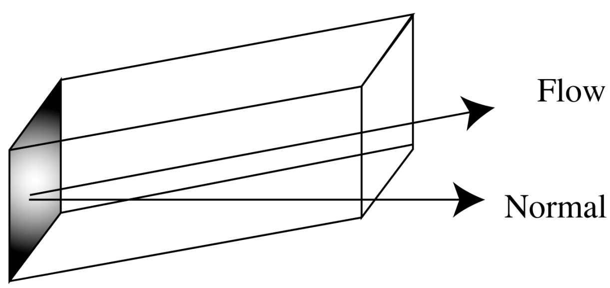
# Lecture 24

# Flux Integrals & Gauss' Divergence Theorem

A classic example of flux is the flow of water. Suppose water is flowing through a hole - say a window in a wall - of area $A$ at some velocity $\mathbf{v}$ which makes an angle $\theta$ with the window's normal vector. Then the volume of water that passes thorough the widow during a time interval $dt$ is

$$V = A\|\mathbf{v}\| \cos\theta dt$$

Figure 24.1: Flow of fluid through a window




**Definition 24.1** *Let $S$ be a surface, and $dA$ an infinitesimal surface element. Then the* **area vector** *of $S$ is*

$$d\mathbf{A} = \mathbf{n}dA$$

*where $\mathbf{n}$ is a unit normal vector to the surface element. A surface is said to be* **oriented** *if an area vector has been defined for its surface.*

Note that there are two possible orientations for any surface, because the surface unit normal can point in either direction. For a closed surface, these would correspond to output and inward pointing normal vectors.

**Definition 24.2** *The* **flow** *through an oriented surface $A$ is the volume of water that passes through that surface per unit time,*

$$\textit{flow} = A\|\mathbf{v}\|\cos\theta dt = \mathbf{v}\cdot\mathbf{A}dt$$

**Definition 24.3** *The* **flux** *through an oriented surface $A$ is the rate of flow through $A$, namely*

$$\textit{flux} = \mathbf{v}\cdot\mathbf{A}$$

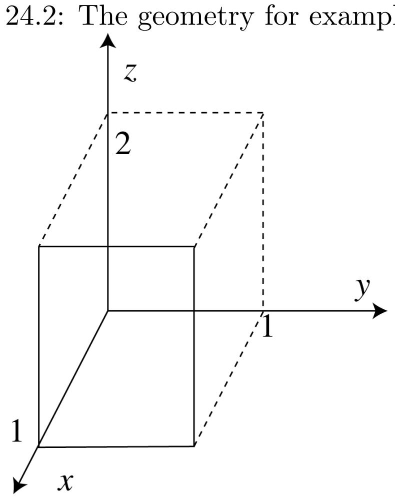
**Definition 24.4** *The* **flux of a vector field F** *through an oriented surface element $dA$ is*

$$\textit{flux} = \mathbf{F}\cdot d\mathbf{A}$$

**Definition 24.5** *The* **flux of a vector field F** *through an oriented surface $A$ or* **flux integral** *is*

$$\textit{flux} = \lim_{n\to\infty}\sum_{i=1}^{n}\mathbf{F_i}\cdot d\mathbf{A_i} = \iint_A \mathbf{F}\cdot d\mathbf{A}$$

Figure 24.2: The geometry for example 24.1.




**Example 24.1** *Find the flux integral of the vector field $\mathbf{F} = (2,3,5)$ through a rectangle parallel to the yz-plane (i.e., normal to the x-axis) with corners*

$$(1,0,0),(1,1,0),(1,1,2),(1,0,2)$$

*as illustrated in figure 24.2*

*Solution.* The dimensions of the rectangle are $1\times 2$ so the area is 2. A normal vector is

$$\mathbf{A} = (2,0,0)$$

hence the flux is

$$\text{flux} = \mathbf{A}\cdot\mathbf{F} = (2,0,0)\cdot(2,3,5) = 4 \;\blacksquare$$

**Example 24.2** *Find the flux integral of*

$$\mathbf{F} = (2, 0, 0)$$

*through a disk of radius 2 on the plane*

$$x + y + z = 2$$

*in the upward direction.*

*Solution.* A normal vector to the plane that is oriented upward is

$$\mathbf{N} = (1, 1, 1)$$

which has magnitude $\sqrt{3}$ and hence a unit normal vector is

$$\mathbf{n} = \frac{1}{\sqrt{3}}(1, 1, 1)$$

The area of the disk is $A = \pi \times 2^2 = 4\pi$. Hence an area vector is

$$\mathbf{A} = \mathbf{n}A = \frac{4\pi}{\sqrt{3}}(1, 1, 1)$$

Therefore the flux is

$$\text{flux} = \mathbf{F} \cdot \mathbf{A} = \frac{4\pi}{\sqrt{3}}(2, 0, 0) \cdot (1, 1, 1) = \frac{8\pi}{\sqrt{3}} \quad \blacksquare$$

**Example 24.3** *Find the outward flux integral of the vector field* $\mathbf{F} = (2 - x, 0, 0)$ *over the face of the cube with one corner at the origin, entirely in the first octant, with each side of length 3.*

*Solution.* Since the vector field is entirely parallel to the $x$-axis, of the six surfaces of the cube, four of them have normals that are perpendicular to the vector field. Since the dot product of these vectors with the vector field is zero, they produce no contribution to the flux. The remaining two surfaces are those perpendicular to the $x$ axis, one at $x = 3$ and the other at $x = 0$. Each of these two surfaces has an area of $A = 9$. At the $x = 3$ surface, $\mathbf{n} = (1, 0, 0)$ and $\mathbf{F} = (-1, 0, 0)$, hence

$$\text{flux} = \mathbf{F} \cdot \mathbf{n}A = (1, 0, 0) \cdot (-1, 0, 0) \times 9 = -9$$

At the $x = 0$ surface, $\mathbf{n} = (-1, 0, 0)$ and $\mathbf{F} = (2, 0, 0)$, hence

$$\text{flux} = \mathbf{F} \cdot \mathbf{n}A = (-1, 0, 0) \cdot (2, 0, 0) \times 9 = -18$$

So the total flux is $-9 - 18 = -27$. $\blacksquare$

We have previously defined the divergence as

$$\nabla\mathbf{F} = M_x + N_y + P_z$$

when $\mathbf{F} = (M, N, P)$. We can use flux integrals to give an alternate definition of the divergences, which gives a more physical description of why it is called "divergence." In fact, the limit in the following theorem was the original definition of the divergence and it was not until sometime later that the derivative formula was derived.

**Theorem 24.1** *Let* $\mathbf{F} = (M, N, P)$ *be a differentiable vector field defined within some closed volume* $V$ *with surface* $A$*. Then the divergence is given by*

$$\nabla \cdot \mathbf{F} = M_x + N_y + P_z = \lim_{V \to 0} \frac{1}{V} \iint_S \mathbf{F} \cdot d\mathbf{A} \tag{24.1}$$

**Example 24.4** *Find the divergence of the vector field* $F = (x, y, z)$*.*

*Solution: Algebraic Method*

$$\nabla \cdot \mathbf{F} = M_x + N_y + P_z = 1 + 1 + 1 = 3$$

*Solution: Geometric Method* Define a volume $V$ of dimensions $dx \times dy \times dz$ that is centered at the point $(x, y, z)$. We can define the following data for this cube.

| Side | Outward Normal | $dA$ | $d\mathbf{A} = \mathbf{n}A$ |
|---|---|---|---|
| $x + dx/2$ | (1,0,0) | $dydz$ | $(1, 0, 0)dydz$ |
| $x - dx/2$ | (-1, 0, 0) | $dydz$ | $(-1, 0, 0)dydz$ |
| $y + dy/2$ | (0, 1, 0) | $dxdz$ | $(0, 1, 0)dxdz$ |
| $y - dy/2$ | (0, -1, 0) | $dxdz$ | $(0, -1, 0)dxdz$ |
| $z + dz/2$ | (0, 0, 1) | $dxdy$ | $(0, 0, 1)dxdy$ |
| $z - dz/2$ | (0, 0, -1) | $dxdy$ | $(0, 0, -1)dxdy$ |

Assuming the center of the cube is at $(x, y, z)$ then we also have the following by evaluating the vector field at the center of each face.

| Side | $\mathbf{F}$ | $\mathbf{F} \cdot d\mathbf{A}$ |
|---|---|---|
| $x + dx/2$ | $(x + dx/2, y, z)$ | $(x + dx/2)dydz$ |
| $x - dx/2$ | $(x - dx/2, y, z)$ | $-(x - dx/2)dydz$ |
| $y + dy/2$ | $(x, y + dy/2, z)$ | $(y + dy/2)dxdz$ |
| $y - dy/2$ | $(x, y - dy/2, z)$ | $-(y - dy/2)dxdz$ |
| $z + dz/2$ | $(x, y, z + dz/2)$ | $(z + dz/2)dxdy$ |
| $z - dz/2$ | $(x, y, z - dz/2)$ | $-(z - dz/2)dxdy$ |

Adding up all the fluxes, we find that

$$\iint_S \mathbf{F} \cdot d\mathbf{A} = 3dxdydz$$

Hence the divergence is

$$\nabla \mathbf{F} = \lim_{V \to 0} \frac{\iint_s \mathbf{F} \cdot d\mathbf{A}}{V} = \lim_{V \to 0} \frac{3dxdydz}{dxdydz} = 3$$

which is the same value we obtained analytically. ■

*Proof of Theorem 24.1.* The proof is similar to the example. Consider a box centered at the point $(x + dx/2, y + dy/2, z + dz/2)$ of dimensions $V = dx \times dy \times dx$. Letting

$$\mathbf{F}(x, y, z) = \big(M(x, y, z),\ N(x, y, z),\ P(x, y, z)\big)$$

the surface integral is

$$\begin{aligned}
\iint_S \mathbf{F}\cdot d\mathbf{A} &= \mathbf{F}(x+dx, y+dy/2, z+dz/2)dydz\cdot(1,0,0)\\
&\quad+\mathbf{F}(x, y+dy/2, z+dz/2)dydz\cdot(-1,0,0)\\
&\quad+\mathbf{F}(x+dx/2, y+dy, z+dz/2)dxdz\cdot(0,1,0)\\
&\quad+\mathbf{F}(x+dx/2, y, z+dz/2)dxdz\cdot(0,-1,0)\\
&\quad+\mathbf{F}(x+dx/2, y+dy/2, z+dz)dxdy\cdot(0,0,1)\\
&\quad+\mathbf{F}(x+dz/2, y+dy/2, z)dxdy\cdot(0,0,-1)\\
&=[M(x+dx, y+dy/2, z+dz/2)-M(x, y+dy/2, z+dz/2)]dydz\\
&\quad+[N(x+dx/2, y+dy, z+dz/2)-N(x+dx/2, dy, z+dz/2)]dxdz\\
&\quad+[P(x+dx/2, y+dy/2, z+dz)-P(x+dx/2, y+dy/2, z)]dxdy
\end{aligned}$$

Hence

$$\begin{aligned}
\frac{1}{V}\iint_S \mathbf{F}\cdot d\mathbf{A} &= \frac{1}{dxdydz}\iint_S \mathbf{F}\cdot d\mathbf{A}\\
&= \frac{M(x+dx, y+dy/2, z+dz/2)-M(x, y+dy/2, z+dz/2)}{dx}\\
&\quad+\frac{N(x+dx/2, y+dy, z+dz/2)-N(x+dx/2, dy, z+dz/2)}{dy}\\
&\quad+\frac{P(x+dx/2, y+dy/2, z+dz)-P(x+dx/2, y+dy/2, z)}{dz}
\end{aligned}$$

Taking the limit as $V\to 0$,

$$\lim_{V\to 0}\frac{1}{V}\iint_S \mathbf{F}\cdot d\mathbf{A} = M_x+N_y+P_z = \nabla\cdot\mathbf{F} \quad \blacksquare$$

**Theorem 24.2 Gauss' Divergence Theorem**. *Let $\mathbf{F}=(M,N,P)$ be a vector field with $M$, $N$, and $P$ continuously differentiable on a solid $S$ whose boundary is $\partial S$. Then*

$$\boxed{\iint_{\partial S}\mathbf{F}\cdot\mathbf{n}\,dA = \iiint_S \nabla\cdot\mathbf{F}\,dV} \tag{24.2}$$

*where* **n** *is an outward pointing unit normal vector.*

*Proof.* Break the volume down into infinitesimal boxes of volume

$$\Delta V = \Delta x\times\Delta y\times\Delta z$$

Then we can expand the volume integral as a Riemann Sum

$$\iiint_V \nabla\cdot\mathbf{F}\,dV = \sum_{i=1}^{n}\Delta V_i\,\nabla\cdot\mathbf{F}_i \tag{24.3}$$

where $\mathbf{F}_i$ is the value of $\mathbf{F}$ in cube $i$. Using the geometric definition of the divergence

$$\nabla \cdot \mathbf{F}_i \approx \frac{\iint_{\partial V_i} \mathbf{F_i} \cdot d\mathbf{A}}{\Delta V_i} \tag{24.4}$$

Hence

$$\iiint\limits_V \nabla \cdot \mathbf{F}\, dV = \sum_{i=1}^{n} \Delta V_i \frac{\iint_{\partial V_i} \mathbf{F_i} \cdot d\mathbf{A}}{\Delta V_i} \tag{24.5}$$

$$= \sum_{i=1}^{n} \iint\limits_{\partial V_i} \mathbf{F_i} \cdot d\mathbf{A} \tag{24.6}$$

Since the surface integrals over bordering faces cancel out (compare with the proof of Green's theorem), all that remains of the sum is over the border of the entire volume, and we have

$$\iiint\limits_V \nabla \cdot \mathbf{F}\, dV = \iint\limits_{\partial V} \mathbf{F} \cdot d\mathbf{A} \ \blacksquare \tag{24.7}$$

**Example 24.5** *Find the surface integral $\iint_S \mathbf{F} \cdot d\mathbf{A}$ of the vector field $F = (0, y, 0)$ over a cylinder of radius 1 and height 2 centered about the z-axis whose base is in the xy plane, where S includes the entire surface (including the top and the bottom of the cylinder).*

*Solution.* Using the divergence theorem,

$$\iint\limits_S \mathbf{F} \cdot d\mathbf{A} = \iiint\limits_V \nabla \cdot \mathbf{F} dV \tag{24.8}$$

$$= \iiint\limits_V (1) dV = \iiint\limits_V dV = \pi(1)^2 \times (2) = 2\pi \ \blacksquare \tag{24.9}$$

# Lecture 25

# Stokes' Theorem

We have already seen the circulation, whose definition we recall here. We will use this definition to give a geometric definition of the **curl** of a vector field, as we did in the previous section for the divergence.

**Definition 25.1** *Let $D \subset \mathbb{R}^3$ be an open set, $C \subset D$ a path, and* **F** *a vector field on $D$. Then the* **circulation of the vector field is**

$$\textbf{circ F} = \oint_C \mathbf{F} \cdot d\mathbf{r} \tag{25.1}$$

*The* **circulation density about a vector n** *is*

$$\textbf{circ}_\mathbf{n}\ \mathbf{F} = \lim_{A \to 0} \frac{\textbf{circ F}}{A} = \lim_{A \to 0} \frac{1}{A} \oint_C \mathbf{F} \cdot d\mathbf{r} \tag{25.2}$$

*We define the* **curl of a vector field at a point** *as vector field having magnitude equal to the maximum circulation at the point and direction normal to the plane of circulation, so that*

$$(\nabla \times \mathbf{F}) \cdot \mathbf{n} = \textbf{circ}_\mathbf{n}\ \mathbf{F} \tag{25.3}$$

**Theorem 25.1** *Let $\mathbf{F} = (M, N, P)$ be a differentiable vector field. Then*

$$\textbf{curl F} = \nabla \times \mathbf{F} = (P_y - N_z, M_z - P_x, N_x - M_y) \tag{25.4}$$

**Theorem 25.2** *Stokes' Theorem. Let $D \subset \mathbb{R}^3$ be a connected set, let $S \subset D$ be a surface with boundary $\partial S$ and surface normal vector* **n**. *If $\mathbf{F} = (M, N, P)$ is a differentiable vector field on $D$ then*

$$\boxed{\oint_{\partial S} \mathbf{F} \cdot \mathbf{T}\, ds = \iint_S (\nabla \times \mathbf{F}) \cdot \mathbf{n}\, dS} \tag{25.5}$$

*where* **T** *is a unit tangent vector of $\partial S$. Here $ds$ is the distance element along $C$ and $dS$ is the surface element on $S$.*

Stokes' theorem as we have presented here is really a special case in 3 dimensions. A more general form would have $S \subset \mathbb{R}^n$ and $\partial S \subset \mathbb{R}^{n-1}$, and would be written using "generalized" surface and volume integrals as

$$\int_{\partial S} \omega = \int_S d\omega \tag{25.6}$$

where $d\omega$ represents the "exterior derivative" of a generalized vector field $\omega$ called a "differential form." We have already met several special cases to equation 25.6:

1. **In one dimension** $f : (a, b) \subset \mathbb{R} \mapsto \mathbb{R}$, $S$ is a line segment and $\partial S$ is the set of endpoints $\{a, b\}$. The derivative is $f'$, so that

$$\int_a^b f'(t)dt = f(b) - f(a) \tag{25.7}$$

   which is the fundamental theorem of calculus.

2. **In two dimensions** we can write the vector field as $F = (M, N, 0)$, $S$ is an area $A \subset \mathbb{R}^2$, $\partial S$ is the boundary $C$ of $A$, and $\mathbf{n} = (0, 0, 1)$ so that

$$(\nabla \times \mathbf{F})) \cdot \mathbf{n} = N_x - M_y \tag{25.8}$$

   Equation 25.5 then becomes

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \oint_C (Mdx + Ndy) = \iint_A (N_x - M_y)dA \tag{25.9}$$

   which is Green's theorem.

3. **In three dimensions** $S = V$, $\partial S = \partial V$, $\omega = \mathbf{F} \cdot \mathbf{n}$ and $d\omega = \nabla \cdot \mathbf{F}$, giving

$$\iint_{\partial V} \mathbf{F} \cdot \mathbf{n}\, dS = \iiint \nabla \cdot \mathbf{F}\, dV \tag{25.10}$$

   which is the divergence theorem (also called Gauss' theorem).

**Example 25.1** *Find* $\oint_C \mathbf{F} \cdot d\mathbf{r}$ *for*

$$\mathbf{F} = \left(yz^2 - y, xz^2 + x, 2xyz\right)$$

*where $C$ is a cirlce of radius 3 centered at the origin in the xy-plane, using Stoke's theorem.*

*Solution.* Letting $S$ denote the disk that is enclosed by the circle $C$ will can the formula

$$\oint_C \mathbf{F} \cdot \mathbf{T}\, ds = \iint_S (\nabla \times \mathbf{F}) \cdot \mathbf{n}\, dS$$

Since $C$, and hence $S$, both lie in the $xy$-plane we have $(n) = (0, 0, 1)$. Hence

$$\begin{aligned}(\nabla \times \mathbf{F}) \cdot \mathbf{n} &= (P_y - N_z, M_z - P_x, N_x - M_y) \cdot (0, 0, 1) \\ &= N_x - M_y \\ &= \frac{\partial}{\partial x}(xz^2 + x) - \frac{\partial}{\partial y}(yz^2 - y) \\ &= z^2 + 1 - z^2 + 1 = 2\end{aligned}$$

Hence

$$\oint_C \mathbf{F} \cdot \mathbf{T}\, ds = 2 \iint_S dS = 2\pi(3^2) = 18\pi \ \blacksquare$$

**Example 25.2** *Find* $\oint_C \mathbf{F} \cdot d\mathbf{r}$ *for*

$$\mathbf{F} = (z - 2y, 3x - 4y, z + 3y$$

*where* $C$ *is a cirlce defined by*

$$x^2 + y^2 = 4, \quad z = 1$$

*using Stoke's theorem.*

*Solution* We are again going to use the formula

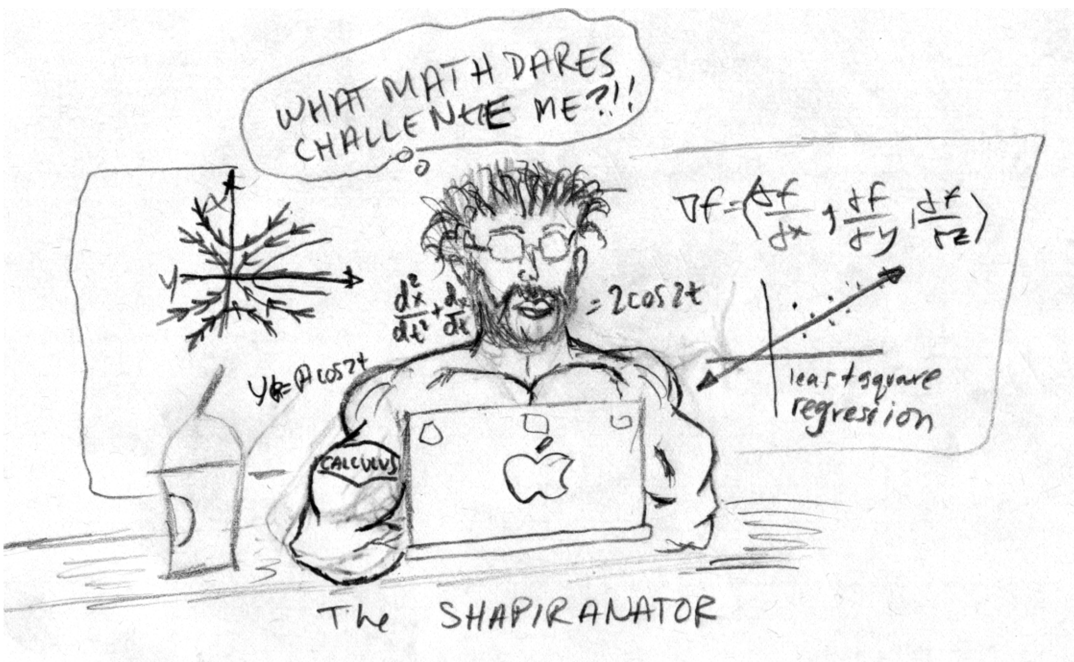
$$\oint_C \mathbf{F} \cdot \mathbf{T}\, ds = \iint_S (\nabla \times \mathbf{F}) \cdot \mathbf{n}\, dS$$

We again have $\mathbf{n} = (0, 0, 1)$, and $C$ is a circle of radius 2. Hence

$$\begin{aligned}(\nabla \times \mathbf{F}) \cdot \mathbf{n} &= (P_y - N_z, M_z - P_x, N_x - M_y) \cdot (0, 0, 1) \\ &= N_x - M_y \\ &= \frac{\partial}{\partial x}(3x - 4y) - \frac{\partial}{\partial y}(z - 2y) \\ & 3 + 2 = 5\end{aligned}$$

Hence

$$\oint_C \mathbf{F} \cdot \mathbf{T}\, ds = 5 \iint_S dS = 5\pi(1^2) = 20\pi \ \blacksquare$$


WHAT MATH DARES CHALLENGE ME?!!
CALCULUS
least square regression
The SHAPIRANATOR